学位論文

Development of a Fluorescent Probe for Visualizing Telomeric Repeat-Containing RNA in Living Cells at the Single Molecule Level

（生細胞内のテロメア RNA を一分子レベルで検出する蛍光プローブの開発）

平成 26 年 12 月博士（理学）申請

東京大学大学院理学系研究科化学専攻

山田　俊理

Development of a Fluorescent Probe for Visualizing Telomeric Repeat-Containing RNA in Living Cells at the Single Molecule Level

A dissertation submitted to the University of Tokyo for the degree of Doctor of Science

Toshimichi Yamada

Department of Chemistry, School of Science

The University of Tokyo

Acknowledgement

This thesis is completed under the supervision of Professor Takeaki Ozawa of Department of Chemistry, School of Science, The University of Tokyo. I wish to express my sincere gratitude to Prof. Ozawa for his continuous guidance, support, advice, and encouragement throughout the present study.

I would like to express my graduate to Dr. Hideaki Yoshimura for their valuable suggestions and help throughout the present work. I would like to thank all the members of Ozawa laboratory for their friendship and valuable discussion in various fields of science.

I very much thank Dr. Nobuyoshi Akimitsu and Dr. Katsutoshi Imamura, Radioisotope Center, The University of Tokyo for considerable support in my experiment technique and continuous stimulating discussion.

I would like to acknowledge the financial support by the Japan Society for the Promotion of Science.

Finally, I would like to thank my parents, and sister for their love and encouragement during all my educational years.

Abstract

Mammalian cells have two types of RNA; messenger RNA (mRNA) and non-coding RNA (ncRNA). mRNA is considered as a genetic transporter which is transcribed from a gene in nucleus and is transported to cytoplasm for producing its protein. Recent RNA imaging studies revealed that the subcellular distribution of mRNA is asymmetry, resulting to local gene expression. However, there is no ideal method to study the dynamics of mRNA, especially at the single molecule level, in living cells. Much less is known about ncRNA. Genome-wide analysis of RNA found that more than half of the transcripts are ncRNA. Among ncRNAs, particular attention has focused on the transcripts referred to as long non-coding RNAs (lncRNAs), operationally defined as being longer than 200 nucleotides. Although, numerous studies showed that lncRNA has a critical role in gene expression, little is known about the mechanism of lncRNA.

In my PhD thesis, I develop a fluorescent probe for live cell imaging of endogenous RNA, applicable both mRNA and lncRNA, with single-molecule sensitivity. By using the probe, I investigated spatiotemporal information of β-actin mRNA and telomeric repeat-containing RNA (TERRA). The probe is designed to specifically target RNA. A domain of an RNA-binding protein, PUM-HD, was mutated to recognize the target sequence in RNA. Each split fragment of enhanced green fluorescent protein (EGFP) is conjugated to amino and carboxyl terminal of mutant PUM-HD. When the two probe molecules come close on the repeat region, EGFP is reconstituted between the adjacent probes.

I applied this probe to β-actin mRNA which produced cytoskeletal actin protein and a long non-coding RNA transcribed from a telomere, telomeric

repeat-containing RNA (TERRA). Single-molecule imaging of β-actin mRNA revealed the localization of β-actin mRNA at the leading edge upon growth stimuli and the direct movement of β-actin mRNA along microtubules. Moreover, live-cell imaging of TERRA at the single-molecule level showed accumulation of TERRA around a telomere, a chromosome end which transcribes TERRA. By visualizing heterogenous nuclear ribonucleoprotein A1 (hnRNPA1), I found that TERRA localized in a telomere-neighboring region trapped diffusive hnRNPA1, thereby inhibiting hnRNPA1 localization to the telomere. Based on the single-particle analysis, I propose a mechanistic model how TERRA functions as a scaffold to hold hnRNPA1 around a telomere, inhibiting the localization of hnRNPA1 to the telomere.

Chapter 1

General Introduction

1-1. Introduction

The Human Genome Project discovered that less than 21,000 protein-coding genes span occupy only 1.2% of the 3 billion DNA bases of the human genome [1]. After the Human Genome project, The Encyclopedia Of DNA Elements (ENCODE) project was initiated to provide a comprehensive picture of all functional elements within human genome by RNA deep-sequencing (RNA-seq) [2, 3]. Two important findings are that at least 75% of the genome is transcribed and that the most of the transcripts do not code proteins, named non-coding RNAs. With the rapid advancement of deep-sequencing and bioinformatics technologies, we are now revealing new RNA species, a functional diversity, as well as RNA-protein interactions [4, 5]. As the results, RNA was found to be a highly versatile molecule, such as coding proteins, serving as structural scaffold, performing enzymatic functions, and regulating gene expression. To perform these multiple functions, RNA production, maturation, interactions, and spatial distribution require tight regulation. However, molecules are fundamentally stochastic in nature, leading to diverse, spatiotemporally inhomogeneous distribution.

Live cell imaging of RNAs at the single molecule level will reveal how individual RNAs get functions by summing the single molecule events. For example, ensemble experiments have shown that β-actin mRNA, which produces the cytoskeletal protein, actin, localizes to the leading edge of the cells [6, 7]. Single molecule imaging of the mRNA found that cytoplasmic distribution of the β-actin mRNA is a cumulative result of individual mRNAs either remaining stationary or undergoing (1) Brownian diffusion, (2) corralled motions, (3) active

transports [8, 9]. Despite the fact that 10-20% of the all β-actin mRNA exhibit biased motion, interrupting active transport significantly alters intracellular distribution [8].

In this chapter, I discuss advantages and disadvantages of the fluorescent probes for imaging RNA in living cells. To overcome the defects of the previous methods, I will propose a new strategy to visualize RNA, which is the main theme for my PhD study. Moreover, I surmise instruments for single molecule imaging and quantitative analysis for single molecule dynamics.

1-2. Methods to visualize RNA in living cells

Conventional techniques for analyzing spatiotemporal information of RNA are *in situ* hybridization in fixed cells and the biochemical fractionation of cells. These methods provide a static picture at the time of fixation or fractionation. In contrast, live-cell imaging techniques provide a temporal and spatial resolution of RNA dynamics. There is a difficulty to label RNA. In the case of visualizing proteins in living cells, a fusion molecule of a fluorescent protein and a target protein is constructed by gene engineering. On the other hand, the fusion molecule of RNA and a fluorophore cannot be generated by gene engineering.

To achieve specific labeling of RNA, an approach using nucleic acid based probe, so called ‘molecular beacon’ has been adapted. The beacons is a DNA probe in the form of a small hairpin loop labeled at the free ends with fluorescent dye and a molecule that quenches the fluorescence emission, respectively. When the DNA probe is bound to the target sequence, the probe

unloops and the quencher is kept away from the dye, leading to fluorescence emission [10, 11] (**Figure 1-1**). The major limitation of the molecular beacon is that molecular beacons cannot freely across plasma membrane. Microinjection is often used to deliver molecular beacons inside cells, but it causes damage to the cells. Even if molecular beacons are introduced into cells, there are several limitations inside the cells. Free molecular beacons enter the nucleus through nuclear pore, perhaps passively, and they are non-specifically captured by nucleic acid binding proteins that are abundant in the nucleus [12]. Moreover, oligodeoxynucleotides with natural backbones are degraded rapidly inside the cell. The degradation products are cleared from the cells with a half-time of 15-30 min [13]. Furthermore, hybrids formed by oligodeoxynucleotides and RNAs are targets for RNase H, a ubiquitous enzyme in the cell that degrades mRNAs in RNA-DNA hybrids [14].

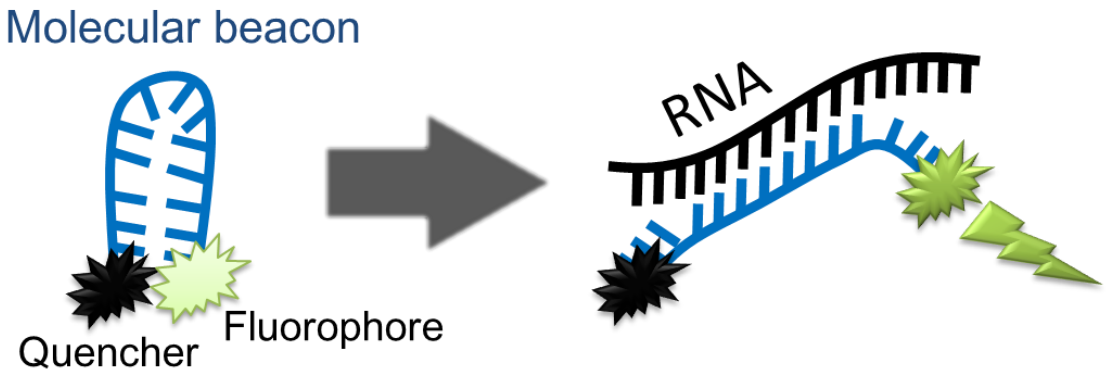

Figure 1-1. Schematic representation of molecular beacon for detecting RNA Donor (green) and quencher (gray) dyes are attached to the probes and interact with RNA as indicated.

Another approach for tagging RNA is the use of an RNA-binding protein fused to a Green fluorescent protein (GFP). The coat protein of bacteriophage MS2, which binds to a unique hairpin in the genomic RNA of the phage with a strong affinity (dissociation constant, 5 nM), was applied to tag RNAs. Two components are necessary for labeling: (1) expression of the target mRNA modified to include the MS2 stem-loop motifs; (2) the MS2 RNA coat protein fused to GFP (**Figure 1-2**). The MS2 system was pioneered by Singer and co-workers in yeast to track the movement of *ASH 1* mRNA [15].

Although the MS2 system has been used widely, this system requires expression of exogenous RNAs including the target sequence of MS2. Engineered gene could be expressed in abnormal amounts. Moreover, binding of many MS2 coat protein-GFP molecules to target RNAs may considerably impact behavior. Another difficulty is that the MS2 coat protein-GFP itself aggregates and fluorescent signals are obtained in cells without the RNA being expressed.

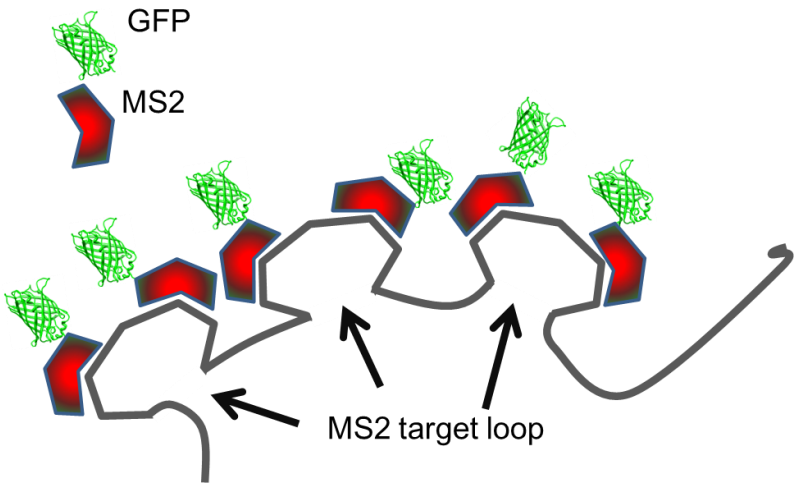

Figure 1-2. Schematic model of the MS2 coat protein-GFP method.
A tandem array of MS2 coat protein-binding sequence is inserted into untranslated region of a target mRNA. MS2 coat protein-GFP molecules bind to each motif and render that RNA more fluorescent than the surrounding.

1-3. Strategy to visualize endogenous RNA at the single molecule level

An approach using split EGFP reconstitution combined with RNA binding protein which recognizes specific RNA sequence [16] has strong advantage to visualize endogenous RNA at the single molecule level (**Figure 1-3**). In this section, I describe a basic concept of the method for visualizing endogenous RNA and also how to achieve single sensitivity in the system.

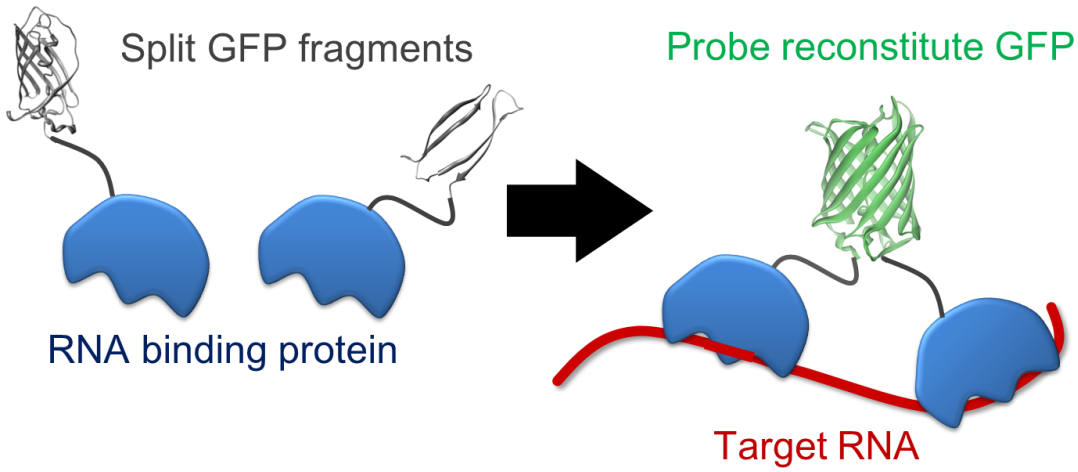

Figure 1-3. Principle to visualize RNA by using RNA binding proteins and split EGFP fragments.

The probe consists of two parts. One is the RNA binding domain which binds to specific RNA sequence. And the other is the split EGFP fragments. Without a target RNA, split EGFP fragment does not emit fluorescence. Upon binding to the target RNA, the probe reconstitutes EGFP.

1-3-1. Advantage of using PUM-HD to recognize endogenous RNA

PUMILIO proteins (PUM) are developmental regulators that control mRNA stability and translation by binding sequences in the 3’ untranslated regions of their target mRNAs. *Drosophila melanogaster* Pumilio (DmPUM) was reported first for regulating abdominal development by repressing the translation of maternal hunchback mRNA [17, 18]. DmPUM represses the target gene expression by binding tandem sequence motifs, Nanos response elements (NREs).

PUM family proteins contain RNA-binding domains, known as the Pumilio homology domain (PUM-HD). The PUM-HD comprises eight sequence repeats, called PUM repeats [19]. Crystal structure of PUM-HD revealed eight α-helical structural modules packed together into a curved shape [20]. Crystal structures of PUM-HD in complex with NRE which includes 5’-UGUAUAUA-3’ sequence showed that PUM proteins contain a domain that generally composed of eight 36-amino-acid repeats; each repeat binds a single nucleotide in its RNA-base target (**Figure 1-4 a**). Amino acids at positions 12 and 16 of the PUM repeat bind each RNA base via hydrogen bonding or van der Walls contacts with a Watson-Crick edge, whereas the amino acid at position 13 makes a stacking interaction with the aromatic rings of the RNA base (**Table 1**). The recognition of RNA is a base specific, such that cysteine and glutamine bind to adenine, asparagine and glutamine bind to uracil, and serine and glutamate bind to guanine (**Figure 1-4 b**). Guided by this recognition code, the specificity of individual repeats can be switched by mutating the two amino acids that make contacts with Watson-Crick edge of the base [21, 22]. Moreover, Filipovska *et. al*

found that a domain which has an arginine at position 16 and glycine/serine at position 12 specifically recognize cytosine [23]. Owing to these comprehensions, PUM-HD can be designed to bind any RNA of interest. Engineered PUM proteins would be tools to manipulate any endogenous RNAs with many potential applications in biotechnology and medicine [24, 25].

Table 1. The amino acids in each repeat interact with a RNA base.

The amino acids interacting with RNA bases are shown. In each repeat of wild type (wt) PUM-HD, three amino acids interact with the specific target RNA base.

PUM Repeat	R8	R7	R6	R5	R4	R3	R2	R1
wt PUM-HD	Q1126	E1083	Q1047	Q1031	Q975	Q939	Q903	Q867
	Y1123	N1080	Y1044	Q1011	H972	R936	Y900	R864
	N1122	S1079	N1043	C1007		C935	N899	S863
RNA	5' U	G	U	A	U/C	A	U	A 3'

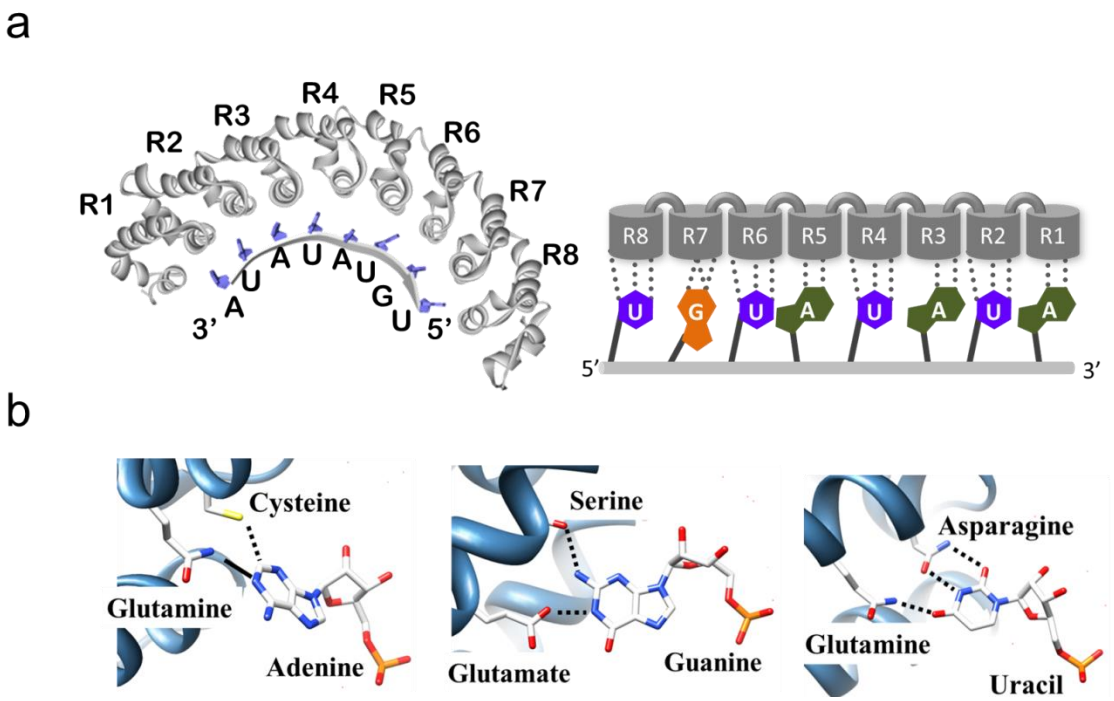

Figure 1-4. **Crystal structure of PUM-HD bound to NRE RNA.**

(a) A crystal structure of the PUM domain bound to the RNA (PDB: 1M8Y, left). Schematic of the recognition of RNA bases in NRE RNA by the PUM-HD repeats (right). (b) Recognition of Adenine, Uracil and Guanine residues. (Left) Interaction of repeat 5 with adenine. Cys-1007 makes a hydrogen bond with the adenine base and Gln 1011 contacts with the adenine base by van der Waals interaction. (Middle) Interaction of repeat 6 with uracil. Asn-1044 and Gln-1047 make hydrogen bonds with the uracil base. (Right) Interaction of repeat 7 with guanine. Ser-1079 and Glu-1083 make hydrogen bonds with the guanine base. Hydrogen bonds are indicated with dotted lines and a van der Waals interaction indicated with a black line.

1-3-2. Reduction of background signals by EGFP reconstitution technique

Protein reconstitution is a reaction in which the proteins divided in two fragments recover the function by approaching each other. In the case of EGFP, a split fragment of EGFP recovers its fluorescence by protein reconstitution when the split fragment contacts the counterpart [26]. Split fragments do not spontaneously interact with one another, but interaction partners fused to the fragments lead an interaction between the split EGFP fragments. The mechanism of protein reconstitution of EGFP has been investigated by *in vitro* studies [27]. These studies revealed that split EGFP fragments are completely folded during the association of the fragments. On the other hand, the formation of chromophore proceeds after the EGFP is folded (**Figure 1-5**). The spectral characteristics of the fluorescent complex and the intact fluorescent protein are indistinguishable, indicating the folding structure and the chromophore of the reconstituted EGFP are likely to be identical to those of intact EGFP [27].

Owing to these features of reconstitution of EGFP, many biologists have been applied this system for identification and characterization of protein-protein interactions in living cells [28]. In my strategy, reconstitution of EGFP is used for selective detection of a target RNAs and for reducing background noise. When EGFP is used as a probe of biomolecules, fluorescence signal from free diffusing EGFP is background noise. Using EGFP reconstitution technique, a probe emits fluorescence signals only when the probe binds to its target molecule. Moreover, a probe applying EGFP reconstitution technique needs two binding factors to

detect its target molecule. The use of two binding factors provide selectivity to the probe.

.

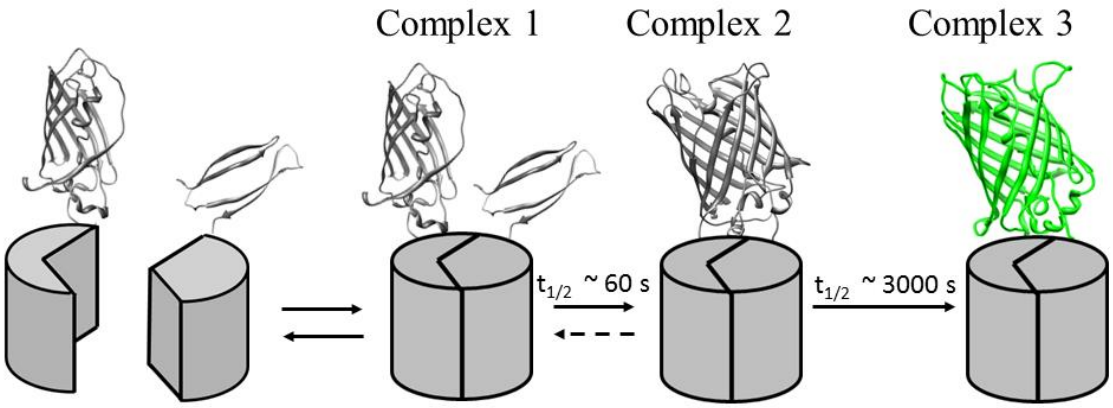

Figure 1-5 | Pathway for bimolecular fluorescent complex formation Initial associations between fusion proteins (complex 1) are mediated by the interaction partners (indicated by gray columns). The association between the fluorescent-protein fragments produces an intermediate (complex 2), which undergoes slow maturation to produce the fluorophore (complex 3).

1-4. Instrumentation for visualizing single molecule

In addition to the labeling strategy, the choice of excitation source, illumination scheme, and detector contribute toward single molecule sensitivity (**Figure 1-6. a**).

1-4-1. Optical apparatus for single molecule imaging

Since lasers emit monochromatic, coherent and collimated light, I basically use lasers for exciting fluorescent probes [29]. The choice of illumination wavelength depends on the spectral properties of the fluorescent probes, but typically spans the visual part of the electromagnetic spectrum (such as 405, 488, 561, 647 nm laser). For detecting dynamics of the single molecule, the detector must have photon counting capabilities, with high-speed recording. To satisfy the conditions, electron multiplying charged coupled devices (EMCCDs) or complementary metal oxide semiconductor (CMOS) has been adapted. Objective lens transmits light from the illumination source and collects fluorescence signal to create a focused image on the detector. For the single molecule imaging, numerical aperture (NA), an indirect measure of photon collection ability, should be high (more than 1.4) and optical magnification must be larger than $\times 60$.

1-4-2. Illumination Geometry

Wide-field epi-illumination is widely used in fluorescence imaging. However, excitation light passes through the entire depth of the cells and out-of-focus fluorescent probes emit background noise, limiting the method to samples with molecules sparsely distributed over all three dimensions (**Figure 1-6. b**). To reduce the background signals, only a thin lamina of 200 nm at the cell surface is illuminated using total internal reflection fluorescence microscopy (TIRF) [30, 31]. TIRF is based on Snell's law, wherein light which is transmitted through a medium of high refractive index (n_i; coverglass) into one of lower refractive index (n_r; water), is totally reflected within the first medium at incident angles (θ_i) greater than the critical angle given by $\theta_c = \sin^{-1} (n_r / n_i)$. TIRF creates a thin lamina of evanescent wave within the aqueous phase, only exciting fluorescent probes within 200 nm of the coverglass-sample interface (**Figure 1-6. c**). TIRF is rarely used in RNA imaging because it only excites the basal plasma membrane, whereas most RNAs are present also throughout the cell which the evanescent wave cannot penetrate. To achieve greater penetration into the cell with low background signals, highly inclined and laminated optical sheet (HILO) [32], also termed variable angle epi-fluorescence microscopy (VAEM) [33]. In this scheme, θ_i is smaller than θ_c, so that the illumination is refracted into the cell at high inclination from the optical axis. Therefore, this geometry illuminates an angled layer within the cell and reducing illumination of molecules not in focus (**Figure 1-6. d**). To visualize RNAs in cytoplasm and the nucleus, I set up VAEM with 405, 488, 561, 647 nm lasers light and an EM-CCD (**Figure 1-6. a**).

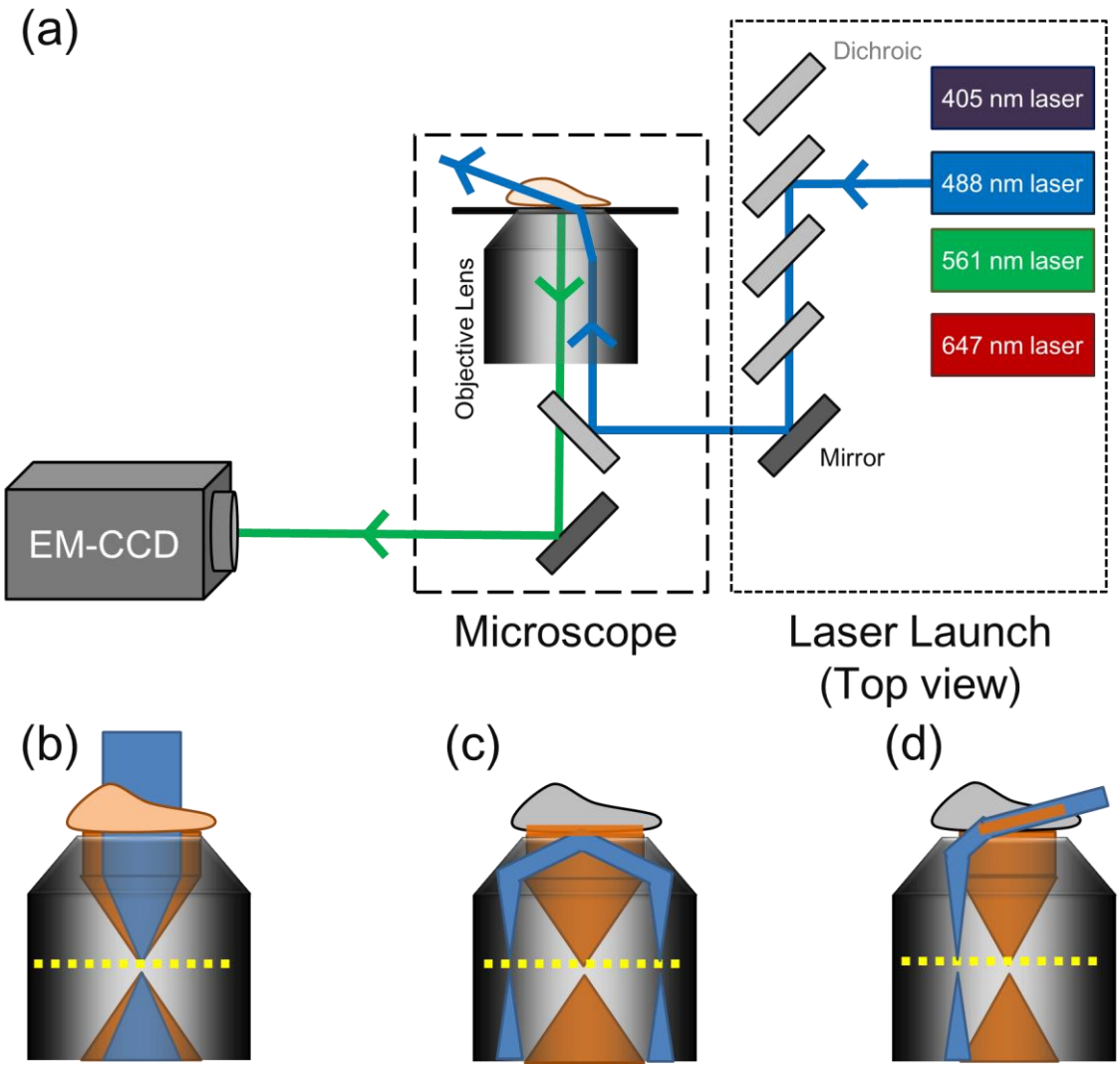

Figure 1-6. A general schematic of home-built microscopy and illumination geometries

(a) The individual laser beams are combined using long-pass dichroic mirrors (DMs). Each DM is chosen with spectral properties specific to the wavelengths that it must reflect and transmit. Here, 488 nm excitation illumination pass is shown. The laser beam are focused on the back-focal plane of the objective and aligned to excite the fluorescent probes with sub-critical, oblique angles. Fluorescence from the probes is detected by an EMCCD. (b) Wide-field, (c) TIRF, and (d) HILO/VAEM geometries are shown. Yellow broken line is a back focal plane of the objective lens. Excitation beam is illustrated in blue and fluorescence from the cells is illustrated in orange.

1-5. Analysis of single molecule dynamics

A major advantage of single molecule imaging is the ability to resolve modes of motions of individual molecules. To quantify individual motion, trajectories of the each molecule are constructed by “single particle tracking” from the fluorescence imaging data. Calculating the mean square displacement (MSD) of the trajectories clearly characterizes the modes of motions and estimates the distribution of quantities characterizing the motions, such as the diffusion coefficient, velocity, and confined size.

1-5-1. Single particle tracking

According to Abbe’s law or Raleigh’s resolution limit, diffraction limits the ability to distinguish two features located closer than half the wavelength of the fluorescence, thereby imposing a theoretical limit on the resolution of 200 nm [34]. Consequently, the image of a single fluorescent probe, a few nanometers in diameter, is spread after passing the microscope optics. The intensity distribution of such a diffraction-limited spot can be mathematically described by a point-spread function (PSF) and approximated as a simple two-dimensional Gaussian function. The center of the Gaussian curve, which coincides with the intensity maximum of the diffraction limited spot, can be localized with accuracy similar to the size of the probe, breaking the diffraction barrier [34, 35]. By connecting each center of the PSF, a trajectory of a fluorescent spot can be constructed.

1-5-2. Mean Square displacement

The mobility of single molecules reflects the mechanical properties of the microenvironment. Since single molecules are under the control of stochasticity, the relevant physical observables are averages. For this reason, the data points within a single trajectory are averaged, yielding the mean-square displacement (MSD) for the trajectory [36, 37]. The generic formula of MSD is

$$\text{MSD}\ (\tau) = \langle \Delta \boldsymbol{r}(\tau)^2 \rangle = \langle [\boldsymbol{r}(t+\tau) - \boldsymbol{r}(t)]^2 \rangle$$

where $\boldsymbol{r}(t)$ is the position of the particle at time t, and τ is the lag time between two positions. The average $\langle \ldots \rangle$ designates a time-average over t.

The MSD characterizes the dynamics of single particles into 4 modes[36], (**Figure 1-7**).

$\langle \Delta \boldsymbol{r}(\tau)^2 \rangle = 4Dt$	Normal diffusion
$\langle \Delta \boldsymbol{r}(\tau)^2 \rangle = 4Dt^{\alpha}$	Anomalous diffusion
$\langle \Delta \boldsymbol{r}(\tau)^2 \rangle = 4Dt + (Vt)^2$	Direct motion with diffusion
$\langle \Delta \boldsymbol{r}(\tau)^2 \rangle = \langle \boldsymbol{r}_c^2 \rangle [1 - A_1 \exp(-4A_2 Dt) / \langle \boldsymbol{r}_c^2 \rangle]$	confined motion

D is diffusion coefficient, V is velocity, $\langle \boldsymbol{r}_c^2 \rangle$ is the confined size, and A_1 and A_2 are constants determined by the confined geometry.

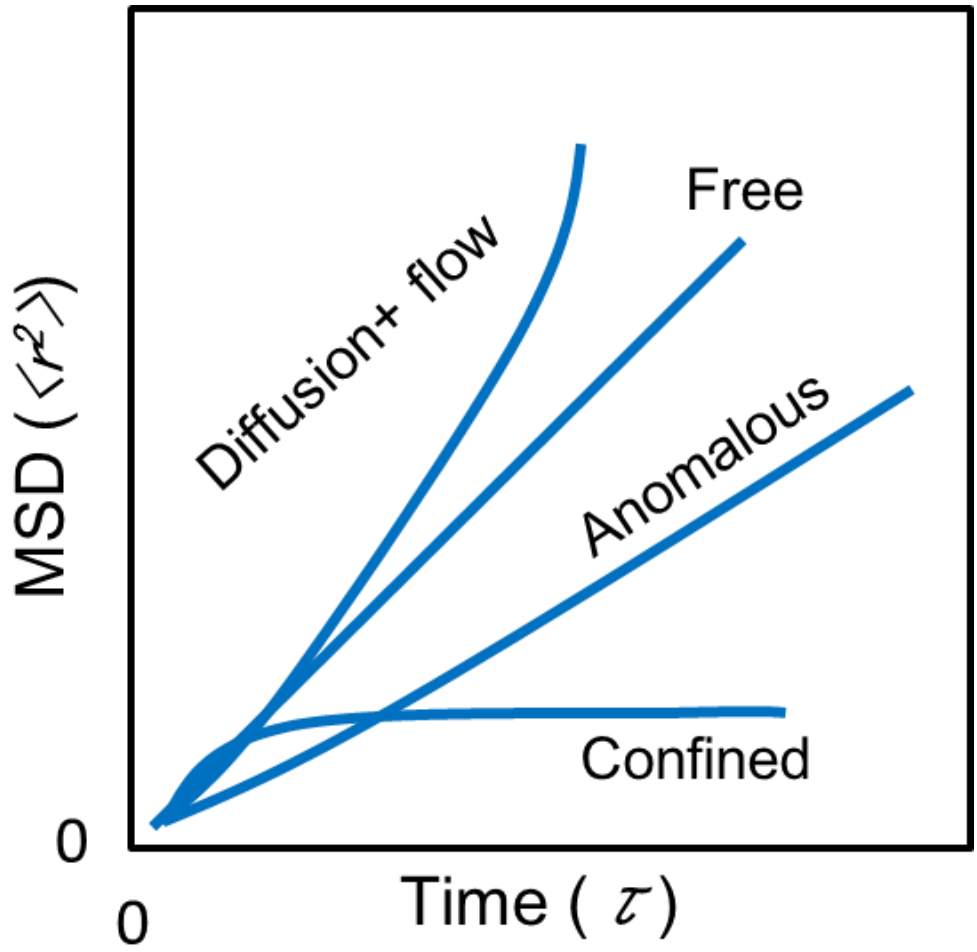

Figure 1-7. Various modes of MSD plotting

The MSD as a function of time for diffusion plus flow, free diffusion, diffusion in the presence of obstacles, and confined motion.

1-7. Conclusion

Single RNA imaging in living cells provides information related to inhomogeneous behavior, kinetic and velocity, and local heterogeneity of target RNA, of which data are lost in ensemble measurement. The probe based on sequence-specific recognition PUM-HD and EGFP reconstitution technique would have an ability to visualize a single endogenous RNA without modification in living cells. To demonstrate the applicability of the probe, I investigated spatiotemporal dynamics of β-actin mRNA in fibroblast. The distribution and movement of β-actin mRNA are already well-studied, and thus I tried to re-characterize β-actin mRNA dynamics by the developed probe; localization of β-actin mRNA at leading edge and direct movement along microtubules.

A large fraction of single molecule studies of RNAs have been used to characterize protein-coding RNAs. However, as I have highlighted in section 1, the majority of RNAs in the eukaryotic cell does not code proteins but rather regulate gene expression via post-transcriptional gene silencing or epigenetic gene regulation. In my PhD thesis, I also develop a method to visualize endogenous long non-coding RNA (lncRNA) with single molecule sensitivity based on the PUM-HD and fluorescence reconstitution technique. To demonstrate the applicability of the method, I visualized Telomeric repeat-containing RNA (TERRA) in living cells.

1-5. Reference

1 Adams, M.D., *et al.* (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651-1656

2 Consortium, E.P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74

3 Consortium, E.P. (2004) The ENCODE (ENCyclopedia of DNA elements) project. *Science* 306, 636-640

4 Tuck, A.C. and Tollervey, D. (2011) RNA in pieces. *Trends in genetics* 27, 422-432

5 Guttman, M., *et al.* (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 458, 223-227

6 Oleynikov, Y. and Singer, R.H. (2003) Real-time visualization of ZBP1 association with β-actin mRNA during transcription and localization. *Current Biology* 13, 199-207

7 Shestakova, E.A., *et al.* (2001) The physiological significance of β-actin mRNA localization in determining cell polarity and directional motility. *Proceedings of the National Academy of Sciences* 98, 7045-7050

8 Fusco, D., *et al.* (2003) Single mRNA Molecules Demonstrate Probabilistic Movement in Living Mammalian Cells. *Current Biology* 13, 161-167

9 Lifland, A.W., *et al.* (2011) Dynamics of Native β-Actin mRNA Transport in the Cytoplasm. *Traffic* 12, 1000-1011

10 Wang, W., *et al.* (2008) Imaging and characterizing influenza A virus mRNA transport in living cells. *Nucleic Acids Research* 36, 4913-4928

11 Tyagi, S. and Alsmadi, O. (2004) Imaging Native β-Actin mRNA in Motile Fibroblasts. *Biophysical Journal* 87, 4153-4162

12 Leonetti, J.P., *et al.* (1991) Intracellular distribution of microinjected antisense oligonucleotides. *Proceedings of the National Academy of Sciences* 88, 2702-2706

13 Fisher, T.L., *et al.* (1993) Intracellular disposition and metabolism of fluorescently-labled unmodified and modified oligouncleotides microijjected into mammalian cells. *Nucleic Acids Research* 21, 3857-3865

14 Walder, R.Y. and Walder, J.A. (1988) Role of RNase H in hybrid-arrested translation by antisense oligonucleotides. *Proceedings of the National Academy of Sciences* 85, 5011-5015

15 Bertrand, E., *et al.* (1998) Localization of *ASH1* mRNA Particles in Living Yeast. *Molecular cell* 2, 437-445
16 Ozawa, T., *et al.* (2007) Imaging dynamics of endogenous mitochondrial RNA in single living cells. *Nature Methods* 4, 413-419
17 Barker, D., *et al.* (1992) Pumilio is essential for function but not for distribution of the Drosophila abdominal determinant Nanos. *Genes & Development* 6, 2312-2326
18 Macdonald, P. (1992) The Drosophila pumilio gene: an unusually long transcription unit and an unusual protein. *Development* 114, 221-232
19 Zamore, P., *et al.* (1997) The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. *RNA* 3, 1421
20 Wang, X., *et al.* (2002) Modular Recognition of RNA by a Human Pumilio-Homology Domain. *Cell* 110, 501-512
21 Cheong, C.-G. and Hall, T.M.T. (2006) Engineering RNA sequence specificity of Pumilio repeats. *Proceedings of the National Academy of Sciences* 103, 13635-13639
22 Lu, G., *et al.* (2009) Understanding and engineering RNA sequence specificity of PUF proteins. *Current opinion in Structural Biology* 19, 110-115
23 Filipovska, A., *et al.* (2011) A universal code for RNA recognition by PUF proteins. *Nature Chemical Biology* 7, 425-427
24 Wang, Y., *et al.* (2013) Engineered proteins with Pumilio/fem-3 mRNA binding factor scaffold to manipulate RNA metabolism. *FEBS Journal* 280, 3755-3767
25 Mackay, J.P., *et al.* (2011) The prospects for designer single-stranded RNA-binding proteins. *Nature Structural & Molecular Biology* 18, 256-261
26 Ghosh, I., *et al.* (2000) Antiparallel Leucine Zipper-Directed Protein Reassembly: Application to the Green Fluorescent Protein. *Journal of the American Chemical Society* 122, 5658-5659
27 Kerppola, T.K. (2006) Visualization of molecular interactions by fluorescence complementation. *Nature Reviews Molecular Cell Biology* 7, 449-456
28 Ozawa, T., *et al.* (2001) Protein splicing-based reconstitution of split green fluorescent protein for monitoring protein-protein interactions in bacteria: improved sensitivity and reduced screening time. *Analytical Chemistry* 73, 5866-5874
29 Grünwald, D., *et al.* (2008) Calibrating excitation light fluxes for quantitative light microscopy in cell biology. *Nature Protocols* 3, 1809-1814

30 Axelrod, D. (1981) Cell-substrate contacts illuminated by total internal reflection fluorescence. *The Journal of Cell Biology* 89, 141-145
31 Betzig, E., *et al.* (2006) Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642-1645
32 Tokunaga, M., *et al.* (2008) Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nature Methods* 5, 159-161
33 Konopka, C.A. and Bednarek, S.Y. (2008) Variable‐angle epifluorescence microscopy: a new way to look at protein dynamics in the plant cell cortex. *The Plant Journal* 53, 186-196
34 Walter, N.G., *et al.* (2008) Do-it-yourself guide: how to use the modern single-molecule toolkit. *Nature Methods* 5, 475-489
35 Ando, R., *et al.* (2007) Highlighted generation of fluorescence signals using simultaneous two-color irradiation on Dronpa mutants. *Biophysical Journal* 92, L97-L99
36 Saxton, M.J. and Jacobson, K. (1997) Single-particle tracking: applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Sructure* 26, 373-399
37 Qian, H., *et al.* (1991) Single particle tracking. Analysis of diffusion and flow in two-dimensional systems. *Biophysical Journal* 60, 910-921

Chapter 2

Method to visualize single β–actin mRNA in living cells

2-1. Introduction

Spatially restricted protein synthesis in cells involves the positioning of mRNAs to where their proteins are required, which results in local gene expression [38-41]. This asymmetrical distribution of mRNA, termed mRNA localization, is now considered as a key step in gene expression because there are several advantages: the spatially restricted gene expression can be achieved with high temporal resolution if local stimuli regulate translations on-instead compared with the case that a signal is delivered to the nucleus to initiate transcription. Moreover, local translation should be thermodynamically efficient than transporting proteins because localized mRNAs can be translated multiple times to generate many copies of proteins. These concepts were mainly established by the RNA imaging studies. Early imaging studies of asymmetric distribution in model system, such as budding yeast [42, 43], *Drosophila* embryos [44, 45], *Xenopus* oocyte [46, 47], mammalian fibroblast [48, 49], and neuron [50-52], revealed that mRNA localization is a conserved phenomenon. Recently, it was shown that during *Drosophila* embryos development about 70% of mRNAs showed distinct spatial patterns [53]. Moreover, half of the neuronal mRNA species in the rat hippocampus are accumulated in axons and dendrites [54]. Although current imaging technology already made a great contribution to discover mRNA localization, more complete understanding of mRNA localization and dynamics demands novel techniques to visualize target mRNAs in living cells.

The ideal methodology for RNA imaging is based on a fluorescent probe that recognizes selectively an endogenous target RNA at the single molecule

level. In this chapter, I present a genetically encoded probe based on sequence-specific binding domain, PUM-HD, and EGFP reconstitution technique. To demonstrate the applicability of this probe, I visualized a β-actin mRNA in living mammalian cells. Using the probe, I showed β-actin mRNA localization in the leading edge of the cell in response to serum stimulation and direct transportation along the microtubule.

2-2. Experimental section

2-2-1. Plasmid Construction.

All cDNAs were cloned using pBlueScript (Stratagene, CA). The cDNA encoding RNA binding domains of human PUMILIO1 (PUM-HD) (828–1176 amino acids) was generated by PCR with a template of PUM-HD. One mutant of PUM-HD (C935E, Q939E), named actPUM1, and the other mutant of PUM-HD (S863N, N899S, Q903E, C935E, Q939E, C1007S, and Q1011E), named actPUM2, was generated using a mutagenic PCR technique. N-terminal EGFP (GN; 1–157 amino acids) and C-terminal EGFP (GC; 158–238 amino acids) were generated by PCR with a template of EGFP. GN and GC were respectively connected with actPUM1 and actPUM2. The cDNA fragments were subcloned into a mammalian expression vector, pcDNA3.1 (+) (Invitrogen Corp., CA). mRFP connected with α-tubulin was generated by replacing the cDNA of GFP in GFP-α-tubulin (Clontech, CA) with a cDNA of mRFP. The cDNA fragments were also subcloned into a mammalian expression vector: pcDNA3.1 (+) (Invitrogen Corp.).

2-2-2. Cell culture and transfection

NIH3T3 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM), (Gibco, CA) supplemented with 10% calf serum (Gibco) at 37°C in 5% CO_2. Before transfecting the plasmids, NIH3T3 cells were grown on 35 mm dishes with glass-coverslip bottoms (Iwaki Glass Co. Ltd., Japan) for imaging and on 10 cm dishes for RT-PCR and IP-RTPCR. The cells were transfected with

plasmids using Lipofectamine LTX (Invitrogen Corp.) according to the manufacturer's protocol.

Serum starvation was performed as follows. After transfection with the probe, the NIH3T3 cells were starved for 24 hr with DMEM containing 0.2% calf serum. Before putting on the microscope, the media were changed to Hanks' Balanced Salt Solution (HBSS; Gibco) containing MEM Amino Acids Solution (Gibco), NEAA (Gibco), 2.5 g/l glucose, 2 mM glutamine, 1 mM sodium pyruvate, and 10 mM 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid (HEPES) (pH 7.4).

To prepare NIH3T3 cells expressing mRFP-tubulin, NIH3T3 cells were transfected with mRFP-tubulin plasmid using Lipofectamine LTX. After 2 days, the cells were replaced in culture medium with 1 mg/ml G418 for 1 week to select colonies expressing mRFP-tubulin. After the selection period, visible colonies were picked and subcultured on new dishes. A cell line expressing mRFP-tubulin was chosen from the picked colonies.

2-2-3. Cell extraction for RT-PCR

NIH3T3 cells were grown on four plates of 10 cm dishes. The cells on the two dishes were transfected with pcDNA3.1 (+), whereas the other two were transfected with the vector including either GN-actPUM1 or actPUM2-GC. Two days after the transfection, the cells were washed with phosphate buffered saline (PBS) buffer (100 mM Na_2HPO_4, 20 mM KH_2PO_4, 137 mM NaCl, 27 mM KCl, pH 7.4) and collected with a scraper. From the half volume of the cells, total RNAs were extracted using the TRIzol (Invitrogen Corp.) according to the

manufacturer's protocol. The total RNAs were suspended in 100 μl of RNase-free water, from which 4 μl was used for the following reverse transcription.

2-2-4. Reverse transcription and polymerase chain reaction (RT-PCR)

Single-stranded cDNAs were prepared using oligo-p(dT) primers and the SuperScript 3 First-Strand Synthesis System according to the protocol (Invitrogen Corp.). The cDNAs were subjected to PCR using either a pair of β-actin primers (5'-AGATGACCCAGATCATGTTTGAG-3' and 5'-CAGTAATCTCCTTCTGCATCCTG-3') or a pair of GAPDH primers (5'-GATGACATCAAGAAGGTGGTGA-3' and 5'-GGTCCAGGGTTTCTTACTCCTT-3'). Each amplified PCR product was separated using agarose gel electrophoresis. Then the existence of β-actin and GAPDH was examined.

2-2-5. Immunostain and fluorescence *in situ* hybridization

NIH3T3 cells expressing GN-actPUM1 and actPUM2-GC were grown directly onto coverslips, washed once in PBS, and fixed for 30 min at room temperature in 1× PBS supplemented with 3.7% formaldehyde. After washing the cells twice with the PBS, the cells were permeabilized by treatment with 70% ethanol overnight at 4°C. The cells were hydrated for 5 min at room temperature in 2× SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0) supplemented with 50% formamide. β-actin mRNAs were hybridized with 5'TAMRA-labeled antisense oligonucleotide probe

(5’-TTAGGTTTTGTCAAAGAAAGGGTGTAAAACGCAGCTCAGTAACAG TCCGC-3’) overnight at 37°C using 40 μl of a mixture containing 10% dextran sulfate, 2 mM vanadyl-ribonucleoside complex, 0.02% RNase-free BSA, 40 μg *E. coli* tRNA, 2× SSC, 50% formamide, 30 ng of the oligonucleotide probe[42]. The cells were washed twice for 30 min in 2× SSC supplemented with 50% formamide at 37°C.

2-2-6. Acquisition of imaging data

Single fluorescence imaging of mRNA was performed using a home-built TIRF microscope constructed on an inverted microscope (IX81; Olympus Corp., Japan). The microscope was equipped with two adjustable laser lines with a 488 nm laser (CYAN-488; Spectra-Physics, CA) for exciting EGFP and a 561 nm laser (JUNO 561; SOC, Japan) for TAMRA and mRFP excitation, and with PlanApo 100× oil immersion objective with a numerical aperture of 1.49. For excitation illumination, the laser beams were placed off the optical axis (but parallel to the optical axis) so that they passed near the edge of the objective lens to make a highly inclined illumination but not total internal reflection, which illuminates a very shallow area of the sample. Emission from the fluorophore was collected by the objective and led to an EM-CCD camera (ImagEM; Hamamatsu Photonics, Japan). The microscope system was controlled with software (Aquacosmos; Hamamatsu Photonics). The single molecule images were captured with software (HC-Image; Hamamatsu Photonics).

2-3. Results

2-3-1. Design of the fluorescent probe for β-actin mRNA imaging

By introducing several amino acids mutations, I developed two mutant PUM-HDs, actPUM1, and actPUM2, which recognize specific region of β-actin mRNA (5'-UGUACGUA-3', and 5'-UGUGCUGU-3', respectively. **Figure 2-1 a**). actPUM1 and actPUM2 were then connected with N-terminal or C-terminal fragments of EGFP, named GN-mPUM2 and actPUM1-GC, respectively (**Figure 2-1 b**). In principle, the probes emit EGFP fluorescence by reconstitution of EGFP fragments upon binding to β-actin mRNA (**Figure 2-1 c**).

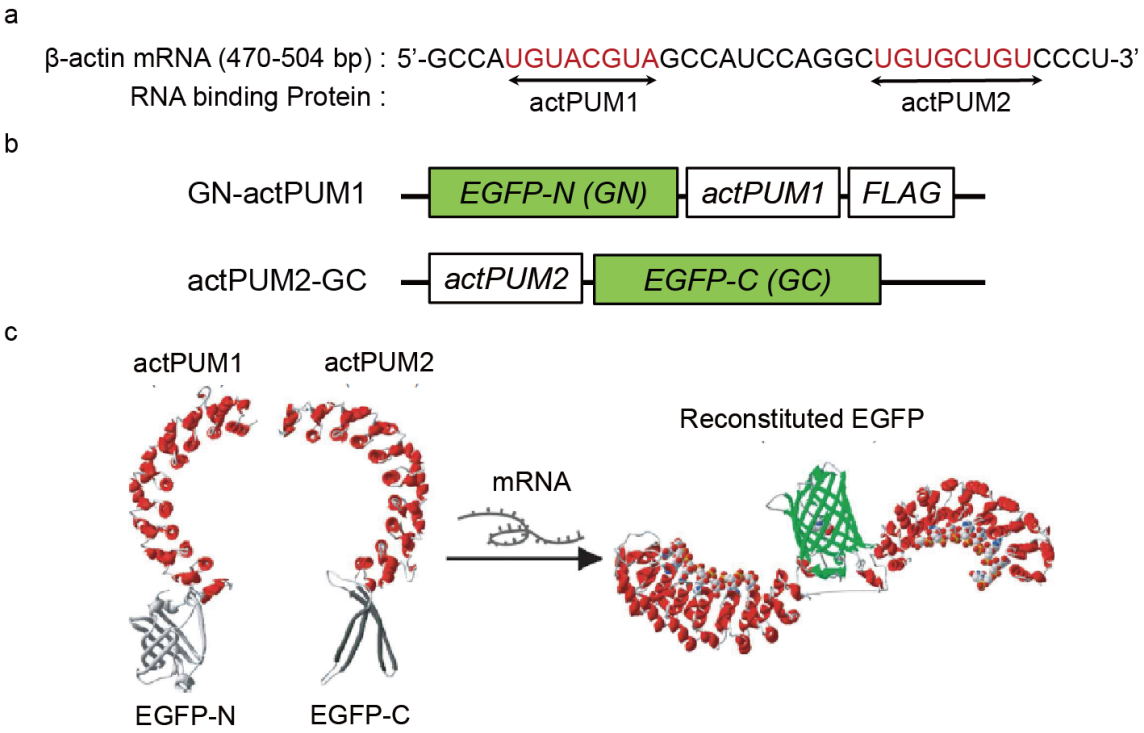

Figure 2-1 Schematic of the probes for visualizing β-actin mRNA

(a) Sequences of β-actin mRNA that are recognized by the PUM-HD mutants. (b) Schematic structures of cDNA constructs. All text in italics refers to the genes of their corresponding proteins. FLAG is an epitope tag for FLAG antibody. (c) Principle for reconstitution of EGFP on β-actin mRNA. Two RNA-binding domains of PUM-HD are engineered to recognize the target mRNA specifically (actPUM1 and actPUM2). In the presence of the target mRNA, actPUM1 and actPUM2 bind to the corresponding mRNA sequences, and the N- and C-terminal fragments of EGFP are brought together, resulting in fluorescence of complement EGFP.

2-3-2. Characterization of the probes

First, I examined binding abilities of GN-actPUM1 and actPUM2-GC to their corresponding RNA sequences by the immunoprecipitation and reverse transcription PCR (IP-RT-PCR) method. GN-actPUM1 and actPUM2-GC were expressed in NIH3T3 cells and were then collected from the cell lysate with either anti FLAG antibody for GN-actPUM1 or anti GFP antibody for actPUM2-GC. The mRNAs binding to each GN-actPUM1 and actPUM2-GC were collected and converted to cDNAs, respectively, from which existence of β-actin mRNAs was examined using PCR. Electrophoresis revealed that a single band originated from β-actin mRNA was obtained (**Figure 2-2**). The band of β-actin mRNA was only detected when the probes were expressed in the cells, indicating that each GN-actPUM1 and actPUM2-GC binds to the β-actin mRNA in the cells.

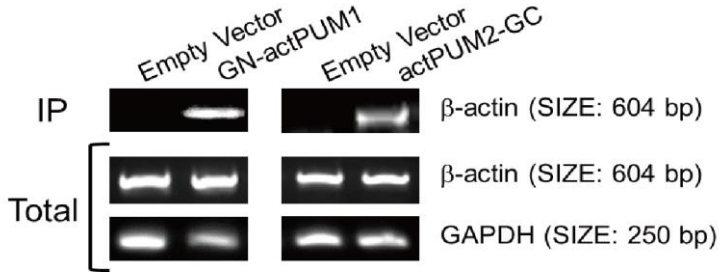

Figure 2-2 Binding ability of the probes to β-actin mRNA.

Immunoprecipitation and reverse transcription PCR. mRNAs extracted from the immunoprecipitates were subjected to RT-PCR using primers of β-actin mRNA (top). The left lane shows the mRNA from the cell transfected only pcDNA3.1 without any insertions. The right lane shows the result from the cell with the probe. The positive RT-PCR control is middle (β-actin) and bottom (GAPDH).

2-3-3. Specific reconstitution of EGFP on β-actin mRNA

Selectivity of the probe was confirmed by investigating whether reconstitution of EGFP fragments occurs upon binding to β-actin mRNA. To probe β-actin mRNAs, a TAMRA-labeled antisense oligonucleotide, which binds to the β-actin 3'untranslated regions, was used. The results of fluorescence *in situ* hybridization (FISH) with NIH3T3 cells expressing the GN-actPUM1 and actPUM2-GC are shown in Figure 2-3 a. EGFP signals were overlapped with TAMRA fluorescence. Moreover, fluorescent spots were localized in the leading edge of the lamellipodium (**Figure 2-3 b**). In the control, I conducted FISH using sense probes under otherwise identical conditions. Compared to antisense probe, much less fluorescent spots were visualized and distributed homogeneously in the cells (**Figure 2-3 c**). These results indicated that the split fragments of EGFP complemented and recovered the fluorescence signal upon binding of the probes to β-actin mRNAs in the cells.

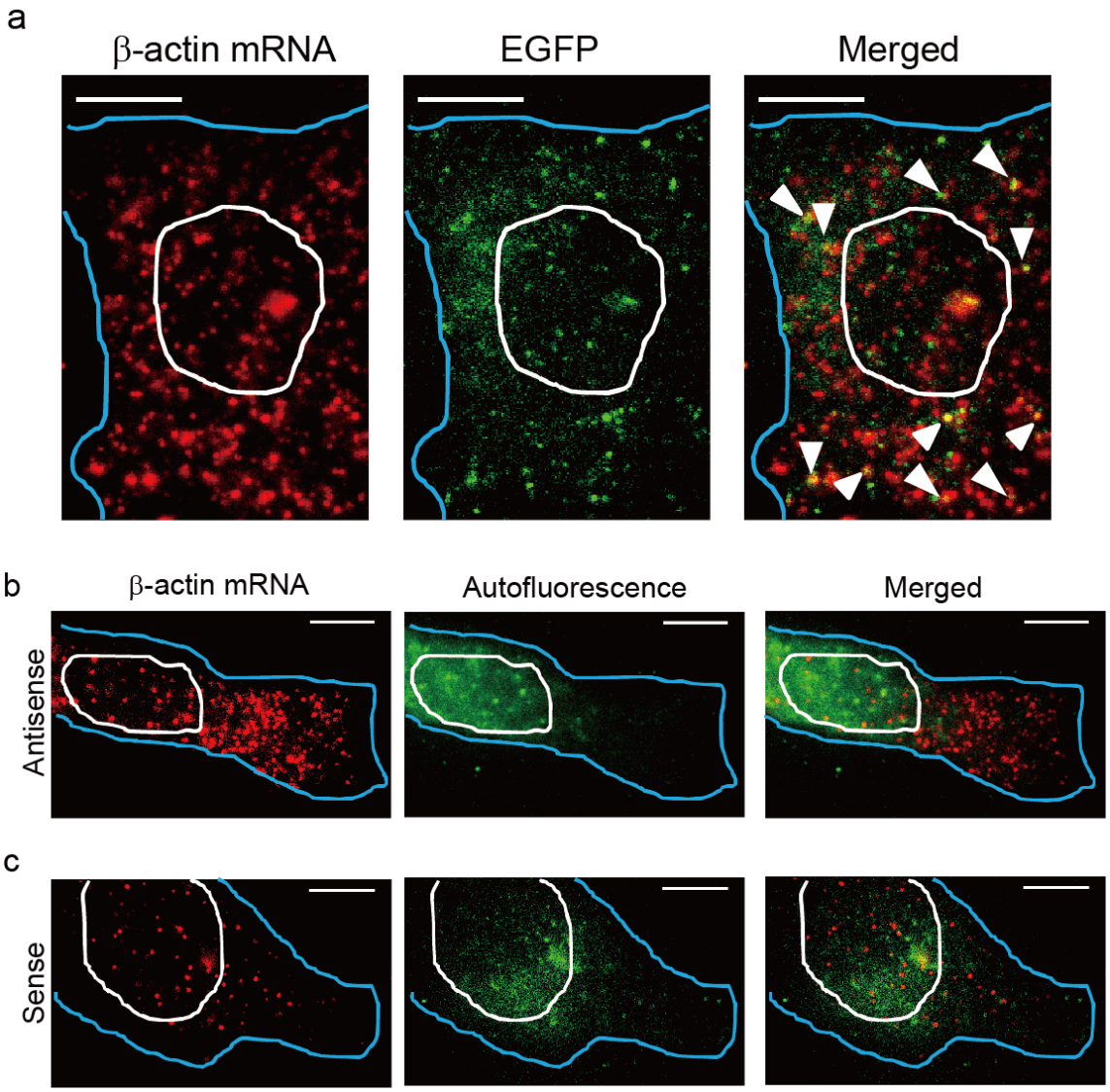

Figure 2-3 Selective binding of the probes to β-actin mRNA

(a) Localization of β-actin mRNA and the probes. β-actin mRNA was visualized with FISH using 5'TAMRA oligonucleotide probe (left) together with EGFP reconstituted from GN-actPUM1 and actPUM2-GC (middle). Red and green dots respectively indicate the localization of β-actin mRNA and reconstituted EGFP. The merged image is shown (right). Arrowheads indicate the colocalized spots of β-actin mRNAs and the reconstituted EGFP. Blue and white lines respectively show the shape of the cell and the nucleus. Scale bars, 8 μm. (b) Fluorescent in situ hybridization using oligonucleotide probe directed against the 3'UTR of β-actin mRNA (b, antisense oligo-nucleotide probe) and its control (c, sense oligo-nucleotide probe). Blue and white lines respectively show the shape of the cell and the nucleus. Scale bar, 8 μm.

2-3-4. Visualization of single β-actin mRNA

I next examined the number of EGFP molecule in a single spot (**Figure 2-4 a**). From the intensity analysis, each fluorescent spot showed small fluctuation of fluorescence intensity and single-step photobleacing occurred (**Figure 2-4 b**), which indicates that the spot consists single EGFP. Moreover, the distribution of the fluorescence intensities of the spots was fitted with a Gaussian distribution (**Figure 2-4 c**), demonstrating that the fluorescent spots consisted of the same number of the EGFP. These two results showed that the probe enabled to visualize β-actin mRNA at the single molecule level.

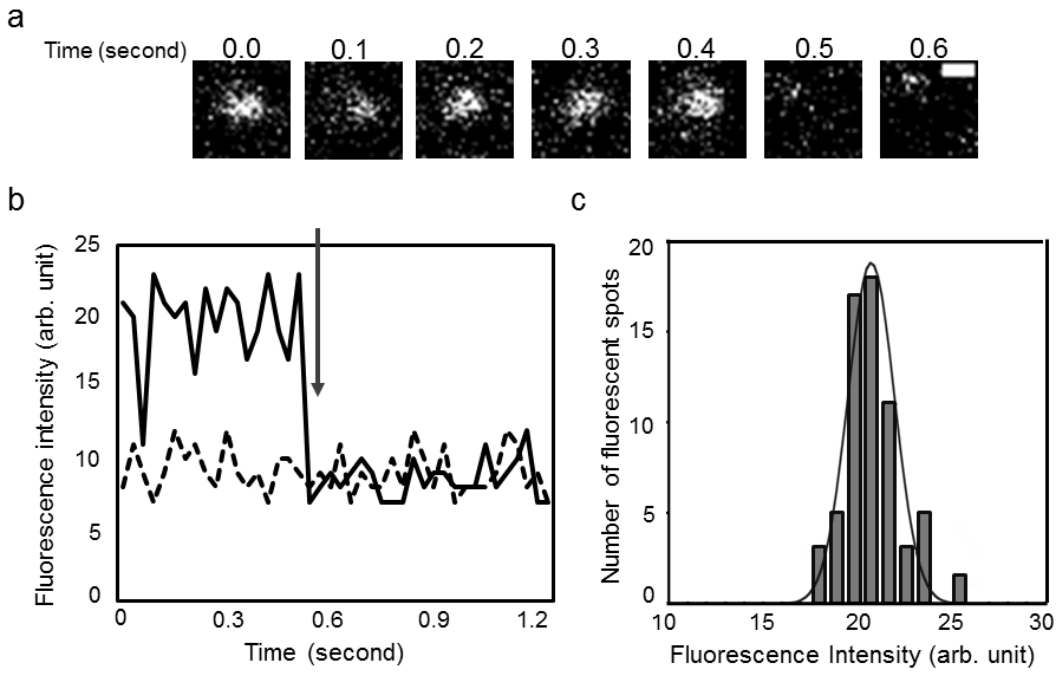

Figure 2-4 Imaging a single β-actin mRNA in living cells

(a) A typical video sequence of a fluorescent spot of EGFP. The probes of GN-actPUM1 and actPUM2-GC were co-expressed in NIH3T3 cells. Scale bar, 400 nm. (b) Time-dependent fluorescence intensity of a fluorescent spot of EGFP. A single-step photobleaching of the fluorescence spot occurred at the time indicated by the arrow. The spot and the background intensities are shown respectively as solid and broken lines. (c) Histogram showing fluorescence intensity of individual spots.

2-3-5. β-actin mRNA response to growth factor

To demonstrate the ability of the probe for investigating spatiotemporal dynamics of β-actin mRNA, NIH3T3 cells expressing the probes were cultured under serum-starved condition. Then fluorescence images of the cells were taken using fluorescence microscopy (**Figure 2-5 a**). To analyze localization of the total β-actin mRNAs labeled with the probes, I calculated the center of all fluorescent spot positions. The center of the fluorescent spots was almost identical in position to the cell centroid. This result suggested that β-actin mRNAs were equally distributed in the cells under the serum-starved condition. Then, I replaced the cell-cultured medium into a 10% serum condition and incubated the cells for 1 hr thereafter. Addition of the serum induced a large change in the location of the fluorescent spots (**Figure 2-5 b**). The center of the fluorescent spots position was separated in 5.0 ± 0.4 μm from the cell centroid (**Figure 2-5 c**). Comparison of fluorescent spots positions between 0.2% and 10% serum conditions clearly demonstrated that β-actin mRNA moved a definite direction inside the cells upon the serum stimulation (**Figure 2-5 d**). The serum effect on the localization of β-actin mRNA in this study was almost consistent with the previous report [48], demonstrating the ability of this method to analyze spatiotemporal dynamics of β-actin mRNA inside the cells in the single-molecule level.

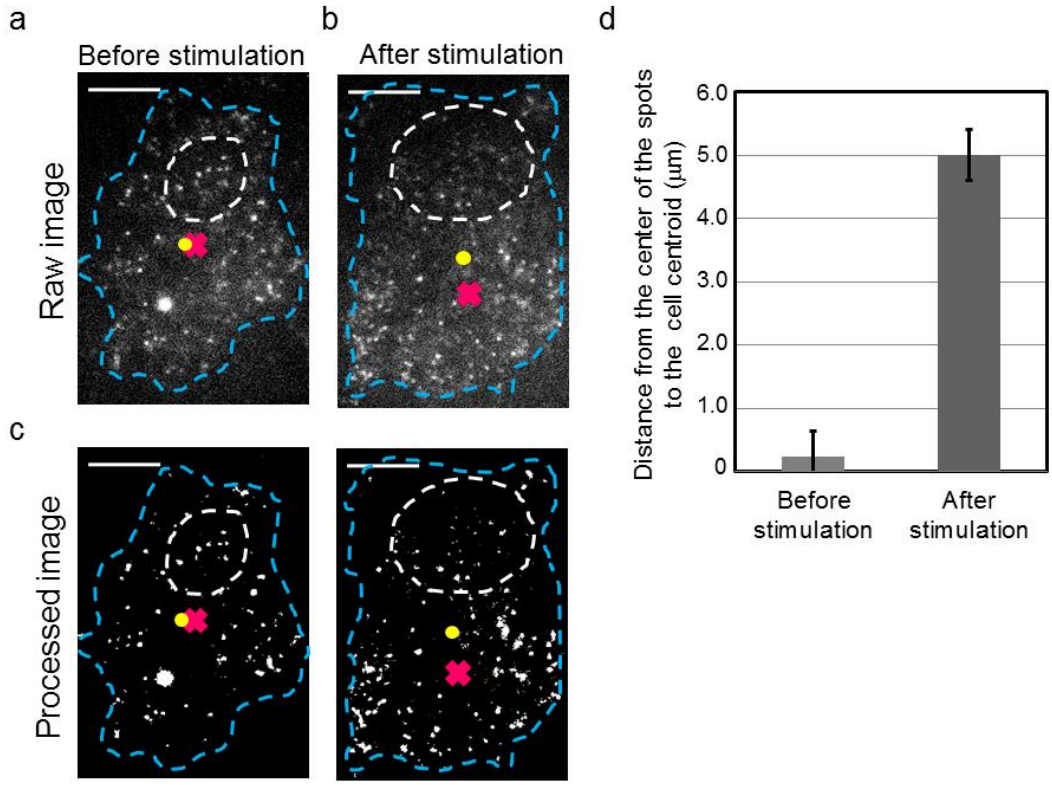

Figure 2-5 Effect of serum on β-actin mRNA localization

NIH 3T3 cells were starved under the condition of 0.2% serum for 24 hr (a) and stimulated by addition of calf serum to the final concentration of 10% (b). The images were deconvolved using the Iterative Deconvolution plug-in of Image J software (c). The yellow dot and red cross respectively represent the cell centroid and the averaged center of the spot positions. Scale bars, 8 μm. (d) Distances between the center of the spot positions and cell centroid were evaluated before and after stimulation with the serum (n=4).

2-3-6. Processive movement β-actin mRNA along microtubules

To verify that the probe visualizes natural movement of β-actin mRNA, I investigated the transportation of β-actin mRNA on microtubules. The probes were expressed in NIH3T3 cells stably expressing tubulin-mRFP. The fluorescent images at lammelipodia of the cells showed that several fluorescent spots of β-actin mRNA were localized on microtubules (**Figure 2-6 a**), and the spots moved along microtubules to the cell periphery (**Figure 2-6 b**). The average velocity was estimated to be 1.78 ± 0.78 μm/s. Moreover, β-actin mRNA movement showed that the velocity was not constant; some β-actin mRNA stayed at the same position and moved at constant velocity thereafter. Processive movement of β-actin mRNA is known to be driven by a motor protein such as kinesin. It moves to the plus end of a microtubule with the velocity of about 1.0 μm/s [55-57]. The velocity of β-actin mRNA I estimated as nearly the same as that of kinesin, indicating that the present method does not disrupt the natural movement of β-actin mRNA through a motor protein on the microtubules in living cells.

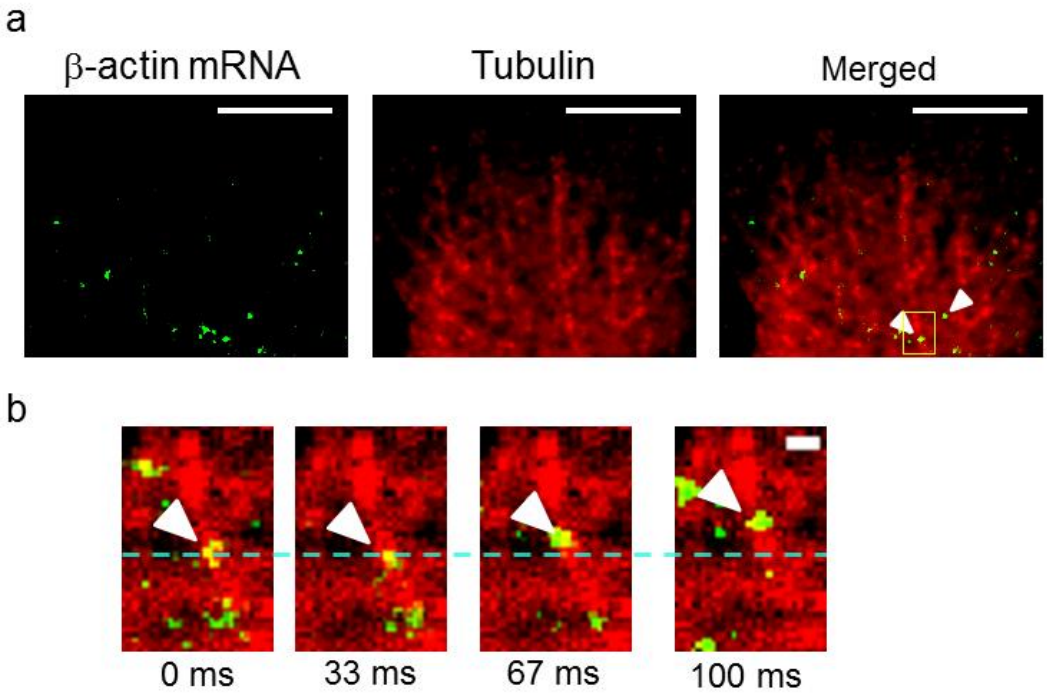

Figure 2-6 Visualization of single β-actin mRNA and microtubule

(a) Localization of β-actin mRNAs and microtubules. The probes of GN-actPUM1 and actPUM2-GC were co-expressed in NIH3T3 cells stably expressing tubulin-mRFP. Images of the RNA probes (green) were deconvolved using the Iterative Deconvolution plug-in of the Image J software. The fluorescence image of tubulin in a lamellipodium is shown as red. White arrowheads in the merged image indicate colocalization of β-actin mRNA with microtubules. Scale bar, 8 μm. (b) Time-lapse images of a single β-actin mRNA on microtubules. The broken blue line shows the start point (at 0 s) of the β-actin mRNA. Scale bar, 800 nm.

2-3-7. Alternative strategy to visualize β-actin mRNA

Although the technique of the EGFP reconstitution reduces background fluorescence and enable single-molecule imaging, there are several limitations. A major limitation is that the EGFP reconstitution takes longer than tens of minutes, which might hamper the analysis of mRNA dynamics. Another limitation is the requirement of the two plasmids for expressing the probes. A possible method to avoid such limitations is use of full-length EGFP because it produces fluorescence right after the maturation of the probe. In addition, only a single expression vector is sufficient to generate the probe.

Here, I briefly summarize an RNA probe consisting of two actPUMs and full-length EGFP. To eliminate background fluorescence, a nuclear localization signal (NLS) was connected with the N-terminus of the probe (NLS-actPUM1-EGFP-actPUM2, **Figure 2-7 a**). I investigated the localization of the probe in living COS7. Although most probes were localized in the nucleus, several spots were detected in the cytoplasm, suggesting that the excess probes that did not bind to β-actin mRNAs were transported from the cytoplasm to the nucleus (**Figure 2-7 b**). I then examined whether the β-actin mRNA probe visualizes colocalization of β-actin mRNA with microtubules. Fluorescence of COS7 cells expressing the probe and tubulin-RFP was observed by epi-fluorescence microscope. I found that EGFP spots were colocalized with tubulin-RFP specifically in the cell peripheral region (**Figure 2-7 c**), demonstrating the colocalization of β-actin mRNA with microtubules. I further revealed that some EGFP showed directional motion in the cytoplasm, of which

the velocity was 1.63 μm/s. This value was comparable to that observed by the probe based on split EGFP (1.78 ± 0.78 μm/s, **Figure 2-7 d**).

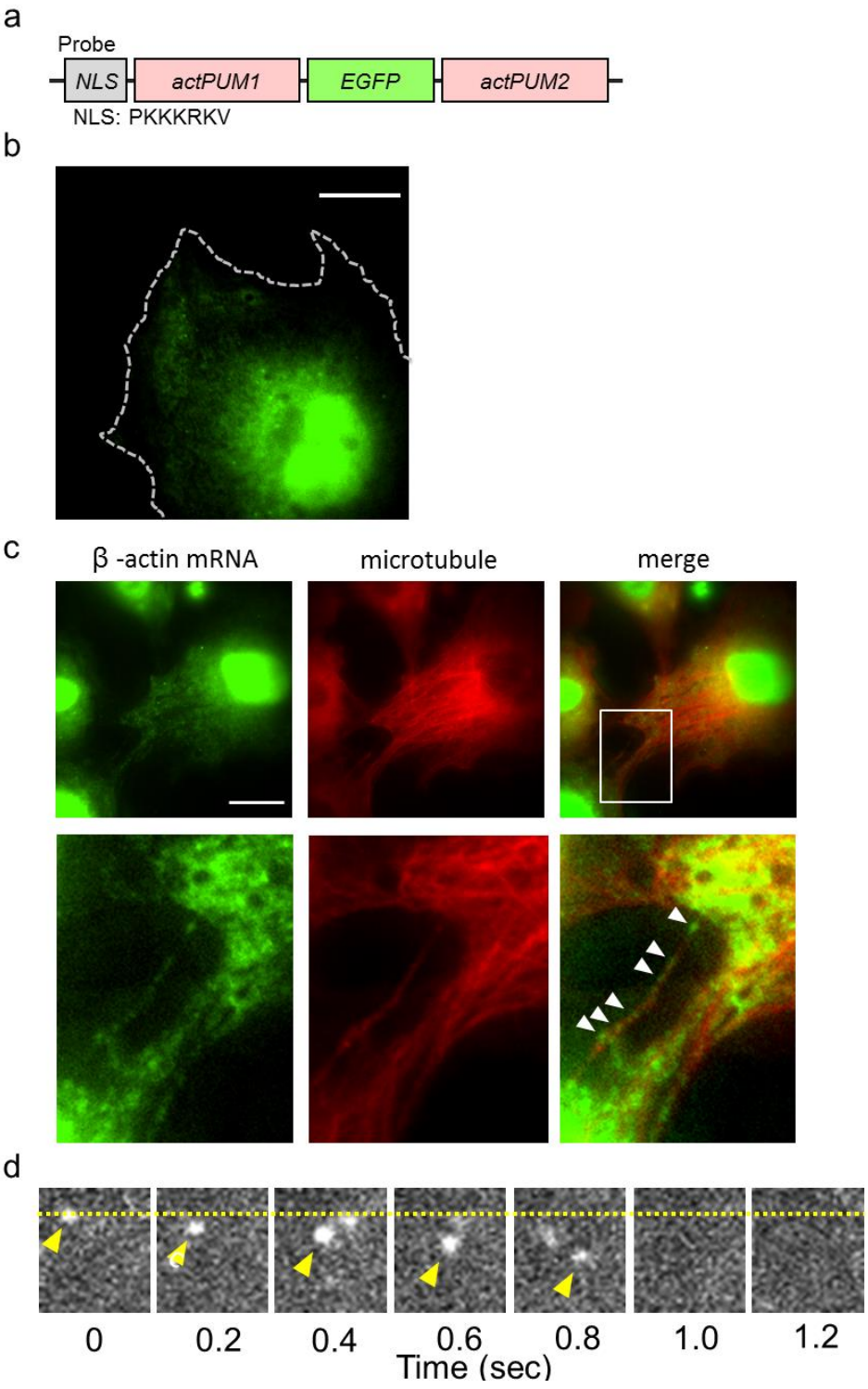

Figure 2-7 Visualization of β-actin mRNA with the probe based on full length EGFP

(a) Schematic structure of the cDNA of the probe. (b) Epifluorescence image of the COS7 expressing the probe. The white broken line represents the cell contour. Scale bar, 20 μm. (b) Epifluorescence image of the COS7 expressing the probe, mRFP-tubulin, and the merged image. Bottom figures are enlarged images in the white box shown in the upper image. Overlapping of the probe on microtubules is indicated by arrowheads. Scale bar, 20 μm.

2-5. Discussion

To establish a method for visualizing endogenous single mRNA in living cell, I developed probes which are constructed by connecting each split fragment of EGFP to RNA-binding domain, PUM-HD. This method has several advantages compared with other methods. One advantage is that a target mRNA is non-engineered and endogenous and therefore this method enables us to observe natural movement of the mRNA in live cells. In the case of MS2 system, addition of a tag sequence to a target mRNA prevents from exploring the endogenous distribution and movement of the target mRNA because the tag sequence may affect the interaction between the mRNA and RNA-binding proteins (RBPs) [58, 59]. Moreover, engineered mRNA is inserted in a random locus of cells and hence its expression level will be different from that of endogenous mRNA. Single mRNA imaging is another advantage of this method. It is possible to acquire information on inhomogeneous behavior, kinetic and velocity, or local heterogeneity of mRNA, which is lost in ensemble measurement. In fact, tracking single β-actin mRNA revealed that β-actin mRNAs move along microtubules toward lamellipodia and its velocity was not constant. The movement of β-actin mRNA showed both processive and static modes, indicating that β-actin mRNA was transported with a motor protein. Third advantage is the relatively small size of the probes. Other methods require many fluorescent molecules to achieve single molecule sensitivity because of high background signals originating from unbound fluorescent molecules [60]. In contrast, this method succeeded in visualizing single mRNA molecule with single fluorescent protein, because the fluorescence complementation and optical

system of TIRF microscopy reduce the background fluorescence signals enough for single molecule imaging. Visualizing with a small tag is preferred because the tags would change the natural conformation of a target mRNA and prevent forming mRNA-protein complexes. In addition, this method is applicable to visualize various kinds of mRNAs. The significant feature of PUM-HD is that I can change its recognition site by introducing some mutations to PUM-HD. In the previous reports, NADH dehydrogenase subunit 6 (ND6) mRNA in mitochondria and RNA of tobacco mosaic virus in plant cells were visualized with the probe made by the same strategy [16, 61]. Thus, this method is widely applicable to tracking various mRNAs of interest.

2-6. Conclusion

Here I demonstrated a method based on PUM-HD and EGFP reconstitution technique to visualize a single endogenous mRNA in living cells. This method presents several advantages compared with previous RNA imaging methods; labeling endogenous mRNA with small tags, high signal-to-noise ratio which enable observation of target mRNAs at the single molecule level, applicability to a wide variety of mRNAs. These advantages in the present method ensure that the method will find widespread use for imaging mRNA. Consequently, using the method will bring great benefits to the studies of mRNA localization and dynamics in living cells.

2-7. Reference

38 Martin, K.C. and Ephrussi, A. mRNA Localization: Gene Expression in the Spatial Dimension. *Cell* 136, 719-730

39 Buxbaum, A.R., *et al.* (2015) In the right place at the right time: visualizing and understanding mRNA localization. *Nature Review of Molecular Cell Biolgy* 16, 95-109

40 Holt, C.E. and Bullock, S.L. (2009) Subcellular mRNA Localization in Animal Cells and Why It Matters. *Science* 326, 1212-1216

41 Weatheritt, R.J., *et al.* (2014) Asymmetric mRNA localization contributes to fidelity and sensitivity of spatially localized systems. *Nature Structual Molecular Biolgy* 21, 833-839

42 Long, R.M., *et al.* (1997) Mating Type Switching in Yeast Controlled by Asymmetric Localization of *ASH1* mRNA. *Science* 277, 383-387

43 Bertrand, E., *et al.* (1998) Localization of ASH1 mRNA Particles in Living Yeast. *Molecular Cell* 2, 437-445

44 Berleth, T., *et al.* (1988) The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. *The EMBO Journal* 7, 1749-1756

45 Gavis, E.R. and Lehmann, R. (1992) Localization of nanos RNA controls embryonic polarity. *Cell* 71, 301-313

46 Weeks, D.L. and Melton, D.A. (1987) A maternal mRNA localized to the vegetal hemisphere in xenopus eggs codes for a growth factor related to TGF-β. *Cell* 51, 861-867

47 Kloc, M. and Etkin, L.D. (1995) Two distinct pathways for the localization of RNAs at the vegetal cortex in Xenopus oocytes. *Development* 121, 287-297

48 Lawrence, J.B. and Singer, R.H. (1986) Intracellular localization of messenger RNAs for cytoskeletal proteins. *Cell* 45, 407-415

49 Huttelmaier, S., *et al.* (2005) Spatial regulation of [beta]-actin translation by Src-dependent phosphorylation of ZBP1. *Nature* 438, 512-515

50 Lin, A.C. and Holt, C.E. (2007) Local translation and directional steering in axons. *The EMBO Journal* 26, 3729-3736

51 Yao, J., *et al.* (2006) An essential role for [beta]-actin mRNA localization and translation in Ca2+-dependent growth cone guidance. *Nature Neuroscience* 9, 1265-1273

52 Holt, Christine E. and Schuman, Erin M. (2013) The Central Dogma Decentralized: New Perspectives on RNA Function and Local Translation in Neurons. *Neuron* 80, 648-657
53 Lécuyer, E., *et al.* (2007) Global Analysis of mRNA Localization Reveals a Prominent Role in Organizing Cellular Architecture and Function. *Cell* 131, 174-187
54 Cajigas, Iván J., *et al.* (2012) The Local Transcriptome in the Synaptic Neuropil Revealed by Deep Sequencing and High-Resolution Imaging. *Neuron* 74, 453-466
55 Hammond, J.W., *et al.* (2009) Mammalian Kinesin-3 Motors Are Dimeric In Vivo and Move by Processive Motility upon Release of Autoinhibition. *PLoS Biol* 7, e1000072
56 Courty, S., *et al.* (2006) Tracking Individual Kinesin Motors in Living Cells Using Single Quantum-Dot Imaging. *Nano Letters* 6, 1491-1495
57 Verhey, K.J. and Hammond, J.W. (2009) Traffic control: regulation of kinesin motors. *Nature Review of Molecular Cell Biolgy* 10, 765-777
58 Shav-Tal, Y., *et al.* (2004) Imaging gene expression in single living cells. *Nature Review of Molecular Cell Biolgy* 5, 855-862
59 Park, H.Y., *et al.* (2014) Visualization of Dynamics of Single Endogenous mRNA Labeled in Live Mouse. *Science* 343, 422-424
60 Tyagi, S. (2009) Imaging intracellular RNA distribution and dynamics in living cells. *Nature Methods* 6, 331-338
61 Tilsner, J., *et al.* (2009) Live-cell imaging of viral RNA genomes using a Pumilio-based reporter. *The Plant Journal : for cell and molecular biology* 57, 758-770

Chapter 3

Development and Characterization of the TERRA probe

3-1. Introduction

The central dogma of gene expression is that DNA is transcribed into messenger RNA (mRNA), which is a template for protein synthesis. Recent advance in genome-wide studies revealed that whereas only 1-3% of the human genome encodes proteins, nearly 60% transcripts non-coding RNA [62]. This fact implies that the central dogma only explains one side of the gene expression. Among non-coding RNAs, particular attention has focused on the transcripts referred to as long non-coding RNAs (lncRNAs), operationally defined as being longer than 200 nucleotides [63-66]. Considering that the prevalence of lncRNA expression in mammalian cells, it has been assumed that lncRNAs constitute a significant fraction of the functional output of mammalian genomes.

Currently, two mechanistic themes of lncRNAs' functions have been emerged [67-69]. **Decoys**: lncRNAs can serve as decoys that interface the binding of regulatory proteins to DNA [70, 71]. **Scaffold**: The lncRNAs can serve as adaptors to bring two or more proeins into discrete complexes [72-75].

Our current understanding of lncRNA has gained from biochemical approaches including northern blotting, RNA microarrays, and high-throughput sequencing. To understand functions of lncRNA in living cells, it is necessary to obtain spatiotemporal information and information of stochastic single-molecule behavior, which can be provided by live cell imaging of endogenous lncRNA at a single molecule level. Live cell imaging of lncRNA has been reported on CCAT1 [76], Xist [77], NEAT1 [78], and TERRA [79] using molecular beacon or MS2 system. These studies revealed distribution of lncRNAs in living cells. Moreover, by using MS2 system, several reports indicated stimulus-dependent and cell

cycle-dependent localization of the lncRNAs to the specific chromosome region. However, there is no direct evidence that lncRNAs act as decoy or scaffold in living cells. To elucidate the molecular mechanism of the lncRNA actions, it is essential to visualize lncRNAs with the single molecule sensitivity.

Here I developed a fluorescent probe (the TERRA probe) for live cell imaging of endogenous Telomeric Repeat-Containing RNA (TERRA) with single molecule sensitivity. TERRA is a lncRNA transcribed from telomere region, the terminal region of mammalian chromosome. TERRA includes hundreds to thousands copies of tandem 5'-UUAGGG-3' repeats and has been proposed to have functions in structure formation and replication of telomere. To visualize TERRA, a domain of RNA-binding protein, PUM-HD, was mutated to recognize tandem repeats in TERRA sequence. Each split fragment of EGFP was conjugated to amino and carboxyl terminal of mutant PUM-HD, respectively. When the two TERRA probe molecules come close on the repeat region, EGFP is reconstituted between the adjacent probes. In this chapter, I describe the strategy to visualize endogenous TERRA at the living cells. Specificity of the TERRA probe on TERRA was evaluated in mammalian cells. Moreover, fluorescence intensity analysis revealed that the TERRA probe visualizes TERRA at the single molecule level.

3-2. Experimental section

3-2-1. Plasmid Construction.

All cDNAs were cloned using pBlueScript (Stratagene, CA). The cDNAs encoding the RNA-binding domain of human PUMILIO1 (PUM-HD) (828–1176 amino acids) and the N-terminal and C-terminal EGFP fragments (GN; 1–157 amino acids, GC; 158–238 amino acids) were generated by polymerase chain reaction (PCR). To generate mPUMt, site-directed amino acid substitutions of PUM-HD (S863N, C935S, Q939E, 971N, Q975S, C1007S, Q1031E, N1043C, S1079N, and E1083Q) were done using a mutagenic PCR technique (Table S1). The cDNAs GN and GC were ligated to mPUMt to yield GN-mPUMt-GC. Three tandem repeats of the NLS sequence DPKKKRKV were connected with the N-terminus of the GN portion. The SNAPf coding region of the hnRNPA1-SNAPf expression vector was prepared with PCR using the pSNAPf-vector (NEB, NA) as the PCR template and primers to add a flexible GGSGGS linker on the N-terminus. The hnRNPA1 cDNA was obtained from Addgene (pET9d-hnRNPA1). hnRNPA1 was fused to the N-terminus of SNAPf. The TRF1 cDNA was generated by PCR and was conjugated to the N-terminus of iRFP 43. The (TTAGGG)13 oligo DNA was purchased from Invitrogen Corp. and was conjugated to the C-terminus of iRFP. NLS-GN-mPUMt-GC and hnRNPA1-SNAPf were subcloned into the mammalian expression vector pcDNA3.1 (+) (Invitrogen Corp., CA). Then iRFP-TRF1 was inserted into the retrovirus expression vector pCLNCX. iRFP-$(TTAGGG)_{13}$ was inserted into a Tet-on inducible vectors, which is an EBV-based episomal vector carrying the tetracycline-regulated expression units, transactivator (rtTA2-M2), and tetO

sequence. This Tet-on vector was kindly provided by Prof. Miwa (Tsukuba University, Japan).

3-2-2. Protein purification

mPUMt gene was subcloned into a pTYB3 vector (New England Biolabs) to produce an N-terminal fusion with the intein/chitin-binding domain. The gene of the fusion protein in the pTYB3 plasmid was expressed in Escherichia coli BL21 strain. The bacterial cells were sonicated in 20 mM sodium phosphate buffer (pH 8.0) containing 1 M NaCl and 0.1 mM PMSF. The obtained lysate was clarified by centrifugation and incubated for 40 min with chitin beads (New England Biolabs). The beads were washed twice with 20 mM sodium phosphate (pH 8.0) buffer containing 0.5 M NaCl and 0.1 mM PMSF, once with 20 mM sodium phosphate (pH8.0), 0.5 M NaCl, and 0.1 mM PMSF, and once with 20 mM sodium phosphate (pH 8.0), 0.15 M NaCl, and 0.1 mM PMSF. Cleavage of the bound fusion protein was initiated by adding 50 mM DTT. During the incubation, the protein solution was purged with argon gas. Recombinant mPUMt was collected and concentrated using Microsep 10 (Miliipore). Protein concentration in the yielded solution was determined by the bicichroninic acid (BCA) assay using bovine serum albumin (BSA) as a standard.

3-2-3. Electron Mobility Shift Assay

RNA oligonucleotides were purchased from Invitrogen and radiolabeled at the 5’ end by using [γ-^{32}P]ATP and T4 polynucleotide kinase (TOYOBO,

Japan) by following the manufacturer directions. Binding reactions included 5 nM radiolabeled RNA and varying concentrations of the protein were incubated in binding buffer (10 mM Hepes, pH 7.4, 50 mM KCL, 1 mM EDTA, 0.01% Tween-20, 0.1 mg/ml BSA, 1 mM DTT). For competition analysis, unlabeled DNA with the TTAGGGTT sequence or RNA with Nanos Response Element (NRE) or the UUAGGGUU sequence was added to the binding mixture. Binding reactions were incubated no more than 2 h at room temperature and immediately analyzed by native PAGE using 6% polyacrylamide gel in 0.5% TBE. The gels were dried and exposed to storage phosphor screens in overnight, and scanned with Fuji 9000.

3-2-4. Cell culture and Transfection

U2OS cells were cultured in Dulbecco's modified Eagle medium (DMEM) (Gibco) containing 10% fetal bovine serum (Gibco) at 37°C in 5% CO_2 atmosphere. For fluorescence imaging, the cells were plated on bottom-glass dishes (Asahi Glass Co., Japan). The cells were transfected with the expression plasmids using Lipofectamine LTX (Invitrogen) according to the manufacturer's protocol. Before image acquisition, the cell cultivation medium was changed to Hanks' Balanced Salt Solution (HBSS) (GIBCO) containing MEM Amino Acids Solution (GIBCO), NEAA (GIBCO), 2.5 g/L glucose, 2 mM glutamine, 1 mM sodium pyruvate, and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.4).

3-2-5. Immunostain and Fluorescence *in situ* hybridization

U2OS Cells were fixed with 15 min incubation at room temperature in 2% PFA in PBS. After 5 min incubation at room temperature in 0.1% sodium citrate/0.1% Triton X-100, the cells were block for 60 min at room temperature in PBS containing 5% (w/v) BSA, 0.1% Tween-20. The cells were incubated for 60 min at room temperature in the presence of mouse anti-GFP IgG (Roche). Alexa-568 goat anti-mouse IgG (Invitrogen) was used as the secondary antibody. PBS containing 0.5% (w/v) BSA, 0.1% Tween-20 was used as washing solution.

U2OS cells grown on bottom-glass dishes were incubated for 7 min in ice-cold CSK buffer (100 mM NaCl, 300 mM sucrose, 3 mM MgCl2, 10 mM PIPES pH 7, 0.5% Triton X-100) containing 10 mM Vanadyl Ribonucleoside Complex (NEB). Cells were then rinsed in PBS, followed by chemical fixation in 2% paraformaldehyde (PFA) in PBS (pH 7) for 15 min and rinsed in 70% ethanol at room temperature. The cells were dehydrated through treatment with an ethanol series (70%, 85% and 100% ethanol for 5 min each) at room temperature, and air-dried. TERRA molecules were hybridized with 5'-TMR-labeled antisense oligonucleotide for 16 hr in 37 °C using 40 μl of a mixture containing 2×Saline Sodium Citrate Buffer (SSC), 2mg/ml BSA, 10% dextrun sulfate, 20% formamide and 10 mM Vanadyl Ribonucleoside Complex. The cell samples were washed three times for 5 min in 2×SSC 50% formamide (pH 7) at 39°C, three times for 5 min in 2×SSC (pH 7) at 39 °C, and 5 min in 2×SSC (pH 7) at room temperature.

3-2-6. Acquisition of imaging data

Single-particle fluorescence imaging of TERRA was performed using a home-built total internal reflection fluorescence (TIRF) microscope system. This system was constructed on an inverted fluorescence microscope (IX81; Olympus Corp., Japan) equipped with laser lines at 488 nm (CYAN-488; Spectra-Physics, CA) and 561 nm (JUNO 561; SOC Corp., Japan) with a PlanApo 100× oil immersion objective with a numerical aperture of 1.49. The laser beams were placed off but parallel to the optical axis to obtain inclined excitation illumination 44,45. Emissions from EGFP and TMR were collected by the objective and were captured by two scientific complementary metal-oxide semiconductor (sCMOS) cameras (ORCA-Flash4.0v2; Hamamatsu Photonics KK, Japan). Then iRFP fluorescence was detected using an electron-multiplying charge-coupled device (EM-CCD) camera (ImagEM; Hamamatsu Photonics KK). Images were acquired using software (MetaMorph; Molecular Devices Corp.) at a frame rate of 10 Hz.

3-2-7. Metaphase telomere FISH

FISH analysis on metaphase spreads were performed as previously described19. Briefly, U2OS cells transfected with empty vector or the TERRA probe were treated with colcemid (0.1 μg/ml; Gibco) for 2 hr to accumulate mitotic cells. Then the cells are trypsinized, resuspended in 75 mM KCl hypotonic solution at 37°C for 30 min, and fixed in fresh 3:1 methanol/acetic acid (v/v) for 4 times. Fixed cells were dropped onto wet glass microscope slides and allowed to dry for overnight. Metaphase spreads were stained with 0.1μg/ml

DAPI and analyzed by FISH using PNA-telomere probe according to manufacturer's instruction (DakoCytomation, Denmark).

3-3. Results

3-3-1. Design of the TERRA probe

To detect UUAGGG-repeats in TERRA, I constructed a mutant PUM-HD (mPUMt) recognizing the RNA sequence of 5'-UUAGGGUU-3' (**Table 3-1**). mPUMt was inserted between the pair of split EGFP fragments (**Figure 3-1 a**). Triple repeats of a nuclear localization signal (NLS) sequence are attached on the N-terminus of split EGFP fragment to translocate the probe into the nucleus (**Figure 3-1 b**). Upon binding of the two TERRA probes to adjacent positions in a TERRA-repeat region, EGFP is reconstituted between the TERRA probes (**Figure 3-1 c**).

Table 3-1.The amino acids mutations to create mPUMt

The amino acids interacting with RNA bases are shown. In each repeat of PUM-HD, three amino acids interact with specific target RNA base. Mutated amino acids and their recognition bases in the PUM-HD repeats in mPUMt are shown in blue and red, respectively. mPUMt was created by 10 mutations (S863N, C935S, Q939E, 971N, Q975S, C1007S, Q1031E, N1043C, S1079N, E1083Q).

wtPUM Repeat	R8	R7	R6	R5	R4	R3	R2	R1
	Q1126	E1083	Q1047	Q1031	Q975	Q939	Q903	Q867
wtPUM	Y1123	N1080	Y1044	Q1011	H972	R936	Y900	R864
	N1122	S1079	N1043	C1007		C935	N899	S863
RNA	5' U	G	U	A	U/C	A	U	A 3'

mPUMt Repeat	R8	R7	R6	R5	R4	R3	R2	R1
	Q1126	Q1083	Q1047	E1031	S975	E939	Q903	Q867
mPUMt	Y1123	N1080	Y1044	Q1011	H972	R936	Y900	R864
	N1122	N1079	C1043	S1007	N971	S935	N899	N863
RNA	5' U	U	A	G	G	G	U	U 3'

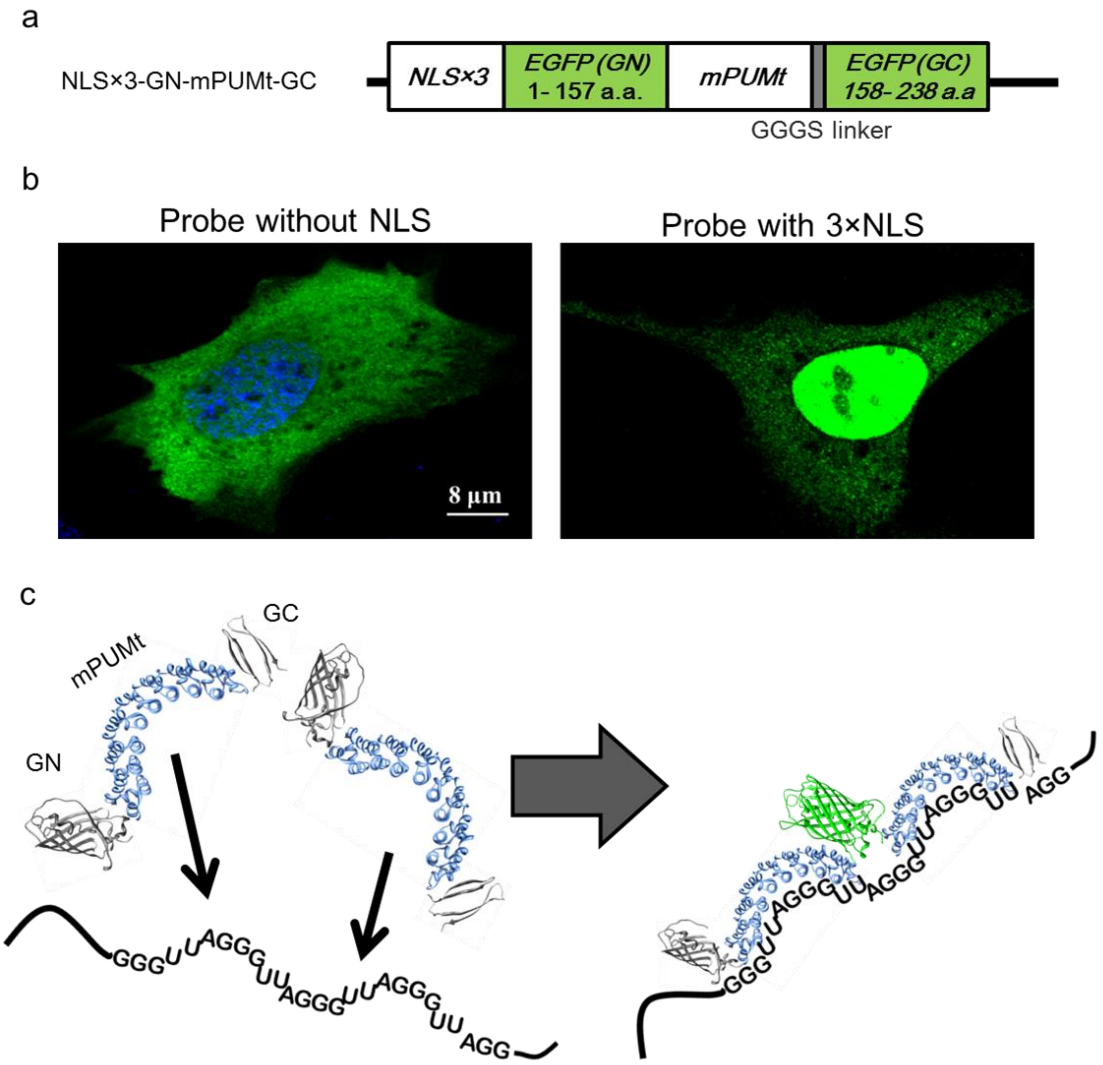

Figure 3-1. Detection of repeat RNA sequence in TERRA based on reconstitution of EGFP and RNA-binding domain, PUM-HD.

(a) Schematic representation of the domain structure of the probe. The mutant PUM-HD binding to repetitive RNA sequence (mPUMt) is sandwiched with EGFP split fragments. GN and GC indicate the cDNA sequence encoding the 1-157 and 158-238 amino acids of EGFP. The construct is fused to three repeats of nuclear localization signals (NLSs). (b) The probes were expressed in U2OS cells and fixed on a glass slide. Fluorescent immunostaining was performed with anti-GFP antibody for each probe. Alexa Fluor 488 labeled second antibody was used to visualize the distribution of the probe. Hoechst 33342 was used for staining the nucleus. The probe analogue without NLS distributed around the cytoplasm (left) and Probe with NLS repeats localized in the nucleus (right). Scale; 8 μm. (c) The functional mechanism of the TERRA probe. Several mPUMt recognize the target repeats sequence of TERRA. Intermolecular EGFP reconstitution is induced between the adjacent TERRA probes.

3-3-2. Characterization of mPUMt *in vitro*

To estimate the affinity of mPUMt to the target RNA sequence, purified mPUMt was subjected to an electrophoretic mobility shift assay (EMSA) using radioisotope (RI)-labeled 5’-UUAGGGUU-3’. In the presence of mPUMt, a band of RI-labeled TERRA RNA was detected at a position of larger molecular weight than that of the free RNA (**Figure 3-2 a**). The dissociation constant between mPUMt and 5’-UUAGGGUU-3’ was calculated to be 0.10 ± 0.01 μM from a plot of average amounts of bound fractions in variety of mPUMt concentrations (**Figure 3-2 b**). Competition with 100- and 1,000-fold excess of unlabeled 5’-UUAGGGUU-3’ prevented the binding of mPUMt to the isotope-labeled RNA (**Figure 3-2 c**). On the other hand, RNA without 5’-UUAGGGUU-3’ and single or double strand DNA with 5’-TTAGGGTT-3’ did not affect the binding of mPUMt to the RNAs (**Figure 3-2 c**). These results indicate that mPUMt specifically interacts with 5’-UUAGGGUU-3’.

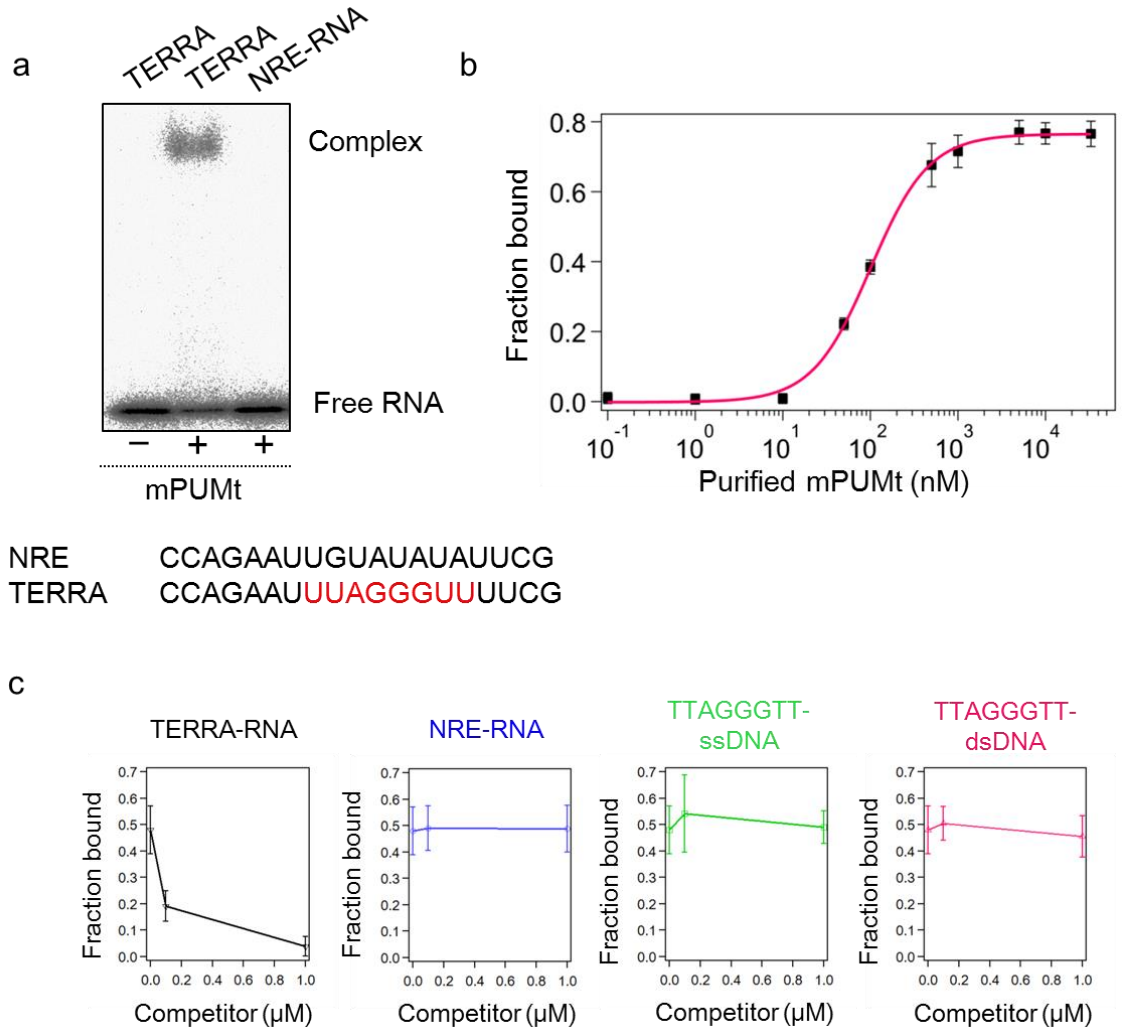

Figure 3-2. Binding affinity of mPUMt to the target RNA sequence

The binding of recombinant mPUMt induces a mobility shift in TERRA. (a) Specific recognition of UUAGGG *in vitro*. Below, sequences of tested TERRA-RNA and NRE-RNA. (b) The percentage of RNA bound to various concentrations of mPUMt. (c) Analysis of competitive RNA binding for mPUMt. Averaged fraction bounds of each condition were plotted versus concentration of competitors (UUAGGGUU-containing RNA (black), NRE sequence containing RNA (blue), single (green) and double (red) strand of TTAGGGTT-DNA).

3-3-3. Reconstitution of EGFP on TERRA in mammalian cells

To confirm the EGFP reconstitution of the TERRA probe with UUAGGG repeats, I generated human osteosarcoma (U2OS) cells carrying transcriptionally inducible iRFP-$(TTAGGG)_{13}$ whose transcription is stimulated by doxycycline (dox). Treatment with dox for two hours led to an approximately two-fold increase in the number of EGFP spots in the nucleus (**Figure 3-3 a, b**). Importantly, EGFP spots also appeared in the cytoplasm where endogenous TERRA does not exist (**Figure 3-3 a**), suggesting that the TERRA probe reconstituted EGFP on the artificially expressed UUAGGG repeats.

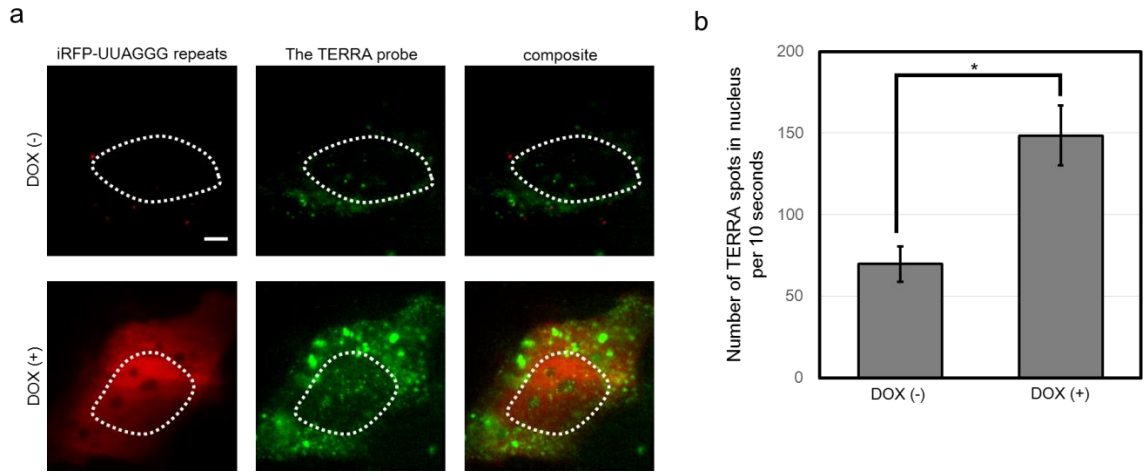

Figure 3-3. EGFP reconstitution of the TERRA probe with artificial UUAGGG repeats
(a) Fluorescence images were obtained from U2OS cells expressing iRFP-$(UUAGGG)_{13}$ (left), the TERRA probe (middle), and the merged image (right). Nucleus is indicated with white broken line. Scale bar, 5.0 μm. (b) Quantification of the number of TERRA spots in the nucleus in the cells without dox treatment (DOX(-)) and with dox treatment (DOX(+)). *P*-value was computed using the Student's *t*-test. * indicates $P < 0.05$.

Next, I examined the ability of the TERRA probe to visualize endogenous TERRA molecules in U2OS cells through a fluorescence in situ hybridization (FISH) experiment. The cells expressing the TERRA probe were chemically fixed, and endogenous TERRA was hybridized with a tetramethylrhodamine (TMR)-labeled oligonucleotide (TMR-$(GGGTTA)_7$). Fluorescence signals of both EGFP and TMR signals were visualized with highly oblique incidence fluorescence microscopy. The colocalization ratio of EGFP spots with TMR spots was 71% (**Figure 3-4**). This high rate of colocalization demonstrates that EGFP of the TERRA probe is reconstituted specifically on TERRA in the cells.

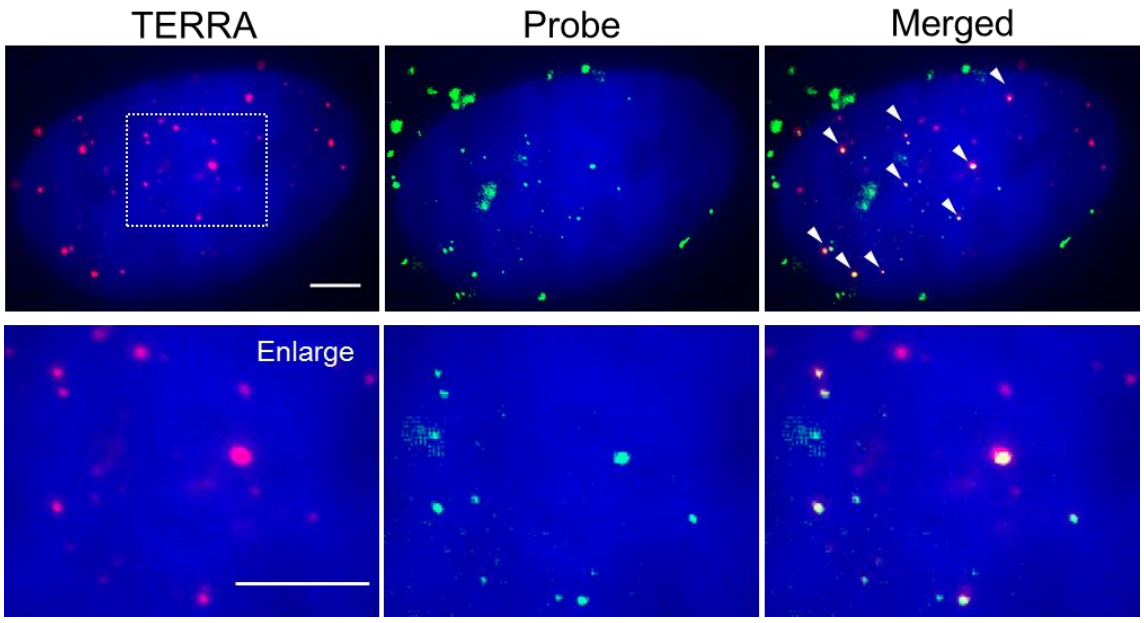

Figure 3-4. Localization of TERRA and the probes

The position of TERRA was visualized by FISH using 5' TAMRA-labeled oligonucleotide in the cell expressing the probes (left). Fluorescence from the probes is also observed in the identical cell (Middle). The white arrowheads in merge image (right panel) indicated that the positions of the probes colocalized with TERRA (right). Bottom panels showed enlarged images of the regions outlined with white dashed lines. The scale bar represents 5 μm.

To further confirm specific EGFP reconstitution on TERRA, I prepared two probe variants; one lacks the N-terminal fragment of split EGFP (mPUMt-GC) and the other containing a β-actin mRNA-recognizing PUM-HD mutant, $mPUM_{ACT}$ instead of mPUMt (3×NLS-GN-$mPUM_{ACT}$-GC). $mPUM_{ACT}$ binds to an RNA sequence of UGUGCUGU [80], of which repetitive sequences do not exist in human transcripts. These variants in U2OS cells were immunostained with Alexa Fluor 568-labeled anti-GFP antibody to show their expression and intracellular distribution. Strong fluorescence signals of Alexa Fluor 568 signals from both variants were observed in the nucleus (**Figure 3-5**), indicating localization of the variants in the nucleus. There was no fluorescence spot of mPUMt-GC and 3×NLS-GN-$mPUM_{ACT}$-GC in the nucleus (**Figure 3-5**), indicating that inter- and intra- molecular reconstitutions of EGFP do not occur spontaneously. Together, these results show that the TERRA probes visualize endogenous TERRA molecules selectively in the cells.

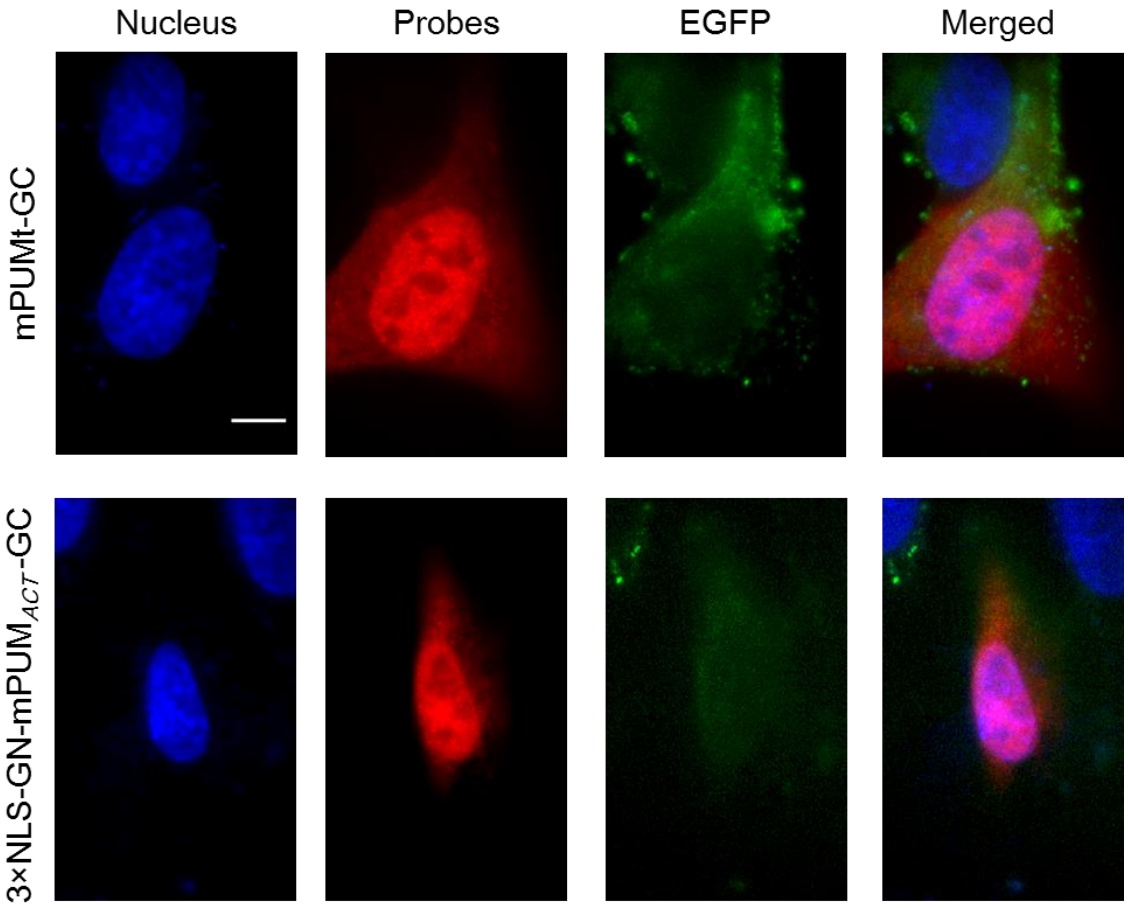

Figure 3-5. Fluorescence signals from the probe analogs and their immunostaining.

The probe analogues were expressed in U2OS cells and fixed on a glass slide. After the fixation, the analogues were immunostained with anti-GFP antibody which binds to full-length EGFP and the C-terminal fragment of EGFP. Alexa Fluor 561 labeled second antibody was used to visualize the distribution of the probe. Hoechst 33342 was used for staining the nucleus. The analogue lacking N terminal fragment of EGFP and NLS partially localized in nucleus (above). Fluorescence signals of EGFP were not detected in the nucleus. The analogue containing mPUM$_{ACT}$ localized to nucleus, fluorescence signal of EGFP was not detected (bottom). Scale; 8 μm.

3-3-4. Visualization of single TERRA molecules

To investigate the applicability of the TERRA probe for single molecule imaging in living cells, I estimated the number of EGFP molecules in the single fluorescent spots. Based on previous information that a fluorescent spot consisting of a single fluorophore shows single-step bleaching [81], I investigated time-course intensities profile of the fluorescent spots. In living cells, fluorescent spots diminished from the focal plane by diffusing in axial direction. To avoid misrecognition of diffusing out from the focal plane as bleaching, I analyzed fluorescence intensities of reconstituted EGFP in fixed U2OS cells. Fluorescent spots of EGFP were fitted with 2D Gaussian function to estimate fluorescence intensities. A fluorescent spot showed fluctuation in its fluorescence intensity and disappeared by single-step bleaching (**Figure 3-6 a**), demonstrating that the spot consists of single EGFP. The histogram of the fluorescence intensities was fitted with a single Gaussian distribution (**Figure 3-6 b**), indicating that the analyzed fluorescent spots consisted of the same number of the reconstituted EGFPs. Taken these two results together, we concluded that the fluorescent spots represented single EGFPs in the fixed cells. To further verify that fluorescent spots in living cells are composed of single EGFPs, the intensity of fluorescent spots in living cells were compared with those in the fixed cells. The histogram of the fluorescence intensities in living cells showed a single Gaussian distribution (**Figure 3-6 b, right**). The peak intensity of fluorescent spots in living cells was 45.2 ± 1.0 a.u. (mean ± s.e.m.), which was substantially the same as that in fixed cells (44.2 ± 0.5 a.u., **Figure 3-6 b, left**), indicating that

each fluorescent spot detected in the living cells consists of a single EGFP molecule.

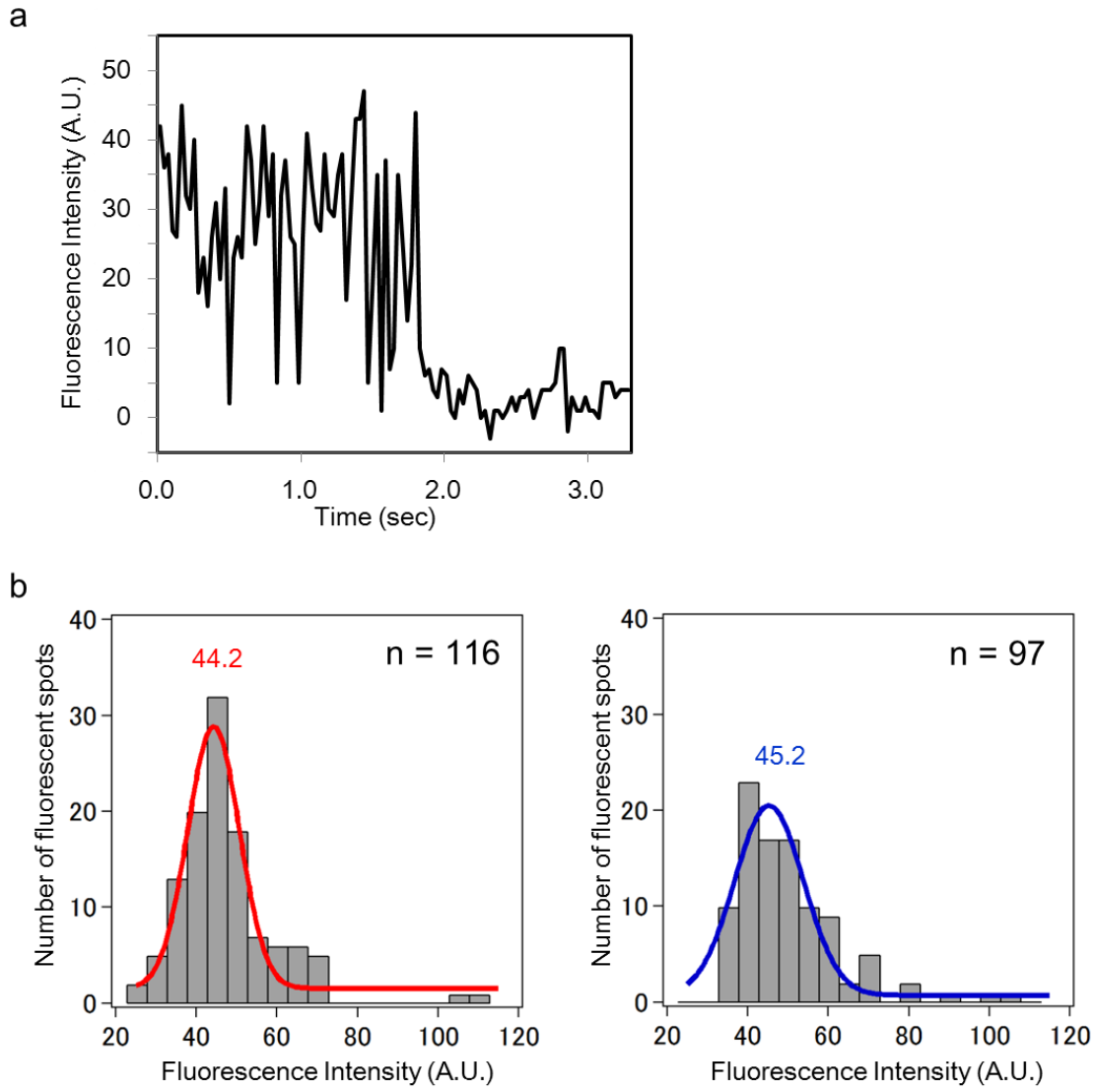

Figure 3-6. Imaging a single TERRA molecule in live cells

(a) Time dependency of fluorescence intensity on an EGFP fluorescent spot in fixed cell. A single-step photobleaching occurred before 1.8 sec. (b) Histograms showing fluorescence intensity distribution of individual spots in fixed cells (left) and in live cells (right). The histograms were fit to a Gaussian distribution (line) with a peak value 44.2 a.u. for fixed cells and 45.2 a.u. for live cells.

Although the spots observed in the cells consist of a single EGFP, multiple EGFP should be reconstituted, in principle, because multiple probes can bind to TERRA. To demonstrate the ability of the probe to reconstitute multiple EGFP, an artificial TERRA containing six repeats of 5'-UUAGGG-3' repeats was expressed in cells. mCherry was conjugated with the UUAGGG repeats to confirm its expression. After the transfection of the TERRA probe and the artificial TERRA, cells were chemically fixed. Fluorescence of EGFP and mCherry was detected by the fluorescence microscopy. EGFP spots were detected from the cells showing mCherry signals (**Figure 3-7 a**). A few EGFP spots showed two-step bleaching. Moreover, step sizes of the fluorescence intensity between the first-step and the second-step was identical (**Figure 3-7 b**). These results demonstrated that that three TERRA probes bind to six UUAGGG repeats and reconstituted two EGFP.

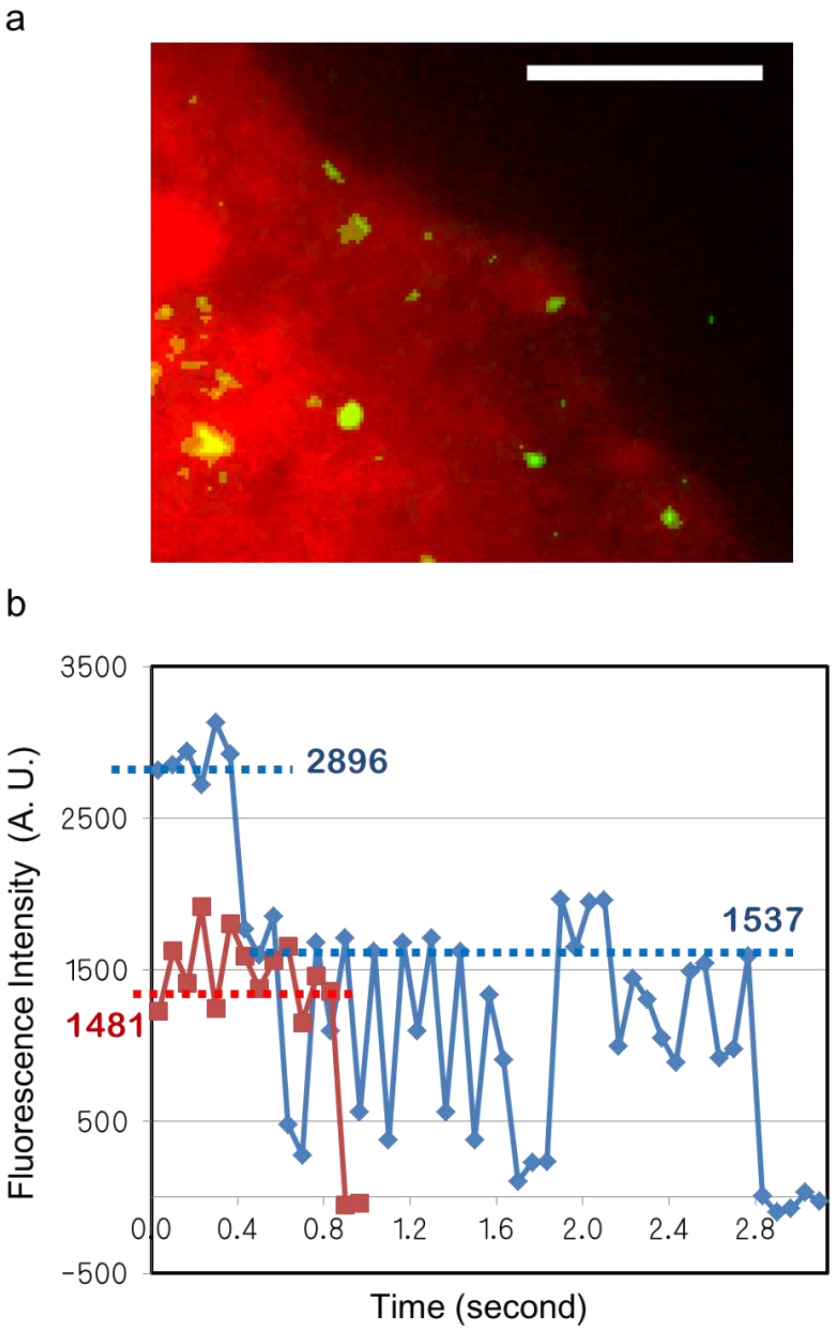

Figure 3-7. Double-step bleaching of the fluorescent spot

(a) Reconstituted EGFP of the TERRA probe on artificial RNA containing six repeats of UUAGGG. Red fluorescence from mCherry indicates the expression of the artificial RNA. Scale, 5 um. (b) A typical fluorescence time trace for the photobleaching of a single EGFP (red) and two EGFP (blue). Broken lines indicate the averaged fluorescence intensity.

3-3-5. The effect of the TERRA probe on telomere physiology

Function of TERRA is considered as maintenance of telomere. To examine whether the TERRA probe inhibits TERRA function, the effects of the TERRA probe on telomere structure and length were examined by the telomere quantitative FISH method (q-FISH). Telomere q-FISH uses fluorescently labeled synthetic DNA mimic which specifically binds to telomere repeats to quantify the number of UUAGGG repeats [82, 83]. To obtain single chromosomes, colcemid was added to the cells expressing the TERRA probes. Colcemid arrest the cells in the metaphase state by disrupting microtubule in mitotic cells. After the fixation of the cells, single chromosomes were extracted by resuspension with hypotonic buffer. The fluorescence of hybridized telomeric repeats was detected by fluorescence microscopy. The expression of the TERRA probe did not cause any significant changes in average telomere length (**Figure 3-8**). Moreover, I investigated the ratio of telomere aberration by calculating the number of a telomere free end (TFE) which was defined as the chromosome end without telomere (**Figure 3-8 b**). The percentage of TFE from the cells with the TERRA probe was 16.8% (n = 184), which was comparable to that without the TERRA probe (12.0%). These results indicated that the TERRA probe does not inhibit the ability of TERRA to maintain the telomere length and structure.

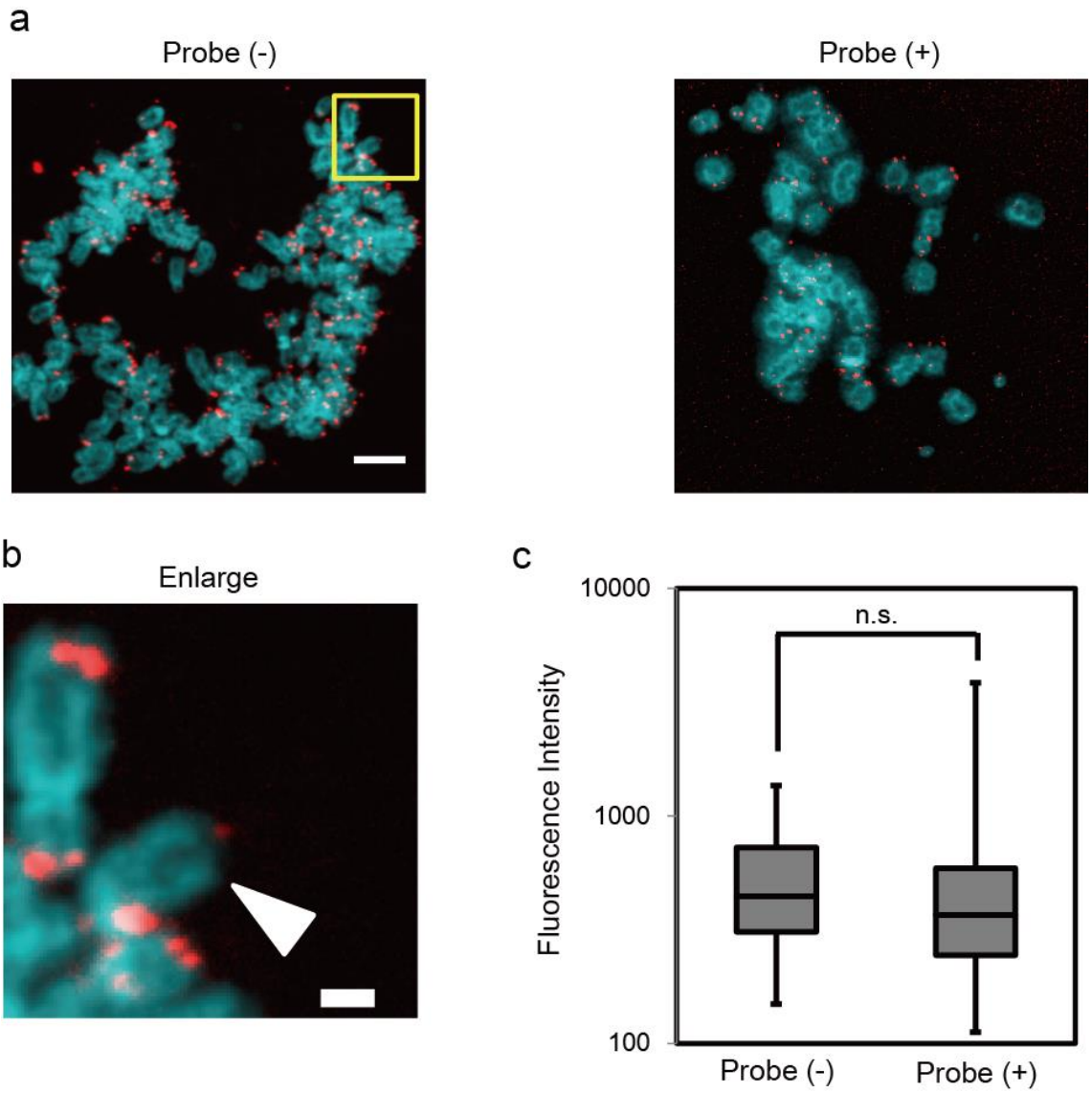

Figure 3-8. The effect of the TERRA probe on telomere length and structure

(**a**) Representative telomere quantitative FISH on metaphase spreads for U2OS cells in the absence (left panel) or presence (right panel) of the TERRA probe. Metaphase spreads were stained for telomeric DNA (red) and counterstained with DAPI (cyan). Scale bar, 10.0 μm. (**b**) An enlarged image of chromosomes with telomere free end (white arrowhead). Scale bar, 1.0 μm. (**c**) Quantification of fluorescence intensity of the telomere for U2OS cells in the presence or absence of the TERRA probe. *P*-value was computed using the Student's *t*-test. n.s. indicates $P > 0.05$.

3-4. Discussion

Here I developed the TERRA probe, a probe for visualizing endogenous TERRA with single molecule sensitivity. In the previous design, two PUM-HD mutants binding to different sites on the target RNA were connected with N- or C-terminal fragments of EGFP, respectively [80]. Contrary to the previous design, I developed a fluorescent probe consisting of a single PUM-HD mutant connected with both N- and C-terminal fragments of EGFP. The newly designed probe comprised of a single protein molecule which is encoded on a single expression plasmid. This is an advantage compared with the previous design because of no need to transfect two expression plasmids.

Specificity of the TERRA probe to TERRA is achieved through PUM-HD and reconstitution of EGFP. The electrophoretic mobility shift assay (EMSA) revealed that mPUMt did not recognize TTAGGGTT-DNA. This result indicated that mPUMt does not bind to telomere in the nucleus. Previously, PUM-HD has been shown to bind the target RNA 2000- fold higher than the equivalent DNA [21]. Moreover, single mutation in the target RNA reduces the affinity of PUM-HD 10- to 1000- fold [21]. This feature of PUM-HD ensures the high specificity of the TERRA probe to TERRA. In addition, reconstitution of EGFP enables 16-base sequences recognition of the probe. By combining split EGFP and PUM-HD, the probe theoretically can distinguish 4^{16} (about 4.3×10^{9}) transcripts, which is more variety than the human genome.

By using the TERRA probe, I visualized TERRA at the single molecule level. In principle, several EGFP should be reconstituted because multiple probes bind to single TERRA. I showed that three TERRA probes bind to artificial

UUAGGG-repeats and reconstitute two EGFP. The reason why single EGFP is reconstituted on the TERRA can be explained by the rigid structure of TERRA, which will not be recognized by PUM-HD. It is known that telomeric repeats form series of rigid structures, named G-quadruplex structures [84, 85]. Previous reports revealed that most of the repeat regions form the G-quadruplex structure and flexible region is infrequent [85]. Considering that PUM-HD can only bind to single-strand of RNA, mPUMt only binds to the flexible region in the telomeric repeats. These results demonstrated that multiple reconstitution of EGFP on TERRA is unrealistic.

3-5. Conclusion

Here I developed the TERRA probe, a probe for visualizing endogenous TERRA with single molecule sensitivity. Mutated PUM-HD (mPUMt) binds to the conserved repeat regions in TERRA, enabling us to target all the sequence variants of TERRA. The EGFP reconstitution approach reduces he background signals from the probe unbound to TERRA, which is suitable for single TERRA molecule imaging.

By mutating the amino acids of PUM-HD, the probe developed in this study has a potential to be applied to RNAs which include repeat region in the sequence. Some pathogenic RNA repeats usually contain hundreds of copies of the trinucleotide, such as CAG repeats in Huntington diseases and CUG repeats in myotonic dystrophy [86]. Importantly, many lncRNAs, such as GRC-RNA [87] and Gomafu [88], contain repeat regions in their sequences.

3-6. Reference

62 Derrien, T., *et al.* (2012) The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Research* 22, 1775-1789

63 Ørom, Ulf A. and Shiekhattar, R. (2013) Long Noncoding RNAs Usher In a New Era in the Biology of Enhancers. *Cell* 154, 1190-1193

64 Nagano, T. and Fraser, P. (2011) No-nonsense functions for long noncoding RNAs. *Cell* 145, 178-181

65 Wapinski, O. and Chang, H.Y. (2011) Long noncoding RNAs and human disease. *Trends in Cell Biology* 21, 354-361

66 Fatica, A. and Bozzoni, I. (2014) Long non-coding RNAs: new players in cell differentiation and development. *Nature Review Genetics* 15, 7-21

67 Wang, K.C. and Chang, H.Y. (2011) Molecular mechanisms of long noncoding RNAs. *Molecular Cell* 43, 904-914

68 Hung, T. and Chang, H.Y. (2010) Long noncoding RNA in genome regulation. *RNA Biology* 7, 582-585

69 Moran, V.A., *et al.* (2012) Emerging functional and mechanistic paradigms of mammalian long non-coding RNAs. *Nucleic Acids Research* 40, 6391-6400

70 Cesana, M., *et al.* (2011) A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell* 147, 358-369

71 Sun, S., *et al.* (2013) Jpx RNA Activates Xist by Evicting CTCF. *Cell* 153, 1537-1551

72 Tsai, M.-C., *et al.* (2010) Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science* 329, 689-693

73 Chu, C., *et al.* (2011) Genomic Maps of Long Noncoding RNA Occupancy Reveal Principles of RNA-Chromatin Interactions. *Molecular Cell* 44, 667-678

74 Kaneko, S., *et al.* (2010) Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes & Development* 24, 2615-2620

75 Rinn, J.L., *et al.* (2007) Functional Demarcation of Active and Silent Chromatin Domains in Human HOX Loci by Noncoding RNAs. *Cell* 129, 1311-1323

76 Kam, Y., *et al.* (2014) Detection of a long non-coding RNA (CCAT1) in living cells and human adenocarcinoma of colon tissues using FIT–PNA molecular beacons. *Cancer Letters* 352, 90-96
77 Ng, K., *et al.* (2011) A system for imaging the regulatory noncoding Xist RNA in living mouse embryonic stem cells. *Molecular Biology of the Cell* 22, 2634-2645
78 Mao, Y.S., *et al.* (2011) Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nature Cell Biolgy* 13, 95-101
79 Cusanelli, E., *et al.* (2013) Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Molecular cell* 51, 780-791
80 Yamada, T., *et al.* (2011) Visualization of Nonengineered Single mRNAs in Living Cells Using Genetically Encoded Fluorescent Probes. *Analytical Chemistry* 83, 5708-5714
81 Sako, Y., *et al.* (2000) Single-molecule imaging of EGFR signalling on the surface of living cells. *Nature Cell Biolgy* 2, 168-172
82 Slijepcevic, P. (2001) Telomere length measurement by Q-FISH. In *Chromosome Painting: Principles, Strategies and Scope* (Sharma, A.K. and Sharma, A., eds), pp. 17-22, Springer Netherlands
83 Deng, Z., *et al.* (2009) TERRA RNA Binding to TRF2 Facilitates Heterochromatin Formation and ORC Recruitment at Telomeres. *Molecular Cell* 35, 403-413
84 Collie, G.W., *et al.* (2010) Electrospray Mass Spectrometry of Telomeric RNA (TERRA) Reveals the Formation of Stable Multimeric G-Quadruplex Structures. *Journal of the American Chemical Society* 132, 9328-9334
85 Collie, G.W., *et al.* (2010) A crystallographic and modelling study of a human telomeric RNA (TERRA) quadruplex. *Nucleic Acids Research* 38, 5569-5580
86 Krzyzosiak, W.J., *et al.* (2011) Triplet repeat RNA structure and its role as pathogenic agent and therapeutic target. *Nucleic Acids Research*
87 Zheng, R., *et al.* (2010) Polypurine-repeat-containing RNAs: a novel class of long non-coding RNA in mammalian cells. *Journal of Cell Science* 123, 3734-3744
88 Tsuiji, H., *et al.* (2011) Competition between a noncoding exon and introns: Gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. *Genes to Cells* 16, 479-490

Chapter 4

Analysis of TERRA Dynamics at the Single Molecule Level

4-1. Introduction

TERRA has been proposed to have multiple functions, such as heterochromatin formation and replication of telomeres [89]. Of the different functions of TERRA, it is suggested that there is a common molecular mechanism; regulating the localization of telomeric related proteins to telomere. In *Saccharomyces cerevisae*, TERRA coordinates the recruitment and activity of telomerase at telomere [79]. In mammalian cells, TERRA maintains heterochromatin and DNA stability at telomere repeats by connecting telomere repeat factor 2 (TRF2) and origin recognition complex (ORC), which is critical for maintaining heterochromatin proteins (HP1) and histone marks (H3K9me3) at telomeres [83, 90]. Moreover, TERRA regulates accessibility of heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) to a telomere during telomere replication. hnRNPA1 localizes to telomere when TERRA expression decreases in late S phase. This hnRNPA1 localization displaces replication protein A (RPA) from telomeric ssDNA. The binding of hnRNPA1 to the telomeric ssDNA is inhibited by TERRA re-accumulation at the telomere, which provides a window for protection of telomeres 1 (POT1), but not RPA, to bind. In this manner, TERRA and hnRNPA1 orchestrate a cell-cycle-regulated RPA-to-POT1 switch on telomeric ssDNA, ensuring orderly telomere replication and capping [91, 92]. Despite the fact that TERRA and chromatin-modifying factors cooperate to regulate telomere states, the function of TERRA in telomere regulation remains elusive. Although the localization and motion of TERRA are governed by physical laws, simple Brownian motion or static confinement of TERRA at telomeres cannot explain the TERRA functions. The underlying mechanism which connects TERRA dynamics to its

function remains unclear. Spatial and temporal information of TERRA, its interacting proteins, and telomeres in living cells will lead to deeper understanding of TERRA mechanism. By using the TERRA probe, I investigated the distributions and motions of TERRA and hnRNPA1 in living cells at single-particle level. Based on the single-particle analysis, I propose a mechanistic model by which TERRA functions as a scaffold to hold hnRNPA1 around a telomere, inhibiting the localization of hnRNPA1 to the telomere.

4-2. Experimental section

4-2-1. Plasmid construction

Fluorescent tags (mCherry, iRFP, and $SNAP_f$) were conjugated to the N terminus of hTRF1 to visualize telomere positions. The fluorescent tags and TRF1 were amplified with PCR and were ligated. The fusion proteins were subcloned into a mammalian expression vector, pcDNA3.1 (+) (Invitrogen).

4-2-2. Cell culture, transfection, and SNAP labeling

U2OS cells were cultured in Dulbecco's modified Eagle medium (DMEM) (Gibco) containing 10% fetal bovine serum (Gibco) at 37°C in 5% CO_2 atmosphere. For fluorescence imaging, the cells were plated on bottom-glass dishes (Asahi Glass Co., Japan). The cells were transfected with the expression plasmids using Lipofectamine LTX (Invitrogen) according to the manufacturer's protocol. SNAP-tag was incubated with 0.1 μM TMR for 0.5 hr in 37 °C for labeling before the observation. The unbound TMR was removed by washing with DMEM containing 10% FBS. Before image acquisition, the cell cultivation medium was changed to Hanks' Balanced Salt Solution (HBSS) (GIBCO) containing MEM Amino Acids Solution (GIBCO), NEAA (GIBCO), 2.5 g/l glucose, 2 mM glutamine, 1 mM sodium pyruvate, and 10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.4). For the heat shock experiment, the cells were incubated at 42°C for 1 hr before the data acquisition. For the heat shock experiment, the cells were incubated at 42°C for 1

hour before the data acquisition. For the doxycycline treatment, the cells were treated with doxycycline 10 ng/mL for 2 hr.

4-2-3. RNA isolation and reverse transcription quantitative real-time polymerase chain reaction (RT-qPCR)

Cells transfected with siControl or siTERRA-1 or -2 were treated with RNAiso Plus (TAKARA, Japan) and were collected to a centrifuge tube containing YTZ ball with a diameter of 5 mm (NIKKATO, Japan). Cells were homogenized by oscillating the tube at the speed of 50 /sec for 2 min. After the homogenization of the cells, $CHCl_3$ were added to the tubes. The tubes were centrifuged at 12,000 g for 15 min at 4°C. The top layers were transferred to new tubes and added an equal amount of isopropanol. After the vortexing, the tubes were centrifuged at 12,000 g for 15 min at 4°C and the supernatant were removed. 75% cold ethanol was added to the tubes. The tubes were centrifuged at 12,000 g for 1 min at 4°C and supernatants were discarded. After the precipitate is dry, isolated RNA was dissolved by RNase-free water.

Isolated RNA was reverse transcribed into cDNA using the SuperScript 4 (Invitrogen Corp.). The cDNA was amplified using the primer set listed in [93]. SYBR Premix Ex Taq2 (Perfect Real Time) (TaKaRa) was used according to the manufacturer's instructions. Quantitative real-time reverse transcription polymerase chain reaction (RT-qPCR) analysis was performed using a Thermal Cycler Dies Real Time system (TaKaRa).

4-2-4. Acquisition of imaging data

Single-particle fluorescence imaging of TERRA was performed using a home-built total internal reflection fluorescence (TIRF) microscope system. This system was constructed on an inverted fluorescence microscope (IX81; Olympus Corp., Japan) equipped with laser lines at 488 nm (CYAN-488; Spectra-Physics, CA) and 561 nm (JUNO 561; SOC Corp., Japan) with a PlanApo 100× oil immersion objective with a numerical aperture of 1.49. The laser beams were placed off but parallel to the optical axis to obtain inclined excitation illumination 44,45. Emissions from EGFP and TMR were collected by the objective and were captured by two scientific complementary metal-oxide semiconductor (sCMOS) cameras (ORCA-Flash4.0v2; Hamamatsu Photonics KK, Japan). Then iRFP fluorescence was detected using an electron-multiplying charge-coupled device (EM-CCD) camera (ImagEM; Hamamatsu Photonics KK). Images were acquired using software (MetaMorph; Molecular Devices Corp.) at a frame rate of 10 Hz.

4-2-5. Analysis of single particle trajectories

Positional differences between the cameras and image distortion resulting from optical aberrations in the fluorescence images were corrected using a pinhole array, as described previously by Koyama-Honda et al. [94]. The pinhole array contained a lattice of 1-μm-diameter holes at 5-μm intervals. A bright-field transmission image of the array was recorded on each camera. The detected pixel positions of the holes were used to determine a set of third-order spline functions, which describe the pixel-to-coordinate relation for each camera. This set of functions was then applied for undistortion of the raw fluorescence images to obtain corrected images.

Before analysis, image noise reduction was performed through Gaussian blur filtering with ImageJ software. The *x* and *y* coordinates of each fluorescent spot were determined using a homemade program based on the cross-correlation method, as described previously [95-98]. In this method, an image and a kernel are described, respectively, as intensity matrices **I** and **K**, which contain both fluorescence signals and background. We set **K** to exhibit a 2D Gaussian distribution with a FWHM of 200 nm for EGFP and TMR and 500 nm for iRFP. The image was scanned by the kernel in one-pixel increments. For each increment, the program calculated a correlation value that describes how well the values in the kernel match those of the underlying image. At the relative shift, the position at which the kernel and the image were most similar was a maximum in the correlation matrix (**X**). Here, the cross-correlation between the kernel (**K**) and the image (**I**) is given as shown below.

(1) $$\mathbf{X}_{x,y} = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} \boldsymbol{I}_{x+i,y+j}$$

In that equation, *x* and *y* denote the distance that the kernel (**K**) has moved over the original image (**I**). In the correlation matrix (**X**), regions with correlation values exceeding those of the background and with areas larger than five pixels were regarded as fluorescent spots. Their positions were determined to be the centroids of the correlation values of the detected spots.

The localization precision of the detected EGFP and TMR spots was estimated from the standard deviations of individual spot positions in chemically fixed cells. To create images of TERRA, hnRNPA1, and telomere without background, Gaussian spots with FWHM of 200 nm were placed on a blank image at the coordinates at which the EGFP and TMR spots were detected. Also,

iRFP spots are shown on the same image as Gaussian spots with FWHM of 500 nm. TERRA foci were defined as the nucleus position in which more than 10 TERRA spots appeared during 8 seconds observation.

The radial distribution function, ρ(*r*), was defined as the time-averaged density of TERRA spots in a region within a distance of *r*-0.25 to *r* (μm) from the center of a telomere. The ρ(*r*) function was calculated using the following equation (2).

(2) $$\rho(\mathrm{r}) = \sum_{t=0}^{T} \frac{N_{\mathrm{r}}(t)}{\pi r^2 - (r-0.25)^2} \times \frac{1}{T}$$

Therein, $N_{\mathrm{r}}(t)$ is the number of detected TERRA spots in the region of *r*–*r*-0.25 μm from a telomere at time *t*. *T* is the observation duration (8 s).

For each fluorescent spot's trajectory, the mean square displacement (MSD) was calculated using the following equation (3).

(3) $$\mathrm{MSD}(n\Delta t) = \sum_{i=1}^{N-n} \frac{(x_{i+n\Delta t} - x_i)^2 + (y_{i+n\Delta t} - y_i)^2}{N-\mathrm{n}}$$

In that equation, Δ*t* is the sampling time interval, x_{i} and y_{i} are the *x* and *y* coordinates of the fluorescent spot at time *i*Δ*t*, and *N* is the total frame number of the trajectory [36, 37]. The diffusion coefficient was derived from the slope of the MSD-Δ*t* plot for 0–300 ms using least-squares fitting. Because of the limitation of the localization precision, the positions of the EGFP spots in chemically fixed cells fluctuated from frame to frame, yielding an "apparent" diffusion coefficient. A histogram of the apparent diffusion coefficients was fitted with a log normal distribution function, from which the 99% confidence limit of the non-mobile spots was inferred. For classification of the fluorescent spot movements into stationary or mobile modes, spots with a diffusion

coefficient smaller than the 99% confidence limit were judged as stationary. The remainders were classified as mobile.

An hnRNPA1 spot and a TERRA spot were regarded as colocalized if the distance between these spots was less than or equal to a threshold value. The threshold was determined as the sum of the localization precisions of EGFP and TMR and the overlay accuracy between the EGFP and TMR images. The localization precisions of EGFP and TMR were estimated from the standard deviations of the spot positions in chemically fixed cells as $(\sigma_x + \sigma_y)/2$, where σ_x and σ_y are the standard deviations of the centroid positions in the x and y dimensions, respectively [99]. To evaluate the overlay accuracy of the images measured on different color channels, TetraSpeck Fluorescent Microspheres on the glass surface were compared after spatial correction. The differences in the centroid positions (Δx and Δy) of the microspheres between color channels were calculated. The overlay accuracies were ascertained as $(\sigma'_x + \sigma'_y)/2$, where σ'_x and σ'_y are the standard deviations of Δx and Δy, respectively. All calculations were done using IGOR Pro software (WaveMetrics Inc., Lake Oswego, OR).

4-2-6. Simulations of free diffusion

Free diffusion was simulated at the diffusion coefficients, 0.24 $\mu m^2/s$, which is the average value of the TERRA. The jumpsize l was obtained for time t = 0.05 sec, according to $l = (4Dt)^{1/2}$. For each jump, a random direction was chosen

based on normal random number generated by Box-Muller's method. Each simulated trajectory consisted of 200 frames.

4-3. Results

4-3-1. Live cell imaging of TERRA and telomere

Using the TERRA probe, which I developed and characterized in chapter 3, I first investigated localization of TERRA and telomeres in living U2OS cells. The TERRA probe was introduced to the cells which stably express near-infrared fluorescent protein (iRFP)-TRF1. Fluorescence signals of EGFP and iRFP were detected by the fluorescence microscope. TERRA was visualized as individual fluorescent particles, whereas the TRF1 localized extensively to telomeres and was observed as punctate fluorescence spots in the nucleus (**Figure 4-1 a**). Because of the weal intensity of the single EGFP spot, background fluorescence hampered to determine exact position of TERRA molecules. To eliminate background fluorescence, I detected the position of individual EGFP spots by using cross-correlation. In this detection system, thickness of spot-detectable area was ± 200 nm from the center of the focus plane, as confirmed by the observation of fluorescent beads with different vertical positions (**Figure 4-1 b)**. Then, the EGFP spots from the raw image were fitted with Gaussian spots with 200 nm diameter, corresponding to the full width at half maximum of the EGFP spots in the raw image (**Figure 4-1a)**. Similarly, Gaussian spots with diameter of 500 nm were applied at the positions of iRFP-TRF1.

To analyze the distribution of TERRA in nucleus, I integrated the time-course observation of TERRA molecules. The time-integration of the Gaussian spots showed that TERRA formed foci in specific regions in the nucleus (**Figure 4-1 c**). The ratio of telomeres colocalizing with TERRA foci was 8.9%.

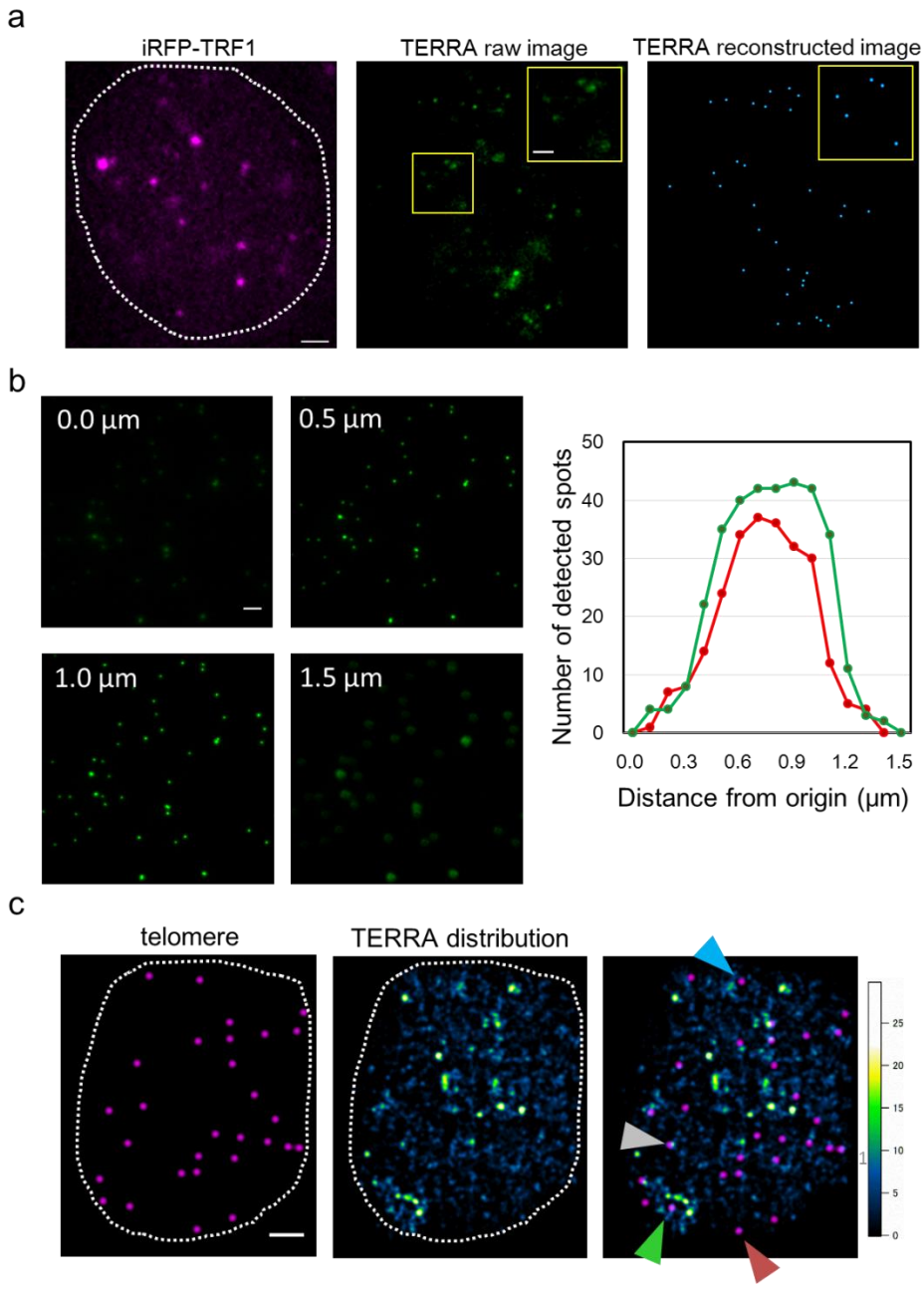

Figure 4-1. Single spot detection of TERRA in a living cell

(a) Live-cell imaging of TERRA molecules. Fluorescence images of U2OS cells expressing iRFP-TRF1 (left), the TERRA probe (middle), and a single-particle reconstruction of TERRA from the middle image (right). The white broken line in the left panel shows the nucleus. Each Gaussian spot in the right image is reconstructed as a spot with 200 nm diameter. Scale bar, 2.0 μm. (b) Images of multicolor fluorescent beads at different focus positions. In each image, the focus position is shown at the upper left (the glass surface is set as 0.0 μm). Scale bar, 2.0 μm. Numbers of detected beads at various focus positions. More than 90% of the beads were detected within ± 200 nm of the surface, but the detection efficiency markedly decreased beyond this region. (c) Overlaid images of fluorescent spots representing telomeres (left) and TERRA (right). Time-integrated images were constructed to show the spatial distributions of telomeres and TERRA foci. The TERRA density is shown in pseudo-color. Scale bars, 4.0 μm.

To determine whether the foci represent the accumulation of TERRA, I observed temperature-dependent alternation of TERRA foci. A previous study showed that thermal stress on mouse embryonic fibroblast cells induces transient increment in the number of TERRA foci [100]. To demonstrate the ability of the probe to visualize thermal stress-induced alternation of TERRA foci, I prepared three cell samples cultured in different temperature conditions; the first was cultured continuously at 37°C, the second was heat shock at 42°C for 1 hour, and the last was cultured at 37°C after the heat shock (**Figure 4-2 a**). From the fluorescence microscopy observation, I constructed time-integrated image by accumulation of time frames to elicit TERRA foci (**Figure 4-2 b**). The treatment of cells at 42°C for 1 h resulted in an increase in the number of TERRA foci in the nucleus (**Figure 4-2 c**). When the temperature was shifted back to 37°C, the

number of TERRA foci decreased to the initial levels (**Figure 4-2 c**). These results demonstrated that the TERRA probe is applicable to analyze dynamics changes in TERRA foci upon different external conditions in living cells.

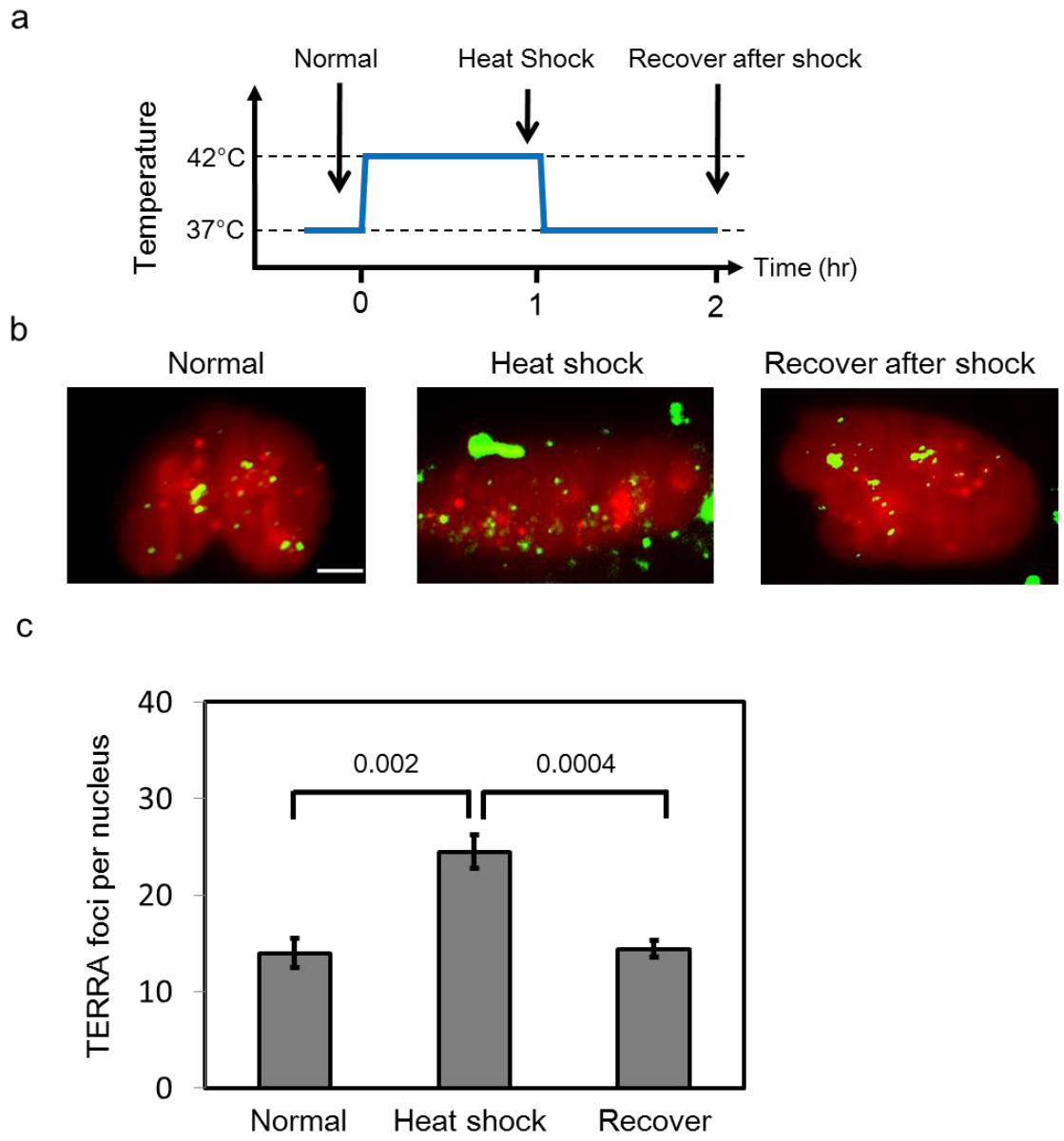

Figure 4-2. Heat shock response of TERRA foci

(a) Scheme of temperature incensement during the heat shock. (b) The distribution of TERRA foci constructed from time integrated image. TERRA foci are shown in green; telomere in red. The scale bar represents 4 μm. (c) Number of TERRA foci in different temperature. Mean numbers of TERRA foci per nucleus (mean ± s.e.m) of each condition are indicated.

4-3-2. Quantification of spatial distribution of TERRA

To quantify the distance between a TERRA focus and a telomere, I calculated a radial distribution function ($\rho(r)$) that evaluates the TERRA density in a region at a specific distance from the telomere center (**Figure 4-3 a**). Upon averaging $\rho(r)$ for all telomeres in the observations, two peaks of condense TERRA density were detected at distance of 0.25 and 1.00 μm ($p < 0.05$, **Figure 4-3 b**). The peak at 0.25 μm indicates the localization of TERRA focus to a telomere. On the other hand, the peak at r = 1.00 μm suggests TERRA concentrates near a telomere. To confirm the specificity of the peak at r = 1.00 μm, I calculated $\rho(r)$ of which the origin (r = 0.00 μm) was selected arbitrary positions outside of the telomeres. Under the condition, a peak was detected only at r = 0.25 μm (**Figure 4-3 c**). In addition, the $\rho(r)$ of hnRNPA1 was calculated, revealing no $\rho(r)$ peak for hnRNPA1 around a telomere (**Figure 4-3 d**). These data suggest that the peak at r = 1.00 μm is characteristic of the TERRA distribution around a telomere. Moreover, the $\rho(r)$ calculated for individual telomeres showed inhomogeneous distributions of TERRA around telomeres (**Figure 4-3 e**). Here, local maximum of $\rho(r)$ was defined as greater than twice the value at $\rho(r > 3.50$ μm). Based on this definition, the $\rho(r)$ was categorized into four patterns: no peak (41%); a single peak at a telomere ($r < 0.50$ μm, 20%); a single peak near telomere ($0.75 < r < 2.00$ μm, 29%); and double peaks at and near a telomere (10%, **Figure 4-3 e**). These results indicate that 39% of telomeres were located within 1–2 μm of a TERRA focus (**Figure 4-3 f**).

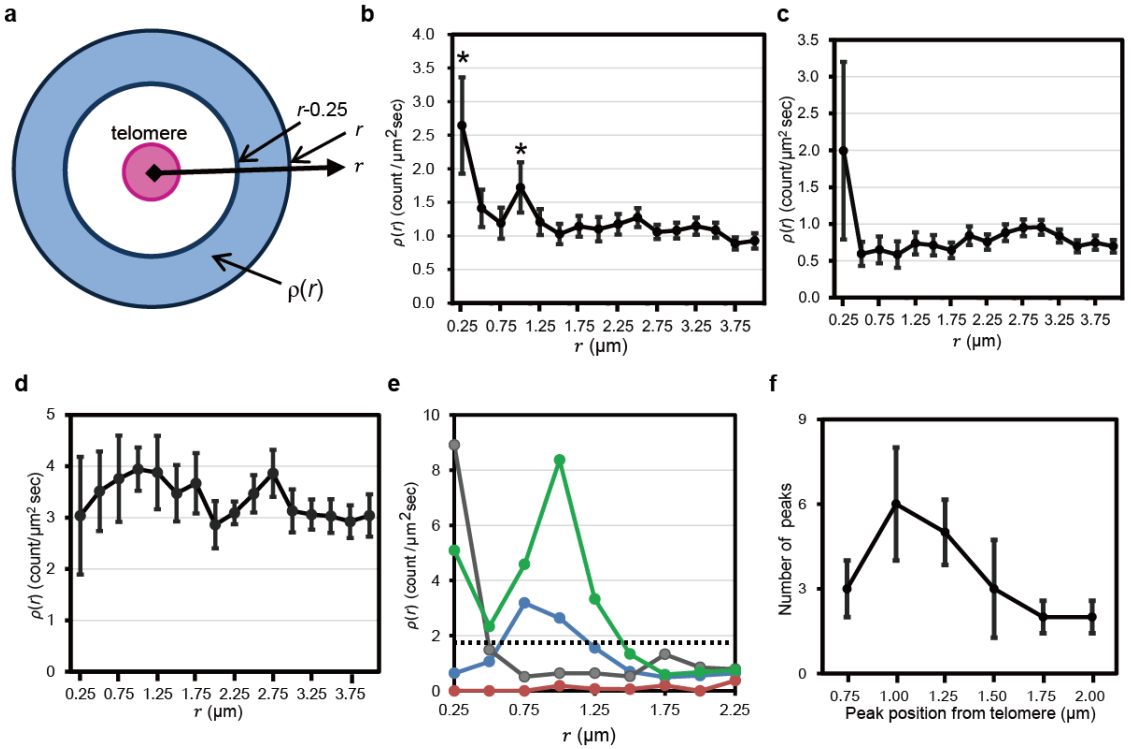

Figure 4-3. Analysis of TERRA distribution

(a) Schematic of the radial distribution function, ρ(*r*). ρ(*r*) is the number of TERRA spots divided by the observation time and area within the region between *r*-0.25 and *r* (μm) (blue) from the center of a reference telomere (magenta). (b) Averaged ρ(*r*) of the TERRA spots for all the telomeres in the cell shown in Figure 2A (±s.e.m., n = 26). ρ(*r* = 0.25 μm) and ρ(*r* = 1.00 μm) exceed ρ(*r* > 3.00 μm) (Student's *t*-test, $p < 0.05$). (c) Averaged ρ(*r*) of the TERRA spots in Figure 4-1c (±s.e.m., *n* = 30) for randomly selected positions. When a position was selected at a TERRA-accumulated spot, ρ(*r* = 0.25 μm) exceeded ρ(*r* > 0.25 μm). (d) Averaged ρ(*r*) of hnRNPA1 (±s.e.m., *n* = 26) for the telomeres in Figure 4-1 c. (e) Plots of ρ(*r*) against *r* for the four individual telomeres shown in Figure 2b. The dotted line represents the threshold for the peak (ρ(*r*) = 1.9). (f) Values for peak positions in ρ(*r*) (*n* = 117) in single cells are shown as a function of the distance from a telomere.

4-3-3. Quantitative analysis of TERRA motion

A notable application of the TERRA probe is analyzing the kinetics of TERRA dynamics at the single molecule level in living cells. To quantify the individual TERRA motions, we tracked the motion of EGFP spots. The motion of each TERRA spot was characterized by its diffusion coefficient (D_{TERRA}). To calculate D_{TERRA} mean-squared displacement (MSD) of the trajectories was calculated as a function of time lag, Δt. The D_{TERRA} values for individual TERRA molecules were widely distributed, ranging from 5.90×10^{-4} to 1.28 $\mu m^2/s$, indicating the existence of multiple modes of TERRA motion. Due to the thermal fluctuation and detection error, D_{TERRA} of a spot restricted at a specific position is not zero. To define a stationary mode, I estimated the apparent diffusion coefficients (D_{fix}) of TERRA in fixed cells. The distribution of D_{fix} was fitted with a lognormal function for which the 99% confidence limit was 5.53×10^{-2} $\mu m^2/s$ (**Figure 4-4 a**). Based on this result, stationary mode was defined as the motion of TERRA spots in living cells with D_{TERRA} values lower than 5.53×10^{-2} $\mu m^2/s$ (**Figure 4-4 b**). From this classification, we found that 29% of the total TERRA were in the stationary mode, whereas the remainder was in a mobile mode. The D_{TERRA} distribution of the mobile TERRA was fitted with a lognormal distribution (mean; 0.15 $\mu m^2/s$, variance; 0.02 $\mu m^2/s$, **Figure 4-4 b**). The mean D_{TERRA} was of the same order of magnitude as that previously reported on poly(A) RNA in living cells [101, 102], suggesting that TERRA in the mobile mode diffuses freely in a manner similar to that of other RNAs in the nucleus. Notably, two modes of TERRA were present at telomeres ($r < 0.50$ μm): a stationary mode (38%) and a mobile mode (62%, **Figure 4-4 c, d**). Numerous

stationary TERRA molecules were localized to telomeres during the observation period (**Figure 4-4 c**). In contrast, mobile TERRA molecules colocalize transiently with a telomere for 0.1–0.3 s and were largely diffuse in the nucleoplasm (**Figure 4-4 c**). Moreover, the ratio of stationary TERRA in the telomere-neighboring region ($0.75 < r < 2.00$ μm) was 42%, which was higher than that in the nucleoplasm (29%, $p < 0.001$), suggesting that the motions of TERRA are confined in the telomere-neighboring region.

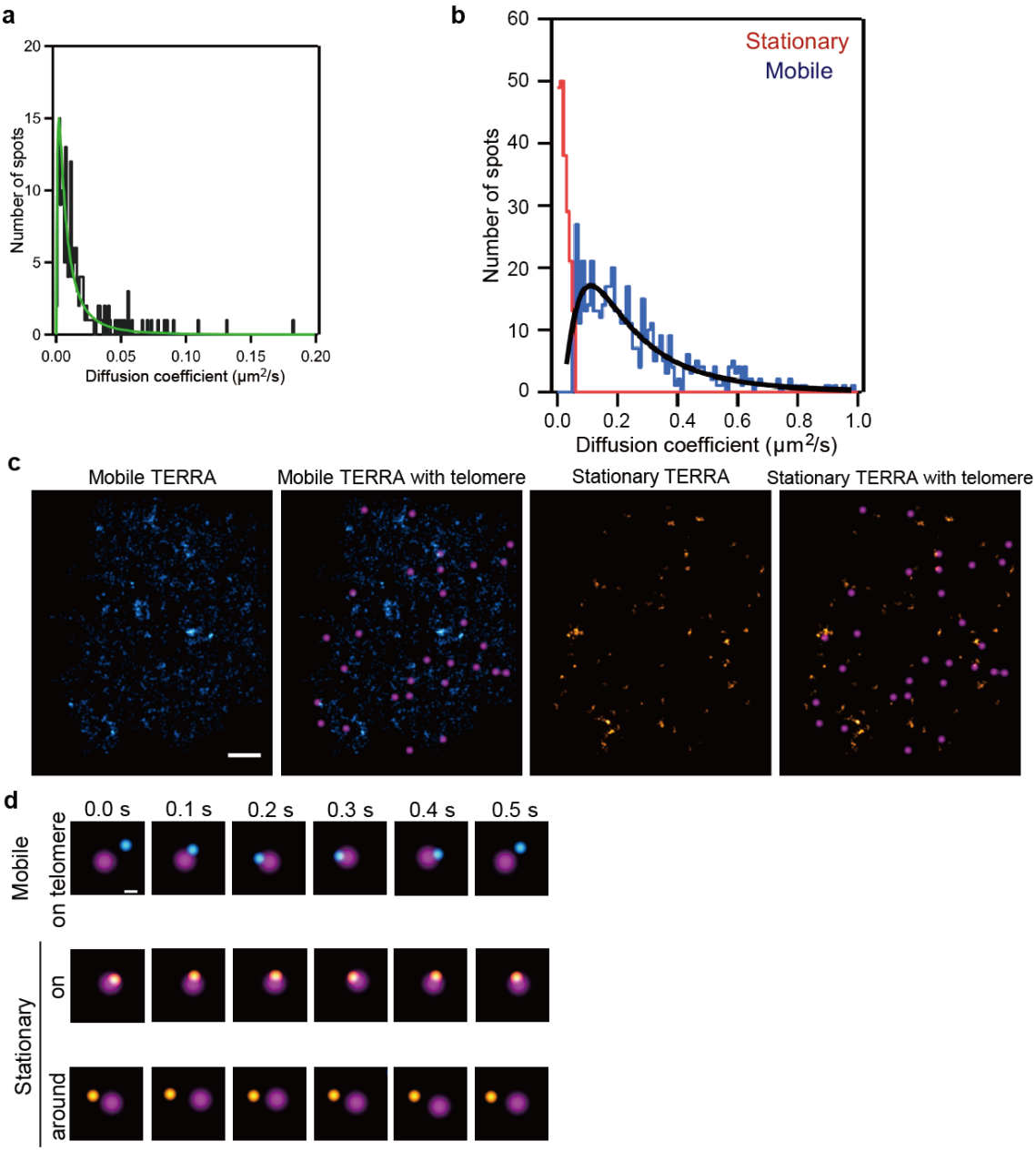

Figure 4-4. Analysis of TERRA motion

(a) Histogram of the diffusion coefficients of fluorescent spots in chemically fixed cells. The distribution of the apparent diffusion coefficients of EGFP spots (n = 181) in chemically fixed cells. Histograms are fitted with lognormal distributions with mean values of 5.32×10^{-3} μm^2/s (variance, 9.33×10^{-3} μm^2/s). (b) Histogram of the diffusion coefficient distribution of stationary TERRA (red, n =200) and mobile TERRA (blue, n=502) with a minimum track length of 7 frames (700 ms). Distribution of the diffusion coefficients of mobile TERRA was fitted with a lognormal function (black line). (c) Spatial distribution of TERRA in mobile and stationary phases. Mobile TERRA spots

with D > 5.53×10^{-2} $\mu m^2/s$ are shown in cyan. The spots in the stationary phase are shown in orange. The images are overlaid with the localization of telomeres (magenta). Scale bar, 2.0 μm. (d) Sequential images of stationary and mobile TERRA on or near a telomere. Scale bars, 0.5 μm.

4-3-4. Live cell imaging of TERRA, hnRNPA1, and telomere

Most of the models of TERRA functions are based on the hypothesis that TERRA works as a scaffold at a telomere [83, 90]. However, the existence of stationary TERRA around telomeres suggests another mechanism of the TERRA function. Recently, it was suggested that TERRA inhibits the localization of hnRNPA1 to telomeres, which cannot be explained using the scaffold model [91]. To assess the possibility of inhibitory activity of TERRA on hnRNPA1, we examined the dynamics of TERRA, hnRNPA1, and telomeres in living cells. Expression plasmids for the TERRA probe and SNAPtag-fused hnRNPA1 were transfected into cells stably expressing iRFP-TRF1. After staining the SNAP-hnRNPA1 with TMR, we observed the three factors by their respective fluorescence signals. As analyzed in TERRA, we classified the motions of hnRNPA1 into stationary or mobile modes (**Figure 4-5 a,b**). This classification revealed that only 29% of the hnRNPA1 on telomeres were stationary and that the remainder were mobile (r < 0.50 μm, **Figure 4-5 c**), indicating that localization of hnRNPA1 to telomeres is transient.

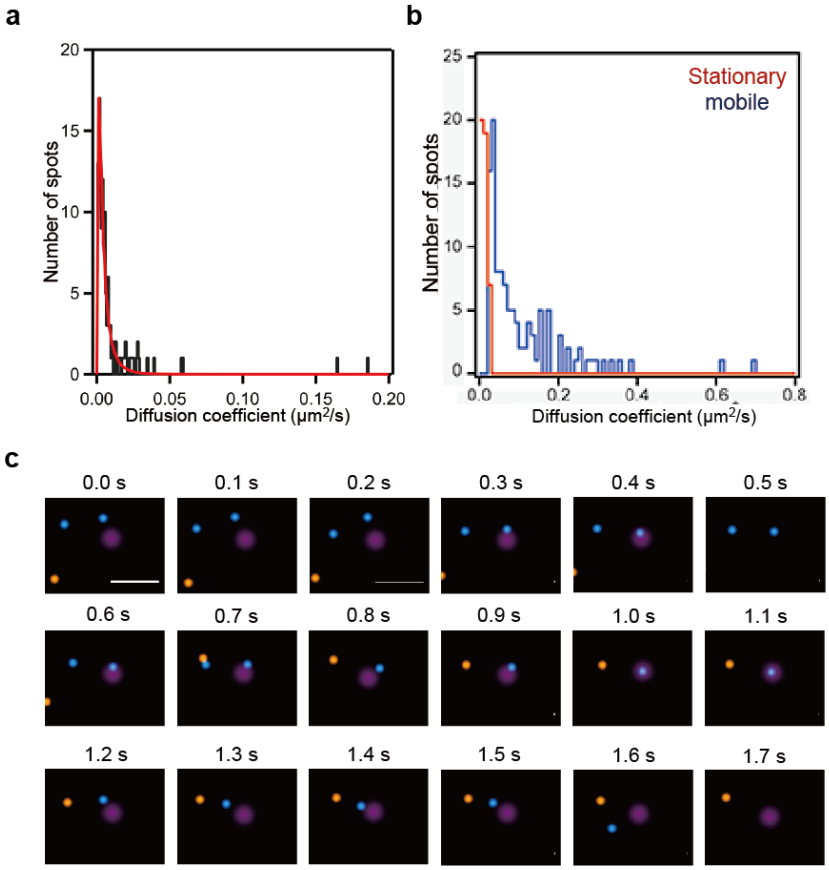

Figure 4-5. Analysis of TERRA motion

(a) Histogram of the diffusion coefficients of fluorescent spots in chemically fixed cells. The distribution of the apparent diffusion coefficients of TMR spots (b, n = 141) in chemically fixed cells. Histograms are fitted with lognormal distributions with mean values of 2.27×10^{-3} μm^2/s (variance, 3.64×10^{-3} μm^2/s) for TMR. (b) Distributions of the diffusion coefficients of stationary hnRNPA1 (orange, n = 46) and mobile hnRNPA1 (cyan, n = 112) on telomeres. (c) Sequential images of mobile hnRNPA1 localized transiently to a telomere (magenta). Scale bar, 1.0 μm.

4-3-5. Definition of TERRA-hnRNPA1 complex in living cells

Next, we calculated the distance between TERRA and hnRNPA1 to detect TERRA-hnRNPA1 complexes in living cells. Colocalization of TERRA with hnRNPA1 was defined according to their localization precisions (32 nm for TERRA and 28 nm for hnRNPA1, **Figure 4-6 a**) and overlay accuracy between EGFP and TMR (30 nm). To ensure whether colocalization resulted from an accidental overlap or complex formation, I investigated the durations of individual colocalization events. To define the accidental overlap of two spots, I evaluated the colocalization of TERRA with H2A, which does not interact each other, lasts only 0.2 sec (**Figure 4-6 b**). In the case of TERRA-hnRNPA1 colocalization, the fraction of events in which the colocalization duration exceeded 0.2 sec was 26% of the total number of colocalization events. This ratio was significantly higher than that for colocalization with H2A (6.1%). Therefore, the colocalization of TERRA with hnRNPA1 lasting longer than 0.2 sec were defined as TERRA-hnRNPA1 complexes (**Figure 4-6 c, d**).

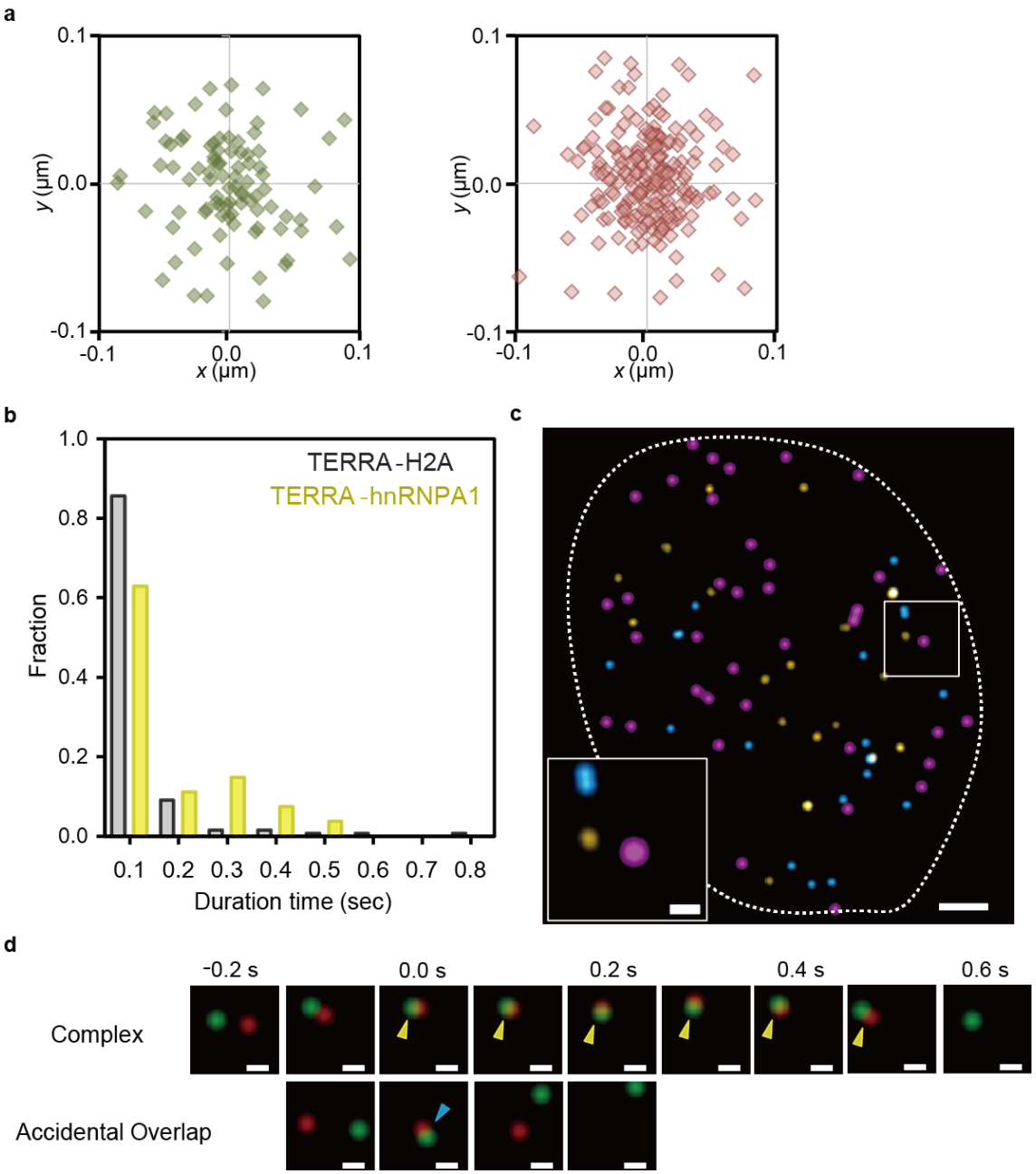

Figure 4-6. Definition of TERRA-hnRNPA1 complex formation

(a) Distribution of the centroid positions of individual EGFP (left) and TMR (right) spots in chemically fixed cells. Each spot gives a cluster of localizations because of the detection error during observations. The localizations of 95 clusters of EGFP and 205 clusters of TMR (indicated as diamonds) were aligned by their centers of mass to generate the overall 2D presentation of the localization distribution. (b) Distributions of the duration of TERRA colocalization with hnRNPA1 (yellow) and H2A (gray). In all, 130 colocalization events were analyzed for the respective proteins. (**c**) Spatial distribution of the colocalization between TERRA and hnRNPA1 in the nucleus.

TERRA-hnRNAP1 complexes are shown as yellow spots. Accidental overlaps are shown as cyan spots. Telomeres are represented by magenta spots. Scale bars, 2.0 μm and 0.5 μm (inset). (**d**) Sequential images of the colocalization of a TERRA spot (green) and an hnRNPA1 spot (red). Overlapping events that exceeded 0.2 s (yellow arrowheads) show TERRA-hnRNPA1 complex formation. Overlap sustained only for 0.1 s (cyan arrowhead) was regarded as an accidental overlap. Scale bars, 0.1 μm.

4-3-6. Distribution and dynamics of TERRA-hnRNPA1 complexes

To analyze the spatial distribution of TERRA-hnRNPA1 complexes, the percentage of the complexes was shown against the distance from a telomere (**Figure 4-7 a**). Importantly, the plot exhibited a local maximum in the telomere-neighboring region ($0.5 < r < 1.5$ μm). We also evaluated the colocalization of TERRA and freely diffusing spots generated by computer simulation. The spatial distribution of the colocalization between TERRA and the freely diffusing spots was homogeneous around telomeres (**Figure 4-7 a**). These results indicate that the formation of TERRA-hnRNPA1 complex occurs intensively in the telomere-neighboring region. Although both TERRA and hnRNPA1 localized to telomeres (**Figure 4-4, 4-5**), the frequency of the complex formation on a telomere was less than 0.1 (**Figure 4-7 a**), which was comparable to that in the nucleoplasm. Based on these results, I hypothesized that TERRA in telomere-neighboring areas, rather than that at a telomere, is responsible for the delocalization of hnRNPA1. To test this hypothesis further, I investigated whether hnRNPA1 localization to a telomere depends on the presence of TERRA in the telomere-neighboring region. In the presence of TERRA in a telomere-neighboring region, the number of hnRNPA1 spots on the telomere was lower by half ($p < 0.05$, **Figure 4-7 b**) than that on a telomere without TERRA around it.

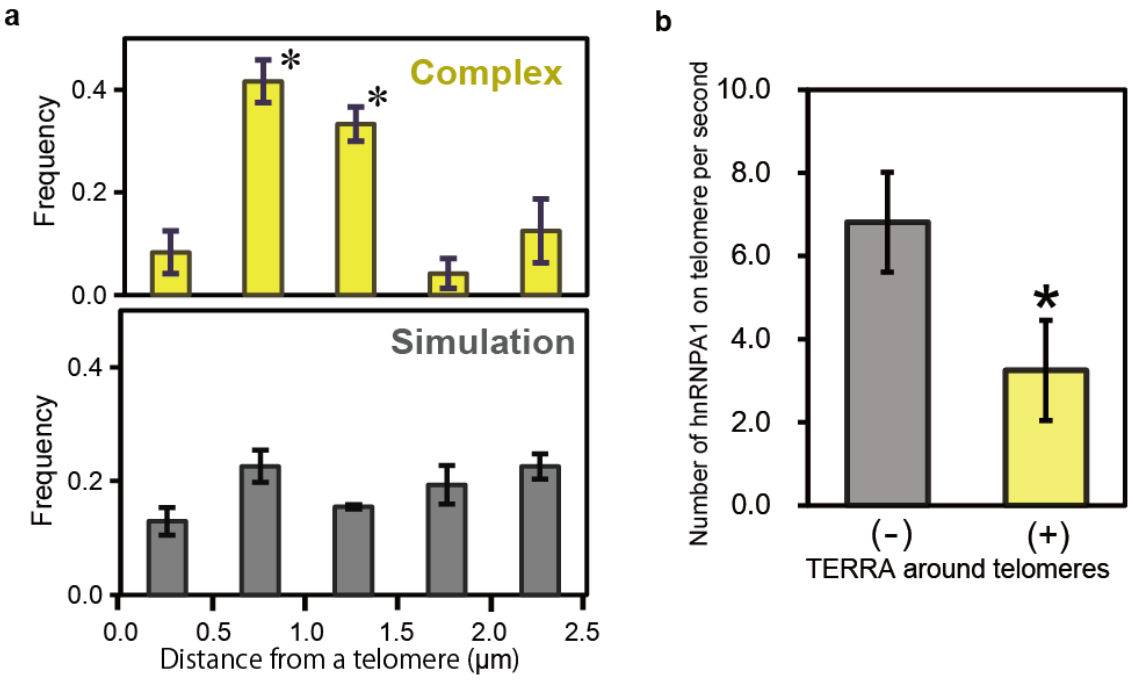

Figure 4-7. TERRA-hnRNPA1 complex around telomeres

(a) Fraction of colocalized regions shown as a function of the distance from a telomere: yellow, TERRA-hnRNPA1 complex formation (* indicates $p < 0.05$, Student's *t*-test); gray, simulated colocalization of TERRA with a freely diffusing molecule. (b) Correlation between hnRNPA1 localization to a telomere and TERRA accumulation around the telomere. The hnRNPA1 spots on telomeres were counted. During the observation, telomeres were categorized into two groups based on TERRA localization: TERRA localized to a telomere-neighboring area (TERRA around telomeres: +, $n = 13$) and no TERRA located around a telomere (TERRA around telomeres: -, $n = 14$). When TERRA was absent around a telomere, the number of hnRNPA1 occurrences was 6.8 spots/s, whereas this value was reduced by half when TERRA was present in the telomere-neighboring area (average: 3.2 spots/s, $p < 0.05$).

Moreover, I explored the ability of TERRA to inhibit hnRNPA1 localization to telomere by depleting endogenous TERRA. Analysis of real-time PCR revealed that two different siRNAs (siTERRA-1 and siTERRA-2) made the TERRA level depleted to ~40% of control levels (**Figure 4-8 a, b**)[93]. The effect of TERRA depletion on hnRNPA1 diffusion was examined by calculating the diffusion coefficients (**Figure 4-8 c**). We found no significant difference in the diffusion coefficients between control and TEERA knock down (KD, **Figure 4-8 c**), suggesting that TERRA has no effect on free-diffusing hnRNPA1 in nucleoplasm. Importantly, the number of hnRNPA1 spots on telomeres in the TERRA KD cells was twice larger than that on a telomere with TERRA around it in the control cells, and was comparable to that on a telomere without TERRA around the telomere (**Figure 4-8 d**). Consequently, I concluded that hnRNPA1 localization to telomeres is prevented when TERRA accumulates around the telomere and forms complexes with hnRNPA1.

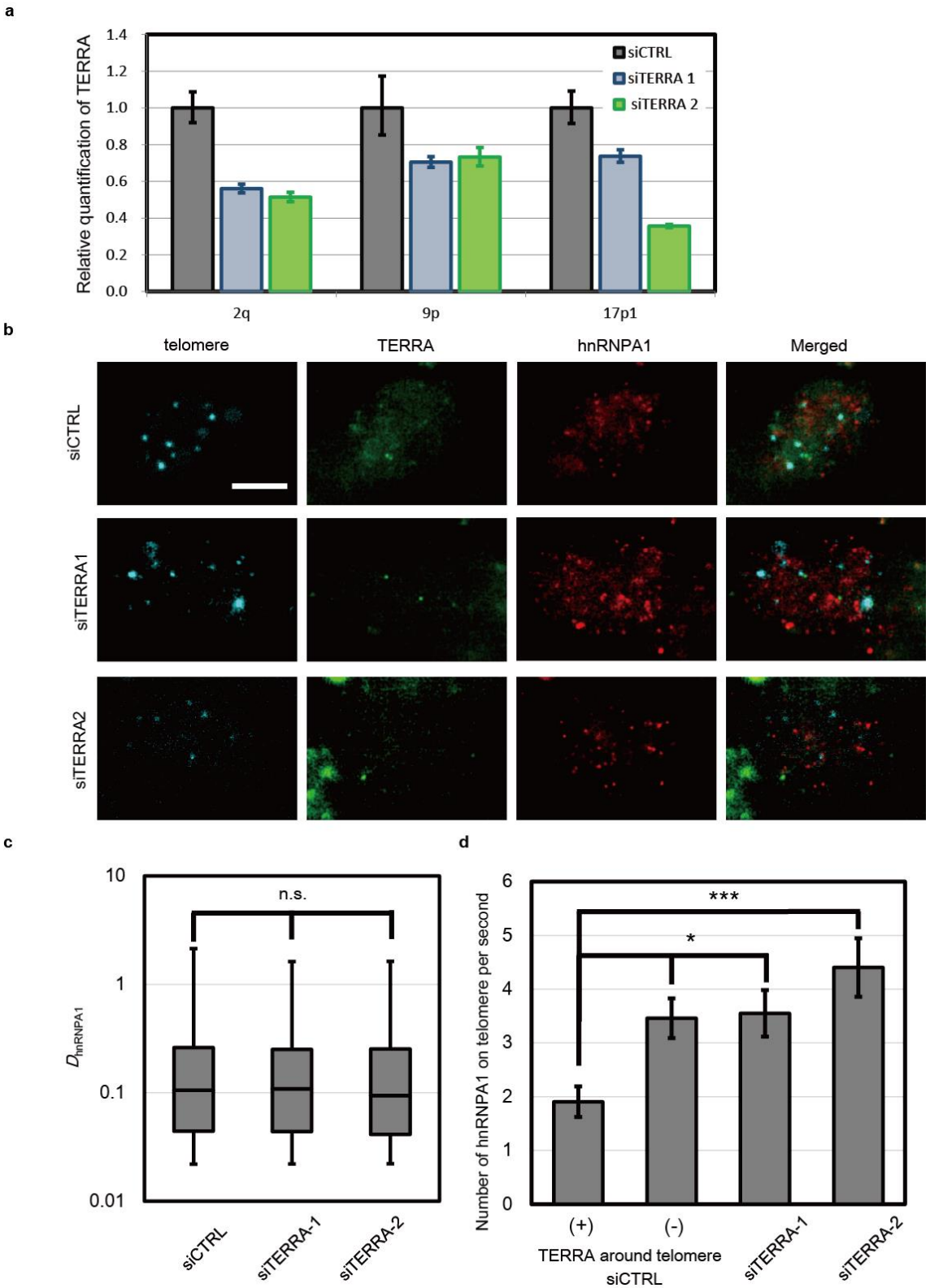

Figure 4-8. The effect of TERRA depletion on hnRNPA1 motion and distribution

(a) U2OS cells were transfected with siCTRL, siTERRA1, and siTERRA2. The amount of TERRA from individual subtelomeres were measured from total RNA and normalized to GAPDH mRNA. The bars represent the average values from three biological and two technical replicates for each sample. (b) Fluorescence images were obtained from U2OS cells expressing iRFP-TRF1 (telomere), the TERRA probe (TERRA), SNAP-hnRNPA1 labeled with TMR (hnRNPA1), and their merged images. The cells were transfected with siCTRL (upper), siTERRA-1 (middle), and siTERRA-2 (right). The scale bar represents 5.0 μm. (c) Quantification of the diffusion coefficients of hnRNPA1 in the cells transfected with each siRNA. n = 513 (siCTRL), 592 (siTERRA-1), 376 (siTERRA-2). (d) The number of hnRNPA1 spots observed at the telomeres in each cell condition. Mean numbers of hnRNPA1 spots per second (mean ± s.e.m) under each condition are indicated. In c and d, *P*-value was computed using the Student's *t*-test. n.s. indicates $P > 0.05$, * indicates $P < 0.05$, *** indicates $P < 0.001$.

To address the mechanism by which the TERRA near a telomere inhibits the localization of hnRNPA1 at the telomere, I analyzed the diffusion coefficients of TERRA and hnRNPA1 that form complexes mutually. Based on the analysis, only 24% of hnRNPA1 was categorized in the stationary mode, in contrast to 88% of TERRA in stationary mode (**Figure 4-9**). This result suggests that complex formation occurs primarily between stationary TERRA and mobile hnRNPA1 in telomere-neighboring regions.

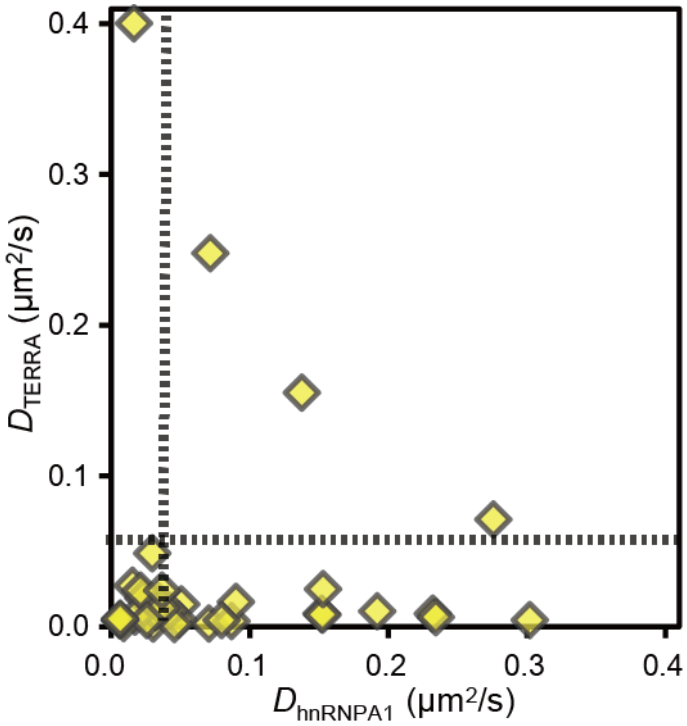

Figure 4-9. motion analysis of TERRA and hnRNPA1 which form the complex

Scatter plot of the diffusion coefficients of TERRA and hnRNPA1 including a period of complex formation within the trajectory. Broken lines show thresholds between the stationary and mobile phases.

4-4. Discussion

In this chapter, I analyzed the distribution and motion of TERRA quantitatively by using the TERRA probe which was developed and characterized in chapter 3. By observing TERRA, hnRNPA1, and telomeres in single living cells, I investigated how TERRA regulates the localization of hnRNPA1 to a telomere. First, I analyzed the density of TERRA around each telomere quantitatively. By calculating the radial distribution function of TERRA around a telomere, the TERRA distribution was categorized into four patterns: absence (40%); at a telomere (20%); near a telomere (0.75–2.00 μm distant from the telomere, 30%); both at and near a telomere (10%). Moreover, analysis of TERRA motion revealed that the population of the stationary TERRA near a telomere was 42%, which is similar to that at a telomere (38%) and higher than that in nucleoplasm (30%). These results suggest that TERRA significantly accumulated in the region apart (approx. 1.0 μm) from a telomere. It has been proposed that telomere-bound TERRA acts as a molecular scaffold to promote recruitment of telomere-binding proteins to the telomere [83, 103, 104]. Therefore, I hypothesized that TERRA may provide scaffolds for telomere-binding proteins in the telomere-neighboring region to regulate the accessibility of the proteins to the telomeres.

I tested the hypothesis by examining whether TERRA in the telomere-neighboring region regulates the localization of hnRNPA1 to the telomere [91, 92]. First, I quantified the density of hnRNPA1 around a telomere, which revealed that hnRNPA1 showed homogeneous distribution around a

telomere in contrast to the case of TERRA. Moreover, analysis of hnRNPA1 motion revealed that a large fraction of hnRNPA1 at telomeres was in mobile mode (71%). The mobile hnRNPA1 molecules were confined transiently at telomeres for only 0.1–0.3 s. These results suggest that binding of hnRNPA1 to a telomere is not static, but rather is in a certain equilibrium state. Intriguingly, the percentage of hnRNPA1 at a telomere decreased with the existence of TERRA in the telomere-neighboring region. These results suggest that TERRA in the telomere-neighboring region inhibits hnRNPA1 localization at the telomere.

I further investigated the mechanism underlying the negative correlation between TERRA near a telomere and hnRNPA1 at the telomere by analyzing the formation of TERRA-hnRNPA1 complexes. Notably, the complex formation occurred specifically in the telomere-neighboring region. This specific formation of TERRA-hnRNPA1 complex raised the possibility that TERRA near the telomere traps hnRNPA1 to shift the equilibrium toward an hnRNPA1-unbound state of the telomere. This idea was supported by results of analysis of TERRA and hnRNPA1 motions in the complexes. Analysis of the diffusion coefficient revealed that 88% of TERRA which formed the complex was in stationary mode, whereas only 24% of hnRNPA1 was in stationary mode, indicating that stationary TERRA traps the diffusing hnRNPA1 in the telomere-neighboring region.

Based on these results, I propose that these characteristic distributions and dynamics of TERRA correlate with its function: TERRA controls the on/off binding of hnRNPA1 to a telomere by acting as a scaffold for hnRNPA1 near the telomere (**Figure 4-10**). In this model, the binding of hnRNPA1 to a telomere

occurs at certain equilibrium in the absence of TERRA. When TERRA localizes to the telomere-neighboring region, the locally accumulated TERRA captures diffusive hnRNPA1, shifting the equilibrium toward the hnRNPA1-unbound state of telomere. Biochemical studies showed that incubation of hnRNPA1-coated telomere DNA with TERRA led to isolation of naked telomere DNA, suggesting that the presence of TERRA at telomeres promote the dissociation of hnRNPA1 from the telomere [91, 92]. Based on this result, they proposed a model by which hnRNPA1 shuttles between a telomere and TERRA dynamically when the concentration of TERRA is high. In this study, we refined this model further by single-particle analysis of TERRA and hnRNPA1: hnRNPA1 shuttles between a telomere and nucleoplasm even without the existence of TERRA; TERRA localized in a telomere-neighboring region, rather than at a telomere, interacts with hnRNPA1 and contributes to hnRNPA1 dissociation from the telomere.

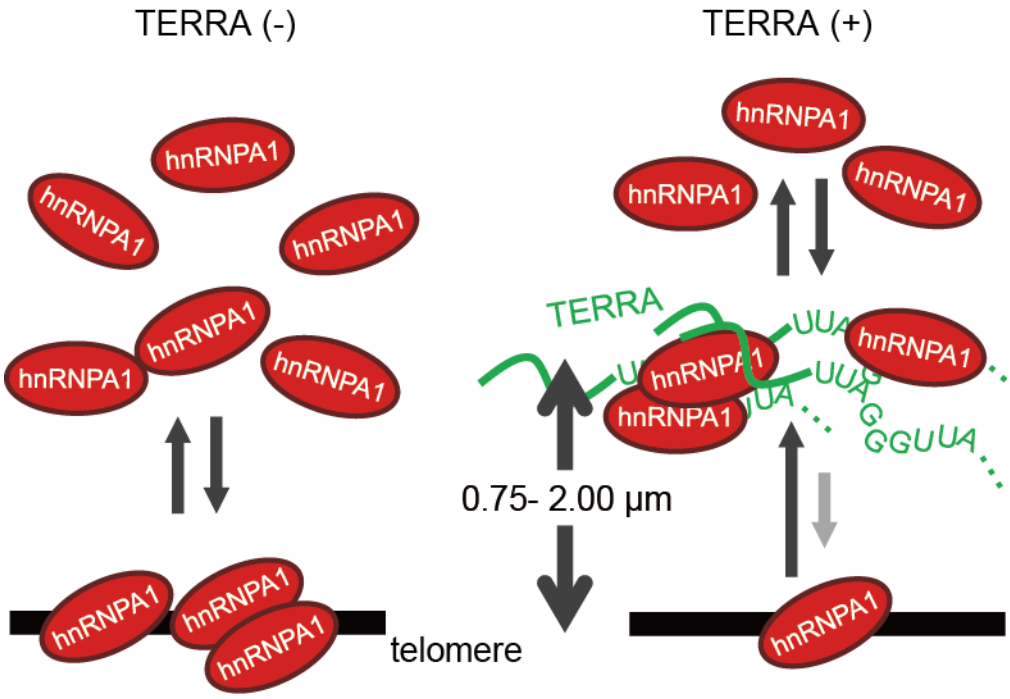

Figure 4-10. model based on the single-particle analysis of TERRA

Model showing how TERRA regulates the localization of hnRNPA1 on a telomere. When TERRA is absent around a telomere (shown as a black bar), the binding of hnRNPA1 to the telomere is at equilibrium (left). When TERRA localizes to the telomere-neighboring region (0.75–2.00 μm from the telomere center), the locally accumulated TERRA captures hnRNPA1, shifting its equilibrium toward the telomere-unbound form (right).

4-5. Conclusion

Here present live-cell imaging of TERRA at single-molecule level. By analyzing the TERRA and hnRNPA1 dynamics, I connected the distribution and motion of TERRA with its function. Here I like to emphasize the mechanism how TERRA, of which motion is governed by physical low, realize the function: local accumulation shift the equilibrium of its-interacting protein. TERRA has diverse function and the model presented in this study cannot explain the whole function. However, in principle, single-particle analysis of TERRA will be applicable to further study of TERRA.

4-6. Reference

89 Luke, B. and Lingner, J. (2009) TERRA: telomeric repeat-containing RNA. *EMBO J* 28, 2503-2510

90 Arnoult, N., *et al.* (2012) Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1α. *Nature Structual Molecular Biolgy* 19, 948-956

91 Flynn, R.L., *et al.* (2011) TERRA and hnRNPA1 orchestrate an RPA-to-POT1 switch on telomeric single-stranded DNA. *Nature* 471, 532-536

92 Redon, S., *et al.* (2013) A three-state model for the regulation of telomerase by TERRA and hnRNPA1. *Nucleic Acids Research* 41, 9117-9128

93 Feretzaki, M. and Lingner, J. (2016) A practical qPCR approach to detect TERRA, the elusive telomeric repeat-containing RNA. *Methods in press*

94 Koyama-Honda, I., *et al.* (2005) Fluorescence Imaging for Monitoring the Colocalization of Two Single Molecules in Living Cells. *Biophysical Journal* 88, 2126-2136

95 Gelles, J., *et al.* (1988) Tracking kinesin-driven movements with nanometre-scale precision. *Nature* 331, 450-453

96 Fujiwara, T., *et al.* (2002) Phospholipids undergo hop diffusion in compartmentalized cell membrane. *The Journal of Cell Biology* 157, 1071-1082

97 Suzuki, K.G.N., *et al.* (2007) GPI-anchored receptor clusters transiently recruit Lyn and Gα for temporary cluster immobilization and Lyn activation: single-molecule tracking study 1. *The Journal of Cell Biology* 177, 717-730

98 Cheezum, M.K., *et al.* (2001) Quantitative comparison of algorithms for tracking single fluorescent particles. *Biophysical Journal* 81, 2378-2388

99 Rust, M.J., *et al.* (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* 3, 793-796

100 Schoeftner, S. and Blasco, M.A. (2008) Developmentally regulated transcription of mammalian telomeres by DNA-dependent RNA polymerase II. *Nature Cell Biolgy* 10, 228-236

101 Politz, J.C., *et al.* (1999) Movement of nuclear poly(A) RNA throughout the interchromatin space in living cells. *Current Biology : CB* 9, 285-291

102 Ishihama, Y. and Funatsu, T. (2009) Single molecule tracking of quantum dot-labeled mRNAs in a cell nucleus. *Biochemical and Biophysical Research Communications* 381, 33-38

103 Azzalin, C.M. and Lingner, J. (2015) Telomere functions grounding on TERRA firma. *Trends in Cell Biology* 25, 29-36
104 Porro, A., *et al.* (2014) Functional characterization of the TERRA transcriptome at damaged telomeres. *Nature Communication* 5, 5379

Chapter 5

General conclusion

Through my PhD thesis, I developed a fluorescent probe based on sequence-specific RNA-binding domain, PUM-HD, and split EGFP fragments. The fluorescent probe developed in this study enable to visualize endogenous target RNA with single-molecule sensitivity. I applied this probe to β-actin mRNA which produced cytoskeletal actin protein and a long non-coding RNA transcribed from a telomere, telomeric repeat-containing RNA (TERRA). Single-molecule imaging of β-actin mRNA showed growth-factor dependent localization of β-actin mRNA to leading edge and directional movement along microtubules. Moreover, single-particle fluorescence imaging revealed that TERRA accumulated in a telomere-neighboring region and trapped diffusive heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1), thereby inhibiting hnRNPA1 localization to the telomere. This result suggests that TERRA regulates binding of hnRNPA1 to the telomere in a region surrounding the telomere, leading to a deeper understanding of the mechanism of TERRA function.

The hallmark of my PhD study is a proposal of a mechanistic model based on the analysis of RNA with single-molecule sensitivity. Currently, most models in life-science are based on the static molecular interaction *in vitro*, which lacks motions of individual RNAs and proteins involved in the models and sub-cellar distribution of molecular interaction. Although the localization and motion of the molecules are governed by physical laws, the dynamics in the models cannot be explained by simple Brownian motion or static confinement. As discussed in chapter 4, I showed that local concentration of TERRA around a telomere is critical for shifting the binding equilibrium of hnRNPA1 to the

telomere. This study is, to my knowledge, the first study that lncRNA function is explained by its distribution and motion at the single-molecule level.

In conclusion, this study demonstrates the potential of live-cell imaging of NRA in studying RNA functions of interest. Single-particle analysis reveals the relation between RNA dynamics and its function, establishing the mechanistic model based on the spatiotemporal information in living cells. The probe developed in this study can be broadly applicable for various RNAs because of its flexibility of PUM-HD recognition. The PUM-based probe with a single-particle tracking technique is a powerful tool for investigating the mechanistic insights into such RNA functions. It can reveal the significance of RNA dynamics in living cells.