Single-cell quantification of the concentrations and dissociation constants of endogenous proteins

Akira T. Komatsubara,†‡§¶1, Yuhei Goto,†§¶1, Yohei Kondo,†§¶, Michiyuki Matsuda,†‡§¶‡ and Kazuhiro Aoki†**

From the †Laboratory of Bioimaging and Cell Signaling, Graduate School of Biostudies, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan, the ‡Division of Quantitative Biology, National Institute for Basic Biology, National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji-cho, Okazaki, Aichi 444-8787, Japan, the §Quantitative Biology Research Group, Exploratory Research Center on Life and Living Systems, National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji-cho, Okazaki, Aichi 444-8787, Japan, the ¶Department of Pathology and Biology of Diseases, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan, the **Imaging Platform for Spatio-Temporal Information, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan, and the ***Department of Basic Biology, Faculty of Life Science, SOKENDAI (Graduate University for Advanced Studies), Myodaiji, Okazaki, Aichi 444-8787, Japan

This work was supported by the Spectrography and Bioimaging Facility and Functional Genomics Facility, NIBB Core Research Facilities. The authors declare that they have no conflicts of interest with the contents of this article.

Received for publication, January 23, 2019, and in revised form, January 30, 2019 Published, Papers in Press, February 9, 2019, DOI 10.1074/jbc.RA119.007685

Supporting Material

Kinetic simulation is a useful approach for elucidating complex cell-signaling systems. The numerical simulations required for kinetic modeling in live cells critically require parameters such as protein concentrations and dissociation constants ($K_d$). However, only a limited number of parameters have been measured experimentally in living cells. Here we describe an approach for quantifying the concentration and $K_d$ of endogenous proteins at the single-cell level with CRISPR/Cas9-mediated knock-in and fluorescence cross-correlation spectroscopy. First, the mEGFP gene was knocked in at the end of the mitogen-activated protein kinase 1 (MAPK1) gene, encoding extracellular signal-regulated kinase 2 (ERK2), through homology-directed repair or microhomology-mediated end joining. Next, the HaloTag gene was knocked in at the end of the ribosomal S6 kinase 2 (RSK2) gene. We then used fluorescence correlation spectroscopy to measure the protein concentrations of endogenous ERK2-mEGFP and RSK2-HaloTag fusion constructs in living cells, revealing substantial heterogeneities. Moreover, fluorescence cross-correlation spectroscopy analyses revealed temporal changes in the apparent $K_d$ values of the binding between ERK2-mEGFP and RSK2-HaloTag in response to epidermal growth factor stimulation. Our approach presented here provides a robust and efficient method for quantifying endogenous protein concentrations and dissociation constants in living cells.

In response to extracellular signals, mammalian cells process information through an intracellular signaling network comprised of chemical reactions, eventually leading to decisions regarding cell fate. Extensive studies of cell signaling have identified numerous proteins, pathways, and feedback or feedforward regulations and have expanded our understanding beyond the simple view of the linear signaling cascade (1). Computer-assisted systems biology approaches may provide a promising strategy for a comprehensive understanding of such complicated cell signaling networks (2). Indeed, various simulation models of signal transduction pathways have been developed in a series of studies over the past 10 years (3–5). Nevertheless, the numerical simulations in most of these systems were conducted with kinetic parameters that have not been measured experimentally; thus, these parameters tend to differ among models even when the reactions themselves are identical. For this reason, experimentally determined parameters are essential for the development of quantitative and reliable simulation models.

The parameters for simulation of cell signaling are roughly classified into four categories: protein concentration, dissociation constant, diffusion coefficient or transport rate, and enzymatic reaction rate parameters. Classically, these parameters have been measured by in vitro biochemical analyses, which require a large number of cells or molecules (6, 7). However, some of these parameters, e.g. protein concentration, are known to show nongenetic cell-to-cell variability (8, 9). Moreover, the quantities of some parameters might differ significantly in vitro and in vivo. For instance, we have shown that the ERK5 MAPK represents entirely different phosphorylation path-
terns, processive and distributive phosphorylation, in an intracellular and in vitro environment, respectively (7, 10). Therefore, it is of critical importance to measure parameters in living cells.

Fluorescence correlation spectroscopy (FCS) and fluorescence cross-correlation spectroscopy (FCCS) are techniques that exploit the fluctuations of fluorescent molecules in the confocal volume (~1 fl) (11–13). FCS examines the autocorrelation function of temporal fluorescence fluctuation, enabling us to determine the number of fluorescent molecules in the confocal volume and the diffusion coefficient. In FCCS, the cross-correlation function is calculated between fluctuations of two different fluorescent species to quantify the extent to which these two species form a complex. We have previously applied FCS and FCCS to mammalian cells exogenously expressing target proteins fused with EGFP and HaloTag and demonstrated that the association/dissociation patterns of the proteins differ from the predicted in vitro $K_d$ (14). Following previous studies, we refer to the effective dissociation constant in living cells as in vivo $K_d$. Note that in vivo $K_d$ includes the effects of posttranslational modification, competitive inhibition by other binders, and all intracellular environments such as molecular crowding, temperature, pH, ionic strength, etc. (15). Therefore, in vivo $K_d$ is a complex quantity that depends not only on the binding energy but also on concentrations of various competitive reactants and environmental conditions. This complicates clear thermodynamic interpretation of in vivo $K_d$ in contrast to true $K_p$, which indicates the affinity of protein–protein interaction under a certain experimental condition. However, measuring in vivo $K_d$, based on FCS and FCCS provides crucial information to quantitatively understand how protein circuits work in living cells.

It has been reported that the endogenous protein concentration and in vivo $K_d$ value were successfully measured in budding yeast with FCS and FCCS (16, 17). However, there have been no reports of the in vivo $K_d$ values based on the measurement of endogenous proteins in mammalian cells, mainly because of the technical difficulties of knock-in (KI) of a fluorescent protein gene to label the protein of interest.

Recent advances in genome editing tools have paved the way for tagging endogenous proteins with fluorescent proteins. These genome editing tools, such as the CRISPR/Cas9 system, enable KI of a gene of interest through DNA double-strand break (DSB) repair mechanisms (18, 19). Homology-directed repair (HDR) is a mechanism by which a homologous template is used as a source of DNA repair. On the other hand, microhomology-mediated end joining (MMEJ) is a mechanism of alternative nonhomologous end joining that also seals DSBs. In contrast to classical nonhomologous end joining, MMEJ repairs DNA DSBs using a 5- to 25-bp microhomologous sequence (20). HDR uses a relatively longer homologous sequence (0.1–10 kbp) to seal the DSB, ensuring error-free repair. Nonetheless, even though MMEJ is an error-prone process of end joining, a recent study demonstrated that MMEJ-mediated KI was more efficient than HDR (21). In this study, we demonstrate a new approach to quantifying the concentration and $K_d$ of endogenous proteins by combining FCS and FCCS with

---

**Figure 1. Outlines of gene KI.** A, schematic of gene KI with an HDR-mediated or MMEJ-mediated KI vector at the human MAPK1 locus. L, left; R, right; A, procedure for mEGFP KI. HeLa cells were transfected with the donor vector and Cas9 vector at day 0, and transfected cells were selected by puromycin. The cell were then subjected to G-418 selection for ~10 days. After G-418 selection, each GFP-positive cell was sorted into a 96-well plate by a FACSAria IIu flow cytometer (BD Biosciences). Single clonal cells were infected with AAV-Cre to remove the P2A-dTKneo-poly(A) (pA) cassette. CRISPR/Cas9-mediated KI of fluorescent proteins in mammalian cells.

### Results

**Design of donor vectors and KI strategy**

First, we attempted to knock in the mEGFP gene at the 3’ site of the human MAPK1 gene encoding ERK2 in HeLa cells. To do this, we constructed a selection cassette for the donor vector that contained the mEGFP gene, a loxP sequence, a porcine teschovirus-1–derived self-cleaving 2A peptide (P2A) sequence, a bifunctional fusion protein between a truncated version of herpes simplex virus type 1 thymidine kinase (dTK) gene, the bacterial neomycin phosphotransferase (neo) gene (dTKneo), a poly(A) addition sequence, and a loxP sequence (Fig. 1A and
CRISPR/Cas9 vector and donor DNA into HeLa cells and then selection cassette had been removed successfully (Fig. 1). Ganciclovir treatment selected cells whose recombination removed the P2A, dTKneo, and poly(A) addition sequence. Of note, it has been reported that neo is preferable for targeting moderate- or low-expression genes with the promoterless targeting vector (25). After KI of the donor vector, Cre-mediated recombination removed the P2A, dTKneo, and poly(A) addition sequence. Ganciclovir treatment selected cells whose selection cassette had been removed successfully (Fig. 1A).

In our experimental schedule, we transfected the pX459 CRISPR/Cas9 vector and donor DNA into HeLa cells and then selected the transfected cells with 1.0 μg/ml puromycin treatment. Three days after transfection, we began to select KI cells with 0.5 mg/ml G418. After selection, the KI cells were sorted by flow cytometry based on mEGFP fluorescence, followed by single-cell cloning for an additional 14 days. The cells were then infected with an adenovirus (AAV), which transiently induced the expression of Cre recombinase, to remove the P2A-dTKneo cassette. One week after AAV infection, the cells were further subjected to negative selection with ganciclovir.

Establishment of KI HeLa cells expressing ERK2-mEGFP

As a proof of concept, we targeted the MAPK1 gene to generate cells expressing ERK2-mEGFP from the endogenous locus. HeLa cells were transfected with pX459-hMAPK1 and a donor vector for HDR- or MMEJ-mediated KI (Fig. 1A). After G418 selection, the genomic DNA was extracted from parental HeLa cells or bulk HeLa cells transfected with the pX459-hMAPK1 and HDR or MMEJ donor vector and subjected to PCR with either the forward (F) primer, the reverse (R) primer or both primers (FR). As expected, PCR products were only observed in the sample from cells introduced with the HDR donor or MMEJ donor when we used both the forward and reverse primers (Fig. 2A), indicating KI of donor vectors into the 3' site of the MAPK1 gene.

Next we used flow cytometry to sort the transfected HeLa cells into two populations based on their GFP fluorescence intensities: a low GFP fluorescence (Low) and high GFP fluorescence (High) group (Fig. 2B). The Low and High cells accounted for 17.2% and 2.0% of the parent population in HeLa cells transfected with the HDR donor vector, respectively (Fig. 2B). 5.6% and 0.1% of the parent population in cells transfected with the MMEJ donor vector were categorized as Low and High cells, respectively. The results of immunoblotting with an anti-GFP antibody and an anti-ERK1/2 antibody demonstrated the expression of ERK2-mEGFP with a molecular mass of ~69 kDa (Fig. 2C, arrowhead) in both Low and High cells transfected with the HDR or MMEJ vector. Intriguingly, in both HDR- and MMEJ-mediated KI, High cells showed stronger nonspecific mEGFP integration bands and weaker specific ERK2-mEGFP integration bands than Low cells (Fig. 2C).

We subsequently isolated several single clones of Low cells by the limiting dilution method. Immunoblotting with the anti-
ERK1/2 antibody demonstrated that isolated cell lines expressed comparable levels of ERK2-mEGFP (Fig. S2A). Western blot analysis revealed that, compared with parental HeLa cells, some of the isolated cells showed only a positive signal of ERK2-mEGFP around 69 kDa (cell lines 13, 16, and 20 with the HDR donor vector and cell lines 1, 3, 5, 10, and 12 with the MMEJ donor vector). The other cells exhibited not only the positive ERK2-mEGFP signal but also apparently negative GFP signals, indicating higher or lower molecular weights than expected (Fig. S2A). Thus, we picked up the former cell clones and confirmed that the positive ERK2-mEGFP signals were also detected with anti-ERK2 antibody (Fig. S2B). The K1 efficiencies were roughly 15% (3 of 20) and 40% (5 of 12) with the HDR and MMEJ donor vectors, respectively (Fig. S2).

For further analysis, HDR-mediated KI cell line 13 (HDR 13) and MMEJ-mediated KI cell line 10 (MMEJ 10) were infected with an AAV expressing Cre recombinase, followed by analysis of expression levels of ERK2-mEGFP (Fig. 2D). Removal of the P2A-dTKneo-poly(A) cassette increased the ERK2-mEGFP levels in both cell lines (2.54-fold increase) (Fig. 2D, left panel). The amount of endogenous ERK2 in control cells was 8.05-fold that of ERK2-mEGFP (MMEJ 10). Of note, the expression levels of endogenous ERK2 did not show a discrete reduction, i.e. a 50% or 100% reduction, but rather exhibited a 40% to 80% reduction in the KI cell lines (Fig. 2D, right panel), suggesting that mEGFP genes were knocked into one or two of the multiple MAPK1 gene loci in HeLa cells, which are known to demonstrate aneuploidy (26).

Establishment of RSK2-HaloTag KI HeLa cells

Because RSK2 is known to interact with ERK2 (27), we next targeted the RSK2 gene for labeling with HaloTag, which covalently linked to HaloTag ligands such as the tetramethylrhodamine (TMR) ligand (28). For HaloTag gene integration, we tested the aforementioned MMEJ donor and precise integration into the target chromosome (PITCh)—like KI vector (21) (Fig. 3A). The PITCh vector includes the insertion cassette, which is further sandwiched by recognition sites of the Cas9—single guide RNA (sgRNA) complex that target the RSK2 gene locus (Fig. 3A, bottom panel). Therefore, the Cas9–sgRNA complex digests not only the RSK2 gene locus but also the PITCh-like donor vector, generating linear double-strand donors within the cells.

To obtain HeLa cells expressing ERK2-mEGFP and RSK2-HaloTag from this endogenous locus, we introduced Cas9 and the sgRNA expression vector (pX459-hRSK2), which targeted a site close to the stop codon of the RSK2 gene, and linearized the MMEJ donor vector (Fig. 3A, top panel) or the PITCh-like donor plasmid (Fig. 3A, bottom panel) into HeLa/ERK2-mEGFP (MMEJ#10) cells (Fig. 2D). After G418 selection and single-cell cloning, several clones were analyzed by PCR using two primers that recognized the 3’ end of RSK2 or the 5’ end of HaloTag (Fig. 3B). Almost half of the clones used in either donor vector showed the expected PCR product (Fig. 3B). However, subsequent DNA sequencing of these PCR products revealed the repeated sequence of the homology arm in PITCh 1, 2, 6, and 7 cells, resulting in a frameshift. MMEJ 1, 4, 6, and 7 cells had no insertion or deletion. In agreement with these data, expression of RSK2-HaloTag was observed only in MMEJ 1, 4, 6, and 7, and PITCh 3 cells by immunoblotting with the anti-RSK2 antibody (Fig. 3C). As with the case of ERK2-mEGFP, removal of the selection marker also enhanced expression of the RSK2-HaloTag (1.89-fold increase) except in MMEJ 4 cells (Fig. 3D). The expression level of the RSK2-HaloTag was 5.14-fold lower than the expression of RSK2 in control cells (Fig. 3D).

Figure 3. HaloTag KI at the human RSK2 locus. A, schematic of the MMEJ-mediated KI vector (top panel) and PITCh-like KI vector (bottom panel) for HaloTag gene KI. The PITCh vector contains the sequence targeted by sgRNA, which recognizes the RSK2 gene. B, after G-418 selection, genomic DNAs were extracted and analyzed by PCR to confirm HaloTag integration. Eight and nine clones were analyzed in clonal HeLa cells transfected with the MMEJ-mediated KI vector (MMEJ, top panel) and PITCh vector (PITCh, bottom panel), respectively. The forward primer and the reverse primer bind to RSK2 and mHaloTag, respectively. The arrowheads indicate the expected PCR products. C, the cell lysates were obtained from parental HeLa cells (−) and the indicated clones of HeLa/RSK2-HaloTag-MMEJ (MMEJ) and HeLa/RSK2-HaloTag-PITCh (PITCh) cells in B and subjected to immunoblotting with an anti-RSK2 antibody. D, cell lysates were obtained from parental HeLa cells (−) and the indicated clones of HeLa/RSK2-HaloTag-MMEJ (MMEJ) and HeLa/RSK2-HaloTag-PITCh (PITCh) cells before (−) or after (+) AAV-Cre infection and subjected to immunoblotting with anti-RSK2 (top panel) and anti-tubulin (bottom panel) antibodies.
Live-cell measurement of concentration and in vivo $K_d$

Subcellular localization of ERK2-mEGFP and RSK2-HaloTag KI HeLa cells

It is well-known that mitogen stimulation induces nuclear accumulation of ERK, causing phosphorylation of several transcription factors and subsequent gene expression (29–31). We picked up the cell line PITCh 3 (HeLa/ERK2-mEGFP/RSK2-HaloTag) and visualized the RSK2-HaloTag by staining with 100 nM TMR ligand. As expected, treatment with FBS and epidermal growth factor (EGF) provoked nuclear translocation of ERK2-mEGFP from the cytoplasm but did not induce subcellular translocation of RSK2-HaloTag-TMR within 30 min (Fig. 4, A and B). The kinetics of nuclear translocation of ERK2-mEGFP were slightly delayed compared with that of endogenous ERK1/2 observed by immunofluorescence (Fig. S3, A and B), suggesting that fusion with mEGFP had a negative effect on the nucleocytoplasmic shuttling of ERK2. RSK2-HaloTag-TMR showed nuclear translocation by FBS stimulation for 4 h (Fig. S3, C and D), consistent with a previous study (32). Fluorescence speckles in KI cells were also observed in parental HeLa cells under our experimental condition (Fig. S3E). Taken together, these data provide evidence that ERK2-mEGFP and RSK2-HaloTag can be imaged by fluorescence microscopy.

Evaluation of measurement noise and intra- and intercellular variability of protein concentration measured by FCS

To quantitatively address the efficacy of endogenous protein measurement by FCS, we assessed the variability derived from labeling efficiency, measurement noise, and intra- and intercellular heterogeneity. First, we found that the labeling of HaloTag-TMR-ligand was saturated by 6 h incubation under our experimental condition (Fig. 5A). Because the HaloTag protein is known to covalently bind to the HaloTag ligand, we assumed nearly 100% labeling efficiency by 6 h. Under this condition, mEGFP-HaloTag-TMR fusion proteins were transiently expressed in HeLa cells and analyzed by FCS (Fig. 5B). The number of mEGFP proteins obtained by FCS was plotted as a function of the number of HaloTag-TMR proteins, showing a linear correlation with a slope of 0.92 (Fig. 5C). If we assume 100% labeling efficiency of HaloTag-TMR, we can derive 92% of the maturation efficiency of the mEGFP fluorophore.

Second, we examined the measurement noise of FCS and intracellular variability of concentration. To calculate measurement noise, single-point FCS time series data were divided into five segments, and each segment was separately analyzed to obtain autocorrelation functions and protein concentrations (Fig. 5D). For intracellular variability, we randomly set five points for FCS measurement in a single cell and determined protein concentrations (Fig. 5E). The results showed that the measurement noise was much smaller than the intracellular variability (Fig. 5F).

Third, to evaluate the difference between the intracellular and intercellular variability, we calculated the absolute difference of the number of molecules between two points (Fig. 5G). As a positive control, we transiently overexpressed mEGFP-HaloTag proteins in HeLa cells, measured the number of mEGFP molecules by FCS, and calculated the differences of intra- and intercellular variability. As expected, the histogram of the intercellular variability showed a broader distribution of the absolute difference of the number of molecules than the intracellular variability (Fig. 5H), indicating that this method is able to discriminate between intra- and intercellular variability. Similarly, we examined the intra- and intercellular variability in ERK2-mEGFP KI cells. Unexpectedly, we did not observe a significant difference between the intra- and intercellular variability of ERK2-mEGFP (Fig. 5I). This observation suggested that the intercellular variability of protein concentration is similar to or smaller than the intracellular variability. This result was also supported by comparing the fluorescence intensity of a pair of pixels with a conventional confocal microscope; the intracellular variability was larger than the intercellular variability in both ERK2-mEGFP and RSK2-HaloTag-TMR (Fig. S4). These results indicate that the intra- and intercellular variabilities of protein concentration measured by FCS are at a similar level when the endogenous promoter of MAPK1 is used.
Live-cell measurement of concentration and in vivo $K_d$ value by FCS and FCCS

After these characterizations of the cell lines and methods, we determined the protein concentrations and the in vivo $K_d$ values of ERK2-mEGFP and RSK2-HaloTag in HeLa/ERK2-mEGFP/RSK2-HaloTag cells by FCS and FCCS, respectively. The autocorrelation functions of ERK2-mEGFP and RSK2-HaloTag-TMR and cross-correlation functions between ERK2-mEGFP and RSK2-HaloTag-TMR were calculated in a single HeLa cell (Fig. 6A). We repeated this analysis with 198 cells and obtained the distribution of the endogenous protein concentration of ERK2-mEGFP and RSK2-HaloTag-TMR (Fig. 6B). The average concentrations of ERK2-mEGFP and RSK2-HaloTag-TMR were determined to be 0.078 $\mu$M and 0.097 $\mu$M, respectively (Fig. 6B). Based on these values, Western blot data (see Fig. 2D for ERK2 and Fig. 3D for RSK2) and the fluorophore maturation efficiency of mEGFP (Fig. 5B; see “Experimental procedures”), the total ERK2 and RSK2 concentrations in a HeLa cell were estimated to be 0.68 $\mu$M and 0.50 $\mu$M, respectively. As mentioned above, the heterogeneity of ERK2 and RSK2 concentration was mainly attributed to the intracellular variability (Fig. 5 and Fig. S4). Nevertheless, a positive correlation between ERK2-mEGFP and RSK2-HaloTag-TMR was found, with a correlation coefficient of 0.64 (Fig. 6B).

The average in vivo $K_d$ value between ERK2-mEGFP and RSK2-HaloTag-TMR was $\sim$0.59 $\mu$M under the basal condition (Fig. 6C). The in vivo $K_y$ values also showed little correlation with the protein concentrations of ERK2-mEGFP and RSK2-HaloTag-TMR (Fig. 6D). We confirmed that ERK2-mEGFP and RSK2-HaloTag-TMR did not bind nonspecifically to HaloTag-TMR and mEGFP, respectively (Fig. S5).

Finally, to take full advantage of this strategy, we quantified the temporal changes in in vivo $K_d$ values at the subcellular level. HeLa/ERK2-mEGFP/RSK2-HaloTag cells were stained with the TMR-HaloTag ligand and treated with EGF, followed by measurement of the protein concentrations and in vivo $K_d$ values.
values in the cytoplasm and nucleus. As demonstrated in Fig. 4, ERK2-mEGFP, but not RSK2-HaloTag-TMR, was translocated from the cytoplasm to the nucleus upon EGF stimulation (Fig. 6, E and F). The nuclear concentration of ERK2-mEGFP after EGF stimulation surpassed the cytoplasmic concentration (Fig. 6E). The in vivo $K_d$ values in the cytosol were higher than those in the nucleus over the course of EGF stimulation (Fig. 6G). Importantly, the $K_d$ value in the cytosol, but not the nucleus, transiently increased 10 min after EGF stimulation (Fig. 6G). These results suggested that the ERK2–RSK2 complex was more stable in the nucleus than in the cytosol under both the unstimulated and stimulated conditions and that EGF stimulation transiently reduced the apparent affinity of the ERK2–RSK2 complex only in the cytoplasm. Another possibility is that the in vivo $K_d$ values changed as a result of changes in the competitive reactions.

The values of the ratio of complex to ERK2-mEGFP did not show any significant change because ERK2-mEGFP per se shuttled between the cytoplasm and nucleus (Fig. 6H). In contrast, the ratio of complex to RSK2-HaloTag-TMR gradually increased in the nucleus (Fig. 6I). This is because the RSK2 concentration did not change, and ERK2 entered the nucleus.

**Discussion**

Here we established a method for quantifying the endogenous protein concentration and dissociation constant in living cells with the CRISPR/Cas9 genome editing technique and FCS/FCCS. In principle, these techniques allow us to quantify these kinetic parameters without any antibodies. Although several research groups have reported methods for endogenous tagging in cultured mammalian cells (21, 33, 34), gene KI is still
Live-cell measurement of concentration and in vivo $K_d$

challenging, even with the CRISPR/Cas9 technique. The KI efficiency is much lower than the random integration efficiency in HeLa cells. Therefore, it is important not only to increase the KI efficiency (35–37) but also to eliminate the negative clones. Further, we found that the selection cassette had a considerable effect on the expression of the tagged protein; the remaining selection cassette reduced the expression level of the protein fused with fluorescent protein (Figs. 2 and 3). Moreover, recently we have seen a reduction in the protein expression level after Cre-mediated selection cassette removal (38). This could be due to the effect of the 3' UTR of the gene that was tagged by KI. For these reasons, removal of the selection cassette by Cre recombination is essential for the quantitative assessment of protein concentration.

We quantified the endogenous protein concentrations of ERK2-mEGFP and RSK2-HaloTag-TMR by FCS (Fig. 6). The average concentration of ERK2-mEGFP measured in this study, 0.078 $\mu M$, was considerably lower than the ERK2 concentration reported previously, 0.68 $\mu M$ for ERK2 in HeLa cells (14). This may have been due to the multiplication of MAPK1 genes because it is evident that the reduction of unlabeled ERK2 by KI of the mEGFP gene was quite limited (Fig. 2). The estimated total ERK2 concentration based on the mEGFP maturation efficiency and Western blotting was 0.68 $\mu M$, which is consistent with the previous data (39). Therefore, gene amplification, which is frequently observed in cancer cells, is a limitation of this technique. On the other hand, the RSK2-HaloTag-TMR concentration, 0.097 $\mu M$, was similar to the previously reported RSK2 concentration of 0.15 $\mu M$ (14) (Fig. 3). The estimated total RSK2 concentration from Western blotting, 0.50 $\mu M$, was comparable with our data. In contrast to MAPK1, KI to the RSK2 gene reduced the endogenous RSK2 protein to half or zero (Fig. 3, C and D), implying monoallelic and biallelic KI, respectively.

Intracellular variability is a technical challenge in the quantification of protein concentration by FCS. The heterogeneity of the subcellular environment, including heterogeneity among the cellular organelles, causes the intracellular variability because a positive correlation was observed between the protein abundance of ERK2-mEGFR and RSK2-HaloTag-TMR (Fig. 6B). FCS measures fluorescence fluctuation in a tiny confocal volume (~1 fl), and therefore the intra-cellular variability is unavoidable. For accurate determination of the protein concentrations in cells, the intracellular variability can be canceled out by averaging repeated measurements at different points in each cell. However, by this approach, photobleaching will be another obstacle for obtaining datasets.

One advantage of this approach is to determine in vivo $K_d$ values for binding of ERK2-mEGFP and RSK2-HaloTag-TMR by FCCS (Fig. 6). The average in vivo $K_d$ value measured in endogenously tagged ERK2 and RSK2 in this study was 0.59 $\mu M$ (Fig. 6C), but this value was comparable with or slightly higher than the in vitro $K_d$ value of ERK2 and RSK1 binding, 0.15 $\mu M$ (39). This difference could be due to competitive binding from endogenous proteins (14). Interestingly, the in vivo $K_d$ value measured in this study was lower than the in vivo $K_d$ value of 1.3 $\mu M$ in overexpressed ERK2-mEGFP and RSK2-HaloTag (14). PEA-15 is known to function as a scaffold to enhance RSK2 activation by ERK (40). Thus, in the previous study, the overexpressed amounts of ERK2-mEGFP and RSK2-HaloTag might have exceeded the amount of endogenous PEA-15, resulting in overestimation of the in vivo $K_d$ values. Further, the in vivo $K_d$ values of ERK2-mEGFP and RSK2-HaloTag-TMR binding were hardly correlated with the expression levels of those proteins (Fig. 6D), suggesting that the binding between ERK2 and RSK2 fluctuates in a manner dependent on unknown mechanisms, such as competitive binding, molecular crowding, post-translational modification, and/or scaffold proteins. In the future, quantitative analysis of the heterotrimer complex could be a fascinating way to verify the scaffolding function of PEA-15 for ERK2-RSK2 binding (41).

After EGF stimulation, the $K_d$ value between ERK2-mEGFP and RSK2-HaloTag-TMR was transiently increased in the cytosol (Fig. 6G). It has been reported that ERK binding is differently regulated among RSK isoforms; RSK1 dissociated completely from ERK1/2 upon EGF stimulation, RSK2 only partially dissociated from ERK1/2, and RSK3 remained bound to ERK1/2 (27). These previous findings are consistent with our finding that a part of the complex of ERK2-mEGFP and RSK2-HaloTag-TMR was dissociated 10 min after EGF stimulation, whereas the other part remained (Fig. 6). The finding that the activation of RSK2 upon EGF stimulation is comparable with that of RSK1 suggested that the phosphorylation of RSK2 by ERK1/2, PDK1, and/or autophosphorylation (27, 42) reduced the affinity to ERK2 to some extent. We should also be careful not to exclude the possibility of changes in competitive binding upon EGF stimulation.

In summary, we established a method for quantifying the endogenous protein concentration and $K_d$ value in a living cell. We believe that these parameters are useful for building more quantitative kinetic simulation models. The kinetic model can be simulated accurately when all of the interacting molecules and their true $K_d$ values are available and explicitly incorporated into the model. However, we suggest that this is not a realistic approach at present because no quantitative map of the whole intermolecular interaction is available. Although in vivo $K_d$ values vary depending on the intracellular environment, the in vivo $K_d$ value will provide better clues for quantitative simulations of intracellular signal transduction (14). Further studies will be needed to improve KI efficiency, the brightness of fluorescent proteins, and quantification of proteins localized at specific cellular compartments, such as the plasma membrane.

Experimental procedures

Design of sgRNAs of CRISPR/Cas9 for gene KI

pX459 (pSpCas9(BB)-2A-Puro) was a gift from Feng Zhang (Addgene plasmid 62988). All guides are designed to overlap the stop codon of MAPK1 or RSK2 so that these guide RNAs do not recognize the target sequences after KI. Guide sequences (without the protospacer adjacent motif (PAM) sequence) were as follows; for MAPK1, TCTTAAATTGTTCAGGTACC; for RSK2, CTCACTGTAGGTCAGGTACC.

Construction of KI donor vectors

Truncated thymidine kinase (dTK) originating from pPB-CAG.OSKM-puDrk (a gift from Kosuke Yusa and Allan Bradley (43)) and the neomycin resistance gene (neo) were conjugated
Live-cell measurement of concentration and in vivo $K_d$

by PCR (Fig. S1). KOD FX neo (TOYOBO, Osaka, Japan) was used for all PCR reactions. The homology arms of the HDR vector were amplified from genomic DNA of HeLa cells extracted with QuickExtract DNA extraction solution (Epicenter, Madison, WI) and attached to an insertion cassette by PCR. 40-bp homology arms for the MMEJ donor were attached by PCR.

Cell culture

HeLa cells were purchased from the Human Science Research Resources Bank (Sennanshi, Japan). HEK293T cells were obtained from Invitrogen as Lenti-X 293 cells (Invitrogen). HeLa cells and HEK293T cells were maintained in DMEM (Wako, Osaka, Japan) supplemented with 10% FBS. For imaging, HeLa cells were plated on 35-mm glass-base dishes (Asahi Techno Glass, Tokyo, Japan). At least 3 h before observation, HeLa cells were maintained with FluoroBrite DMEM (Life Technologies) supplemented with 1% GlutaMAX (Life Technologies) and 2% FBS.

For HaloTag imaging, the cells were incubated with 100 nM HaloTag-TMR for at least 18 h. After that, the cells were washed twice with PBS and then incubated with FluoroBrite DMEM supplemented with 1% GlutaMAX for 30 min. The medium was again replaced with the same medium.

Establishment of the KI cell line

HeLa cells were plated on a 24-well plate and transfected with 1 μg of pX459 vectors and 50 ng of KI donor DNA by using polyethyleneimine “Max” MW 40,000 (Polyscience Inc., Warrington, PA), followed by puromycin selection. Three days after transfection, transfected cells were seeded onto 35-mm dishes and treated with 0.5 mg/ml G-418 (Invivogen, San Diego, CA). For HaloTag imaging, the cells were incubated with 100 nM HaloTag-TMR for at least 18 h. After that, the cells were washed twice with PBS and then incubated with FluoroBrite DMEM supplemented with 1% GlutaMAX for 30 min. The medium was again replaced with the same medium.

Adeno-associated viruses expressing Cre and removal of the selection marker

The complementary DNA of Cre was inserted into pAAV-MCS (Cell Biolabs Inc., San Diego, CA), generating pAAV-Cre. HEK293T cells were co-transfected with pAAV-Cre, pAAV-DJ, and pH Helper to produce a recombinant AAV expressing Cre recombinase. Three days after transfection, HEK293T cells were collected and resuspended in 1 ml of DMEM, followed by four freeze–thaw cycles. After the final thaw and centrifugation, 10 μl of supernatant was added to each well of a 24-well plate to remove the selection marker. At least 6 days after AAV-Cre infection, the cells were selected with 50 μM ganciclovir (Wako).

Immunoblotting

Cells were lysed in 2× SDS sample buffer. After sonication, the samples were separated by 5–20% or 7.5% SDS-PAGE (Nacalai Tesque Inc., Kyoto, Japan) and transferred to polyvinylidene difluoride membranes (Millipore, Billerica, MA). After blocking with Odyssey blocking buffer (Tris-buffered saline, LI-COR Biosciences Inc., Lincoln, NE) for 1 h, the membranes were incubated with primary antibodies diluted in blocking buffer, followed by secondary antibodies diluted in blocking buffer. Fluorescence levels were detected by an Odyssey IR scanner (LI-COR Biosciences Inc.). The following antibodies were used in this study: anti-GFP antibody (632375, Clontech, Palo Alto, CA), anti-p44/p42 MAPK (ERK1/2) antibody (4695, Cell Signaling Technology, Danvers, MA), anti-RSK2 (sc-9986, Santa Cruz Biotechnology, Santa Cruz, CA), anti-HaloTag antibody (G9281, Promega, Madison, WI), anti-α-tubulin (PM054, MBL), IRDye680LT goat anti-rabbit IgG antibody (925-68021, LI-COR Biosciences, Inc.), and IRDye800CW donkey anti-mouse IgG antibody (925-32212, LI-COR Biosciences, Inc.).

Imaging by confocal microscope

Cells were imaged with a laser-scanning confocal microscope (FV1200, Olympus) equipped with gallium arsenide phosphide detectors. The excitation lines were set at 473 nm and 559 nm. The excitation beam was reflected by a DM 405/488/559 dichroic mirror and focused by an oil immersion objective lens (Uplanapo ×60, numerical aperture 1.35, Olympus). The emitted light was detected through a band pass filter with wavelengths of 495 to 540 nm for mEGFP and 575 to 630 nm for HaloTag-TMR.

FCS and FCCS data were obtained as described previously (14). Briefly, cells were imaged with a laser-scanning confocal microscope (FV1000 or FV1200, Olympus) equipped with gallium arsenide phosphide detectors. The excitation lines and power were set at 473 nm (0.1%) and 559 nm (0.1%). The excitation beam was reflected by a DM 405/488/559 dichroic mirror and focused by an oil immersion objective lens (Uplanapo ×60, numerical aperture 1.35, Olympus). Time series data of fluorescence fluctuation were obtained by point scan (2.0 μsec/1 sampling) in photon count mode, and total time series data were typically 8,000,000 points. Of note, the first 1,600,000 data points were used for the FCS and FCCS analysis to exclude the effect of photobleaching.

FCS and FCCS data were analyzed as follows. First, to correct intensity decay because of photobleaching, we scaled the intensity traces so that the corrected traces had constant means and fluctuations in time (See Equations 21 and 22 in Ref. 44). Then we used the multiple tau Python package (version 0.1.9) developed by Paul Müller (https://pypi.org/project/multipletau/, accessed January 9, 2018) to compute auto- and cross-correlations. The numerical correlation functions were fitted by the theoretical model

Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party–hosted site.
Live-cell measurement of concentration and in vivo $K_d$

5.1. ERK1 and tubulin were used as loading controls for ERK2 and RSK2, respectively. Based on these values, the total ERK2 concentration was estimated as follows: 0.078 / 0.92 × 8.0 = 0.68 μM. The total RSK2 concentration was estimated as follows: 0.097 × 5.14 = 0.50 μM.


Acknowledgments—We thank the members of the Matsuda laboratory and Aoki laboratory for helpful input. We also thank A. Kawagishi, K. Hirano, N. Nishimoto, E. Ebine, and K. Onoda for technical assistance.

References
Live-cell measurement of concentration and in vivo $K_d$


