Plasmid-DNAgram: 세포 내 다양한 형광단백질쌍의 분자간 FRET을 이용한 언어 퍼즐 문제 해결

김수동[3] 이은석[3] 박형주[1] 이승희[1] 김민정[1] 강봉균[1, 3] 장병탁[2, 3]

[1]서울대학교 생명과학부, [2]서울대학교 컴퓨터공학부, [3]서울대학교 대학원 인지과학 협동과정

[1]{phj2, sabet1, minjkim1, kaang}@snu.ac.kr; [3]{sdkim, eslee, btzhang}@bi.snu.ac.kr

Plasmid-DNAgram: Anagram Solving by Molecular Interactions in Living Cells by FRET between Variants of the Green Fluorescent Protein

Su Dong Kim[3], Eun-Seok Lee[3], Hyungju Park[1], Seung-Hee Lee[1], Min-Jeong Kim[1], Bong-Kiun Kaang[1, 3] and Byoung-Tak Zhang[2, 3]

[1]School of Biological Sciences, Seoul National University

[2]School of Computer Science & Engineering, Seoul National University

[3]Interdisciplinary Program in Cognitive Science, Seoul National University

요 약

Plasmid-DNAgram, a variant of DNAgram, adopted the concept of DNAgram that linguistic information could be encoded on bio-molecular sequences. For example, a DNA restriction map of "*Bsi*WI - *Xho*I - *Eco*RV - *Dra*I" could be regarded as a word, "word." Plasmid-DNAgram aims to do linguistic information processing *in vivo*. To demonstrate this purpose, we used the fused catalytic and regulatory subunits of PKA to GFP mutants with excitation and emission spectra suitable for FRET. The indication that FRET between the tagged PKA subunits occurs was obtained by co-transfecting COS-7 cells or HEK293T cells with RⅡ-CFP and C-YFP, exciting at 458 nm, the optimum for CFP, and collecting emissions at 480 and 550 nm for CFP and YFP, respectively. CFP could be regarded as a sign of adjective, and YFP as a noun-tag. Normal expression-level of fluorescent protein means that it is carrying correct semantic information. The physiological modulation of FRET would be able to correspond with the pragmatics. In conclusion, our results indicate that imaging of FRET between variants of GFP in living cells could be considered as a simple linguistic information processing *in vivo*.

1 INTRODUCTION

Linguistic metaphors have been woven into the fabric of molecular biology since its inception. In 2002, David Searls reviewed the notion of DNA as language and the genome as the 'book of life' [1]. That is, primary DNA sequence corresponds to the lexical 'string' level of linguistics, while syntax can be seen as dealing with structures not only in an informational but also a conformational sense, when one views parse trees as capturing distant dependencies induced by interactions within 'folded structures'. Arguably, the function of a protein can be seen as its semantics, albeit determined not by logic and experience but by thermodynamics and biochemistry [2].

Lewis Carroll (1832-1898), the author of the *Alice* books, had created a vast range of games and puzzles that depend upon wordplay of various kinds, especially the manipulation of alphabetic symbols in diverse contexts. David Searls rediscovered that meaning of Carroll's wordplays toward modern computational biology [3].

Inspired by Carroll's most famous poem, "Jabberwocky [4]," Eun-Seok Lee suggested DNAgram, a molecular simulation of anagram problem solving [5]. In this study, we describe Plasmid-DNAgram, a variant of DNAgram.

1.1 Plasmid-DNAgram

DNAgram, a molecular simulation, models how people solve anagram puzzle [5]. DNAgram demonstrates the procedure of extracting plausible English word answer by molecular computing, which exposes new information and provides new perspective on existing anagram creation models and its methods.

Plasmid-DNAgram, a variant of DNAgram, adopted the concept of DNAgram that linguistic information could be encoded on bio-molecular sequences. Plasmid-DNAgram aims to do linguistic information processing *in*

vivo. To perform Plasmid-DNAgram, we assigned 26 characters of alphabet to 26 kinds of DNA restriction enzyme site. For example, a DNA restriction map of "*Bsi*WI - *Xho*I - *Eco*RV - *Dra*I" could be regarded as a word, "word." Similarly, a DNA restriction map of "*Xho*I - *Kpn*I - *Eco*RI" could be regarded as a word, "one." If we shuffled the DNA restriction sites, "*Xho*I - *Kpn*I - *Eco*RI," assigned for a word "one," we could easily get its anagrams, "*Kpn*I - *Eco*RI - *Xho*I " for "neo", and "*Eco*RI - *Xho*I - *Kpn*I" for "eon." Each DNA sequence could be inserted into the multicloning sites (MCS) of various fluorescent protein-expression vectors. And then each subcloned plasmid DNA could be transfected into the living cells. We can observe the resulting proteins fused with the variants of fluorescent protein expressed from these vectors within the transfected cells. We tagged a target protein to a fluorescent protein as a part of speech. For example, cyan color of cyan fluorescent protein (CFP) could be regarded as adjective, and yellow color of yellow fluorescent protein (YFP) could be regarded as noun. If we could detect fluorescent resonance energy transfer (FRET) between these fluorescent protein pairs in living cells, we conclude that we are able to perform a simple syntactic information processing *in vivo*.

1.2 FRET between variants of the GFPs

Fluorescent resonance energy transfer (FRET), which relies on the distance-dependent transfer of energy from a donor fluorophore to an acceptor fluorophore, is one of the few tools available for measuring nanometer scale distances and changes in distances, both *in vitro* and *in vivo* [6]. Genetically encoded dyes, such as green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* and its varients (blue, cyan, and yellow), have revolutionized the ability to perform FRET *in vitro* and especially in living cells [7]. These proteins form FRET pairs with each other or with conventional organic dyes, and can be attached to many proteins of interest, usually at the N- or C-terminus. Tsien *et al.* [8] and other workers have developed several GFP FRET constructs that are used to monitor the biochemical environment inside living cells.

Two pairs of GFP mutants have excitation and emission properties favorable for FRET: blue fluorescent protein (BFP) with GFP and cyan fluorescent protein (CFP) with yellow fluorescent protein (YFP) [9]. More recently, novel YFPs and red fluorescent proteins (RFPs) have been cloned from corals of the *Discosoma* genus [10], producing yet another FRET pair, GFP-RFP [11].

Protein kinase A (PKA), a holotetramer is composed of two regulatory (R) and two catalytic (C) subunits [12]. Each subunit of cAMP-dependent PKA could be labelled with a different fluorescent protein such as CFP or YFP capable of fluorescence resonance energy transfer (FRET) in the holoenzyme complex R_2C_2 [13].

2 MATERIALS AND METHODS

2.1 Construction of GFP-tagged Proteins

Both constructs, pcDNA3-RⅡ-CFP and pcDNA3-C-YFP, were kindly provided by M. Zaccolo (University of Padua, Padua, Italy) [14]. We regarded that RⅡ-CFP and C-YFP were encoded simple wordplays described above.

2.2 HEK293T cell transfection

The methods for the culture of COS-7 cells or human embryonic kidney (HEK293T) cells and their transient transfection with pcDNA3-RⅡ-CFP and pcDNA3-C-YFP were similar to those described by Nashmi *et al* [15]. Briefly, cells were grown in Dulbecco's modified Eagle's medium (DMEM; Sigma) supplemented with 10% fetal bovine serum (FBS). Cells were split twice per week. COS cells or HEK293T cells were plated onto 35mm Petri dishes with glass coverslip bottoms that were coated with poly-$_D$-lysine (MatTek), incubated at 37℃ in an incubator until they were 60-80% confluent, and then transiently were transfected with pcDNA3-RII-CFP and pcDNA3-C-YFP using Lipofectamine 2000 (Invitrogen). For each dish, 1 μg of plasmid DNA was mixed with transfection reagent and medium according to the manufacturer's protocol.

2.3 Imaging

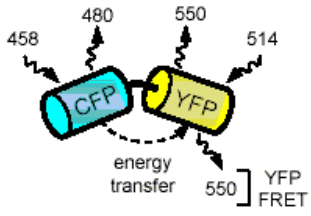

Fig. 1. Schematic illustration of FRET between CFP and YFP. This schematic shows fluorescence resonance energy transfer from the donor CFP fluorophore (blue) to the acceptor YFP fluorophore (yellow). Figure adapted from ref. [17].

Images were obtained with a confocal laser-scanning microscope (LSM510, Carl Zeiss) using laser excitation at 458 nm (CFP) and 514 nm (YFP) (Fig. 1). Fluorescence were measured as previously described [16]. All fluorescence data were collected with a high-resolution (1.4 NA), 60x oil-immersion DIC lens. The software controlled 510 LSM confocal microscope enabled timed regional photobleaching performed

selectively at 514 nm with a mean reduction of 525-575 nm (YFP) emissions to less than 15 % in the photobleached region [17]. Images were acquired before and after photobleaching.

3 RESULTS

3.1 Construction of GFP-tagged Proteins

Both constructs, pcDNA3-RⅡ-CFP and pcDNA3-C-YFP, were kindly provided by M. Zaccolo (University of Padua, Padua, Italy) [14]. Zaccolo and Pozzan cloned CFP and YFP in the frame at the 3'-end of the rat PKA regulatory (R) β-subunit in the constructs RⅡ-EBFP [13] or the mouse PKA catalytic (C) subunit in the constructs C-S56T [13] in place of EBFP and S65T, respectively [14]. Then they cloned both chimeras into pcDNA3 (Invitrogen) as *Hin*dⅢ / *Xba* Ⅰ fragments.

3.2 Expression of GFP-tagged PKA

We transfected pcDNA3-RⅡ-CFP pcDNA3-C-YFP into stable cell lines, such as HEK293T cells and COS-7 cells. When RⅡ-CFP and C-YFP were transiently coexpressed, the fusion proteins localized in the cytoplasm and were clearly and consistently excluded from the nucleus (Fig. 2). This result is consistent with the report that the cAMP target enzyme, PKA, is largely compartmentalized within the three-dimensional matrix of the cell via anchoring to A-kinase anchoring proteins (AKAPs) [18].

3.3 GFP-tagged PKA generate FRET

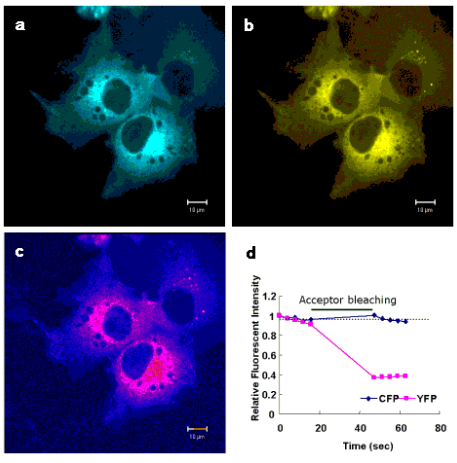

Fig. 2. COS-7 cells transfected with RⅡ-CFP (A) and C-YFP (B). The colocalization of the two fluorescent PKA subunits is shown in (C). Plots of YFP and CFP intensities during acceptor photobleaching (D). Each scale bar represents 10 μm.

The indication that FRET between the tagged PKA subunits occurs was obtained by co-transfecting COS-7 cells or HEK293T cells with RⅡ-CFP and C-YFP, exciting at 458 nm, the optimum for CFP, and collecting emissions at 480 and 550 nm for CFP and YFP, respectively (Fig. 1, Fig. 2). Furthermore, we can monitor PKA regulatory (R) and catalytic (C) subunit dissociation due to a rise in the intracellular cAMP concentration, $[cAMP]_i$, by measuring changes in FRET between the CFP and YFP [14].

4 DISCUSSION

Our aim was to demonstrate simple linguistic information processing *in vivo*. For this purpose, we used the fused catalytic and regulatory subunits of PKA to GFP mutants with excitation and emission spectra suitable for FRET developed by Zaccolo and Pozzan [14].

In this study, called Plasmid-DNAgram, lexical information is encoded on DNA restriction map in MCS of plasmid DNA. Practically, we choose variants of fluorescent protein-expression vector for operating linguistic information. Each subcloned plasmid contains its specific lexicon within MCS, and its products expressed in living cells are tagged with the fluorescent protein representing its own part of speech. For example, "*Kpn*I - *Eco*RI - *Bsi*WI" for an adjective, "new," and "*Dra*I - *Xho*I - *Xho*I - *Eco*RV" for a noun, "door." If we shuffled these DNA restriction sites, we could easily get their anagrams, "*Xho*I - *Kpn*I - *Eco*RI," for an adjective, "one," and "*Bsi*WI - *Xho*I - *Eco*RV - *Dra*I" for a noun, "word." CFP could be regarded as a sign of adjective, and YFP as a noun-tag. Normal expression-level of fluorescent protein means that it is carrying correct semantic information. When the two binding partners of protein tagged with the fluorescent protein FRET pairs form their appropriate conformational complex, we can measure FRET between these molecules *in vivo* as the signal that a proper syntactic phrase was constructed. FRET efficiency is altered in response to signaling molecules such as cAMP and Ca^{2+} [19]. Its physiological modulation would be able to correspond with the pragmatics.

5 CONCLUDING REMARKS AND FUTURE WORKS

In conclusion, our results indicate that imaging of FRET between variants of GFP in living cells could be considered as a simple linguistic information processing *in vivo*.

Unfortunately, the FRET pairs of GFP are very limited now. However, efforts to engineer new fluorophores and reporter classes must continue [20]. If we have much more FRET pairs of GFP, we could manage to process more complex syntactic structures *in vivo* via Plasmid-DNAgram.

ACKNOWLEDGMENTS

This research was supported in part by the Ministry of Education & Human Resources Development under the BK21 Program, the Ministry of Commerce, Industry and Energy through the MEC project, and the NRL Program from Korean Ministry of Science & Technology. The IMBG and the ICT at Seoul National University provided research facilities for this study.

REFERENCES

1. Searls, D. B. The language of genes, *Nature*, 420:211-217, 2002.
2. Searls, D. B. Reading the book of life, *Bioinformatics*, 17(7):579-580, 2001.
3. Searls, D. B. From Jabberwocky to Genome: Lewis Carroll and Computational Biology, *J. Comput. Biol*, 8(3):339-348, 2001.
4. Carroll, L. *The Complete Works of Lewis Carroll*, Modern Library, New York, 1936. Reprinted in paperback by Penguin, London, 1999.
5. Lee, E. -S. DNAgram: Molecular Simulation of Anagram Problem Solving, *Master's Thesis*, *Seoul National University*, 2003. (in Korean)
6. Selvin, P. The renaissance of fluorescence resonance energy transfer, *Nat. Struct. Biol.*, 7(9): 730-734, 2000.
7. Tsien, R. Y.The green fluorescent protein, *Annu. Rev. Biochem.*, 67:509-544, 1998.
8. Tsien, R. Y., Miyawaki, A. Seeing the machinery of the living cells, *Science*, 280:1954-1955, 1998.
9. Siegel, R. M. *et al.* Measurement of molecular interactions in living cells by fluorescence resonance energy transfer between variants of the green fluorescent protein, *Sci. STKE*, 2000:38, pl1, 2000.
10. Matz, M. V. *et al.* Fluorescent proteins from non-bioluminescent Anthozoa species, *Nat. Biotechnol.*, 17:969-973, 1999.
11. Mizuno, H. *et al.* Red fluorescent protein from *Discosoma* as a fusion tag and a partner for fluorescence resonance energy transfer, *Biochemistry*, 40:2502-2510, 2001.
12. Meinkoth, J. L. *et al.* Signal transduction through the cAMP-dependent protein kinase. *Mol. Cell.Biochem.* 127:179–187, 1993.
13. Zaccolo, M. *et al.* A genetically encoded, fluorescent indicator for cAMP in living cells, *Nat. Cell Biol.*, 2:25-29, 2000.
14. Zaccolo, M., Pozzan, T. Discrete microdomains with high concentration of cAMP in simulated rat neonatal cardiac myocytes, *Science*, 295:1711-1715, 2002.
15. Nashmi, R. *et al.* Assembly of $\alpha 4\beta 2$ Nicotinic acetylcholine receptors assembled with functional fluorescently labeled subunits: Effects of localization, trafficking, and nicotine-induced upregulatoin in clonal mammalian cells and in cultured midbrain neurons, *J. Neurosci.*, 23(37):11554-11567, 2003.
16. Centonze, V. E. *et al.* Fluorescence resonance energy transfer imaging microscopy, *Methods Enzymol.*, 360:542-560, 2003.
17. Staruschenko, A. *et al.* Fluorescence resonance energy transfer analysis of subunit stoichiometry of the epithelial Na^+ channel, *J. Biol. Chem.*, April 19, 2004 [Epub ahead of print].
18. Shillace, R. V. Scott, J. D. Organization of kinases, phosphatases, and receptor signalling complexes, *J.Clin. Invest.*, 103(6):761-765, 1999.
19. Sekar, R. B., Periasamy, A. Fluorescence resonance energy transfer (FRET) microscopy imaging of live cell protein localizations, *J. Cell Biol.* 160(5):629-633, 2003.
20. Zhang, J. *et al.* Creating new fluorescent probes for cell biology, *Nat. Rev. Mol. Cell Biol.* 3(12):906-918, 2002.