Thesis abstract

Arranging multiple types of enzymes in defined space by modular adaptors

(モジュール型アダプターを利用した複数種類の酵素の特異的な配置)

NGUYEN MINH THANG

Research background

Bioenergy is one of the ideal sources of renewable energy because living organisms are optimised for usage of energy and materials during the long time evolution. Organisms rely on the well-ordered system of proteins and enzymes that drives the complex metabolic pathways. The well-ordered natural enzyme system is usually compartmentalized in many levels such as clustered enzymes, enzymes cascades, multi-enzyme assemblies and organelles. This ubiquitous phenomenon allows an organism to utilize energy and materials in an optimized condition. Therefore, studying the enzyme compartmentalization will open a new dimension for enzyme technology. The DNA origami technique provides an ideal scaffold to assemble enzymes or proteins of interests (POIs) in a spatially well-defined manner (P. W. K. Rothemund, *Nature*, **2006**, *440*, 297). Cascade reactions by two or three enzymes on various DNA scaffolds have been studied to demonstrate enhancements of the cascade reaction depending on the interenzyme distance (A. Rajendran, *et al*, *ChemBioChem* **2017**, *18*, 696). However, assembling proteins on the DNA scaffold remains challenging because proteins are fragile and require mild treatments in short time. Although a number of strategies have been developed to assemble POIs on a DNA scaffold, these methods often accompany with disadvantages, such as low assembly yields, reduction of the enzyme activity, and requirements of long preparation time, indicating them to be not suitable for assembly multiple enzymes.

Problems of the previous methods

The previous study in the Morii laboratory has applied the sequence-specific DNA binding proteins as adaptors for locating POIs at the target positions on the DNA in short time (E. Nakata, *et al*, *Angew. Chem., Int. Ed.* **2012***, 52,* 2421). A drawback of this adaptor system is the reversible nature of DNA-adaptor complex, which causes incomplete loading of POI to the

desired position on the DNA scaffold. To overcome this issue, a modular adaptor (ZF-SNAP) consisting of a sequence-specific DNA binding zinc finger adaptor zif268 (ZF) and a self-ligating protein-tag (SNAP-tag) has been developed to expedite the efficient formation of a covalent linkage between the individual modular adaptor molecule and the programmed position modified with a substrate of protein-tag on the DNA scaffold (E. Nakata, *et al*, *Chem. Commun.,* **2015,** *51*, 1016). To assemble several POIs on the DNA scaffold at their specific positions in functional forms, modular adaptors with orthogonality and fast kinetics for the covalent bond formation are required.

Purpose of the research

This research aims to develop a simple and efficient method for locating multiple enzymes on a DNA scaffold by using the modular adaptor strategy. The orthogonal modular adaptors were developed by the combination of DNA binding proteins and protein-tags. A set of orthogonal modular adaptors was used to co-assemble three enzymes derived from the natural xylose metabolic pathway on a DNA scaffold to study the efficiency of enzyme cascade rations by the enzymes in close proximity.

Contents of the research

In Chapter 2, I first expanded the combinations of DNA binding protein (zif268, AZP4) and the protein-tag (SNAP-tag, CLIP-tag, and Halo-tag) to select an orthogonal set of modular adaptors to locate more than two types of enzymes on a DNA scaffold. The possible six types of modular adaptors were classified into two categories: those sharing the same DNA binding protein with different protein-tags and those sharing the same protein-tag with different DNA binding proteins. Among them, I constructed four modular adaptors: ZF-SNAP, AZ-SNAP, AZ-CLIP, AZ-Halo, which were divided to two groups: one with the same protein-tag (ZF-SNAP and AZ-SNAP) and one with the same DNA binding protein (AZ-SNAP, AZ-CLIP, and AZ-Halo). With these two groups, I asked a question whether the orthogonality could be dictated by the specificity of DNA binding domain or by the substrate selectivity of protein-tag.

I first checked the affinity of DNA binding domain from AZP4 (AZ) to oligodeoxynucleotide (ODN) containing the specific sequence for AZP4 (ODN-AZ). The

equilibrium dissociation constants (K_D) for AZ-SNAP, AZ-CLIP, and AZ-Halo with the matched ODN were determined. The K_D values for AZ-SNAP, AZ-CLIP, and AZ-HALO were 61 nM, 65 nM, and 118 nM, respectively. For the unmatched complexes, AZ-SNAP, AZ-CLIP, and AZ-Halo did not afford measurable dissociation constants for ODN-ZF ($K_D > 1000$ nM). These results indicated that the affinity and selectivity of DNA binding protein were not influenced by conjugation of the protein-tag domain, which was consistent with our previous reports on the affinity of DNA binding proteins with the introduction of various POIs. Kinetics of the covalent bond formation (k_{app}, $M^{-1}s^{-1}$) by modular adaptors with the matched and the unmatched pairs were also investigated. The second-order rate constants of the matched reactions were determined to be 10^5 - 10^6 $M^{-1}s^{-1}$, respectively, almost in the same order of magnitude. Such similar rates of covalent bond formation by modular adaptors are convenient for the co-assembly of POIs at the same time. The k_{app} for the unmatched pairs were significantly reduced for ZF-SNAP, AZ-CLIP, and AZ-HALO, about two to three orders of magnitudes to that of the matched pair.

The denaturing PAGE and AFM imaging analyses were used to study the orthogonal reaction by four types of modular adaptors (ZF-SNAP, AZ-SNAP, AZ-CLIP, and AZ-Halo). The study indicated that the modular adaptors with the same DNA binding protein (AZ-SNAP, AZ-CLIP, and AZ-Halo) showed good orthogonality; the better result was obtained by the combination of ZF-SNAP, AZ-CLIP, and AZ-HALO, where the orthogonality was governed both by the substrate selectivity of protein-tag and the sequence specificity of the DNA binding module. AFM imaging showed that the co-assembly yields of ZF-SNAP, AZ-CLIP, and AZ-Halo were more than 90% in 5 min incubation, which is the highest yield among the currently available methods for co-assembling three proteins on a DNA scaffold.

Chapter 3 focused on developing a general method to assemble a number of POIs orthogonally on a DNA scaffold. Kinetic aspects of covalent bond formation between the modular adaptor and its substrate modified oligodeoxynucleotide (ODN) suggested a new strategy for expanding the number of orthogonal modular adaptors by combining the same protein-tag with different sequence-specific DNA binding proteins. This next generation of modular adaptors was termed as "sequence-specific modular adaptors", in which the covalent bond formation occurred

exclusively at the specific sequence of DNA binding protein in the substrate modified ODN. The sequence-specific reaction of modular adaptors depends on the balance between the binding rate constants of the DNA binding protein and the rate constant for the covalent bond formation of the protein-tag module to its substrate. The study in Chapter 2 indicated that ZF-SNAP and AZ-SNAP did not show good orthogonality even though the ZF and AZ binding modules showed good selectivity to their respective ODNs. This was due to the reactivity of SNAP-tag to its substrate benzylguanine (BG) being too fast ($k_{cov} \approx 10^4$ $M^{-1}s^{-1}$). The sequence-specific reaction of modular adaptor depended on the balance between the affinity of the DNA binding protein and the reactivity of the protein-tag. The SNAP-tag derived and the CLIP-tag derived modular adaptors reacted to benzylcytosine (BC) with slower kinetic, <10 $M^{-1}s^{-1}$ and 10^2 $M^{-1}s^{-1}$, respectively. Therefore, modular adaptors with SNAP- or CLIP-tag could show the orthogonality to BC modified ODNs. Combinations of SNAP-tag and CLIP-tag with three different DNA binding proteins (ZF, AZ, and GCN4) generated two sets of modular adaptors: SNAP-tag derivatives (ZF-SNAP, AZ-SNAP) and CLIP-tag derivatives (ZF-CLIP, AZ-CLIP, and CLIP-G). The SNAP-tag derivatives and BC modified ODNs showed reasonable orthogonality but the reaction was quite slow. Combination of the CLIP-tag derivatives and BC modified ODNs were preferable with its fast reaction and good orthogonality. By rationally selecting the combination of protein-tag and its substrate from the standpoint of reaction kinetics, I successfully developed the sequence-specific modular adaptors (ZF-CLIP, AZ-CLIP, and CLIP-G), which were derived from three DNA binding proteins (ZF, AZ, and G), respectively, and the same protein-tag (CLIP-tag). We could co-assemble the three modular adaptors on a 3-well DNA scaffold with a co-assembly yield of 87%, while the modification yield at the unmatched position was 11%. In the presence of 200 mM NaCl, dissociation constants (K_D) for the complexes of modular adaptor and ODNs lacking the substrate modification were significantly increased, while the rate constants for the covalent bond formation by protein-tag to the fluorophore modified substrate remained similar. The rate constants (k_{app}) for the reaction of modular adaptor and the substrate modified ODN were also reduced, while the ratio of k_{app} for the matched over unmatched pair increased by more than two orders of magnitudes. This study indicated that the orthogonality was governed by K_D of the

complex or the sequence selectivity of the DNA binding module. Indeed, when three modular adaptors (ZF-CLIP, AZ-CLIP, and CLIP-G) were co-assembled on a DNA scaffold, the yield at the unmatched position on a DNA scaffold was reduced to less than 9% while the co-assembly yield at the matched positions was maintained almost 90%.

I next tested the robustness of the orthogonal modular adaptors by using AZ-CLIP (AC) and AZ-Halo (AH) to assemble the unstable enzyme xylulose kinase (XK) on DNA scaffold in Chapter 4. I succeed to assemble and to characterize XK in the form of adaptor fused enzyme (AC-XK and AH-XK). The cascade reaction of two enzymes derived from xylose metabolic pathway was previously activated on a DNA scaffold (T. A. Ngo, *et al., J. Am. Chem. Soc.* **2016**, *138*, 3012). Here, I further expanded an artificial xylose metabolism pathway by three adaptors fused enzymes: ZS-XR (ZF-SNAP fused xylose reductase), G-XDH (GCN4 fused xylitol dehydrogenase), and AC-XK on a DNA scaffold. The efficiency of the three enzymes cascade reaction was significantly enhanced when these three enzymes were placed at close proximity. The highest cascade reaction efficiency was observed when the three enzymes were placed in 10 nm distance. This study is the first example to investigate the cascade reaction by three enzymes derived from the natural metabolism pathway on a DNA scaffold.

Significance of the research

The methods developed in this thesis solved many current problems encountered in the co-assembly of multiple enzymes on a DNA scaffold. Co-assembly of three proteins on a DNA scaffold with the co-assembly yield more than 90% has demonstrated, which is the highest co-assembly yield reported to date. The method described here also significantly reduced the preparation time of the co-assembly from days to minutes. Furthermore, the number of modular adaptors will be easily increased to locate more than three POIs on a DNA scaffold. By demonstrating the cascade reaction by three enzymes on DNA scaffold, I confirmed the proximity-driven efficient cascade reaction. This is the first study of the cascade reaction by three enzymes derived from the natural metabolic pathway on a DNA nanoscaffold.

Future perspectives

In this research, I developed an effective method for simultaneously locating multiple proteins on a DNA scaffold. This method facilitated the study of the artificial cascade metabolic pathway on a DNA nanoscaffold. The number of orthogonal modular adaptors will be easily expanded to realize more complicated artificial metabolic pathway on surface-confined or three-dimensionally shaped DNA scaffolds. The modular adaptors are also applicable for assembling protein complexes on a designed DNA scaffold to mimic *in vivo* protein assembly systems to study the spatial and stoichiometric effect on the protein compartmentalization that occurred inside the cell. In addition, the principle found in this study on the rational design of the combination of protein-tag and its substrate for the sequence-specific modular adaptors would be also applicable for designing almost any kind of the recognition-driven modifiers, such as the drugs with covalent bond formation ability.

List of publications

1. **Design of modular protein-tags for the orthogonal covalent bond formation at specific DNA sequences**

   Thang Minh Nguyen, Eiji Nakata, Masayuki Saimura, Huyen Dinh, and Takashi Morii

   J. Am. Chem. Soc. **2017**, *139*, 8487

   (Chapter 2, 4)

2. **Development of sequence-specific modular adaptors**

   Thang Minh Nguyen, Eiji Nakata, Masayuki Saimura, and Takashi Morii

   To be submited

   (Chapter 3)

Other Publication

3. **DNA binding adaptors to assemble protein of interest on a DNA scaffold.**

   Eiji Nakata, Huyen Dinh, Thang Minh Nguyen, and Takashi Morii.

   Methods Enzymol. vol. 617, **2018** *(in press)*

4. **Review article: DNA-binding proteins as the potential adaptors to locate functional protein on DNA nanoscaffold**

   Tien Anh Ngo, Huyen Dinh, Thang Minh Nguyen, Fong Fong Liew, Eiji Nakata, and Takashi Morii.

   Manuscript in preparation