Construction of biomacromolecular assemblies with spatially arranged functional units to assess the cellular functions

ZHENGXIAO ZHANG

Kyoto University

Preface and Acknowledgements

The studies in this thesis have been carried out under the direction of Professor Takashi Morii at the Graduate School of Energy Science, Kyoto University, from Nov. 2017 to Jun. 2022. The studies focus on the construction of biomacromolecular assemblies with spatially arranged functional units to assess the cellular functions.

First of all, I wish to express my grateful acknowledgment to Professor Takashi Morii for his guidance, valuable suggestions, fruitful discussions and encouragement throughout the work. I am deeply grateful to Associate Professor Eiji Nakata, Junior Associate Professor Arivazhagan Rajendran, Dr. Shun Nakano and Dr. Dinh Thi Thu Huyen for their kind advice and helpful discussions.

I also would like to express my appreciation to Professor Isao Saito for his helpful suggestions and encouragements.

I thank the past and the present members of Professor Morii's laboratory for their suggestions, cooperation, and discussions. In particular, I express my appreciation to Associate Professor Eiji Nakata, Junior Associate Professor Arivazhagan Rajendran, Dr. Shun Nakano, Dr. Dinh Thi Thu Huyen and Dr. Masayuki Saimura, for their technical supports, kind guidance, fruitful discussions, and encouragements. I also thank Dr. Nguyen Minh Thang, Dr. Shunsuke Tajima, Dr. Peng Lin, Mr. Hiroaki Konishi, Mr. Musashi Shimizu, Mr Yoshitaka Toyoshima, Mr Tomohiko Wakisaki, Ms Khongorzul Gerelbaatar, Mr Kouhei Muroi, Ms Morita Yuki, Mr Daiki Seko, Mr Kirankumar Krishnamurthy, Mr Shiwei Zhang, Mr Taiga Baba, Ms Yaqi Zhang, Ms Wanqing Hou, Mr Futa Komatsubara, Mr Yuya Shibano, Ms Mei Nakabayashi, Ms chardet Crystalle,

Ms Asif Mashal for their helpful comments and discussions. I also express my great appreciation to my friends, colleagues and classmates, past and present, in Kyoto University.

Heartfelt gratitude to Ms. Kaori Hashimoto and Ms. Yukie Kajikawa whose kindness and generosity of spirit made my stay and study in Japan unforgettable and pleasant.

I would like to express my gratitude and appreciation to my thesis committee members, Professor Masato Katahira and Professor Takashi Sagawa for their faithful discussions and insightful suggestions. I thank to Professor Kazunari Matsuda, Professor Shiroh Futaki, Program-Specific Associate Professor Hisaaki Hirose and Dr. Jan Vincent V. Arafiles for the collaboration work and supports in my projects.

Finally, I would like to express my deep appreciation to my mother Wei Deng and my father Xiaoyu Zhang for their constant encouragements and full support.

ZHENGXIAO ZHANG

Laboratory of Biofunctional Chemistry

Department of Fundamental Energy Science

Graduate School of Energy Science

Kyoto University

June 2022

Table of Contents

CHAPTER 1

General introduction

1.1 Conjugation of functional molecules

Biomolecules, such as carbohydrates, proteins, lipids and nucleic acids for the basics of all life beings, occur in great diversity with different properties and activities to mediates all the biological processes in a cell, for instance catalysis, transport, structure and motion. Importantly, functional biomolecules are featured by complexity of conformation and diversity of reactive properties to direct versatile biomolecular interaction and reaction.[1–4] Those active biomolecules enable multiple modifications via the noncovalent binding or the covalent conjugation by other functional molecules, such as fluorophores, small ligands, aggregation-induced modules, etc.[5–7] Such conjugation processes incorporating given functional units to biomacromolecules would further facilitate biomolecule-assisted cellular studies (Figure 1.1).[8–10]

The construction of functional biomolecules requires highly specific conjugation reactions to incorporate the functional units of interests on the biomolecules with selective control and spatial arrangement. Over past two decades, many methods have been reported to tether a relatively simple set of functional groups to biomolecules through a limited variation of chemical bonding strategy (Figure 1.2).[11,12] For those strategies to proceed the bioconjugation, achievement of precision in the chemical reactions remain to be a persistent challenge. The precision requires not only the specificity of a given type of conjugation reaction, but also the orthogonality to other types of chemical reactions: the specificity in modifying the target site and orthogonality in discriminating the target site among multiple groups.[13] Tradionally, the specificity and the orthogonality of reactions completely relied on the selectivity of the chemical reaction, namely chemoselectvity, based on the reactivity of various nucleophilic residues or incorporated noncanonical functional groups.[14] In this case, the orthogonality of chemical reactions limited the usage of multiple reactions for bioconjugation. Therefore, the affinity and specificity of the complex formation with the biomolecule should play a role in controling

the selectivity of conjugation reaction to expand the types of selective conjugation reactions for different biomolecular targets.[15–17] Beside the specificity and the orthogonality of reactions, it is also important to consider the influence on bioactivity and performance of target biomolecules modified by different bioconjugation methods.[18,19]

As an example showing the crucial role of biomolecular interaction for efficient modification of biomolecules, in our laboratory, we took advantage of the specific interaction between RNA and a peptide to covalently conjugate the RNA molecule and the fluorophore-modified Rev peptide in the ribonucleopeptide (RNP) complex.[20] An RNA containing the RRE (Rev responsive element) sequence forms a stable and structurally well-characterized RNP complex with the Rev peptide.[21] By tuning the ligand binding and fluorescence characteristics of the fluorescent RNP, a fluorescent RNP receptor library for the screening of RNP sensors with distinct optical response can be constructed for the target ligand (Figure 1.3a).[22–24] In order to improve the usage of RNP sensor at its low concentration, covalently crosslinking the RNA and the peptide subunits of RNP receptor was proved as a novel strategy to produce stable fluorescent RNP sensors (Figure 1.3b).[20] To construct a stable fluorescent RNP sensor, the 3'-terminal of RNA subunit and the C-terminal of Rev peptide were modified by a 3'-dialdehyde group and a hydrazine group, respectively, and they were covalently tethered by a hydrazone bond formation. It was demonstrated that the covalently conjugated RNP sensors showed enhanced thermostability and enabled the detection of ligands at lower concentrations of sensors. The conjugation reaction of a truncated peptide possessing only the linker moiety reacted with the RRE-containing RNA scaffold showed a significant decrease in the yield (Figure 1.3c).[25] This result supported the notion that the specific complex formation by RRE RNA and the Rev peptide mainly contributed to increase the reaction yield. The specific recognition of RRE RNA by the Rev peptide is the main force to drive the efficient linkage formation between peptide and RNA. Promotion of the conjugation

reaction by the non-specific interaction was also observed, indicating the complicated properties of the intramolecular modification reactions. The specific recognition of biomolecules would be a critical interaction to achieve the precision and the efficiency for functionalization through the affinity-directed conjugation.

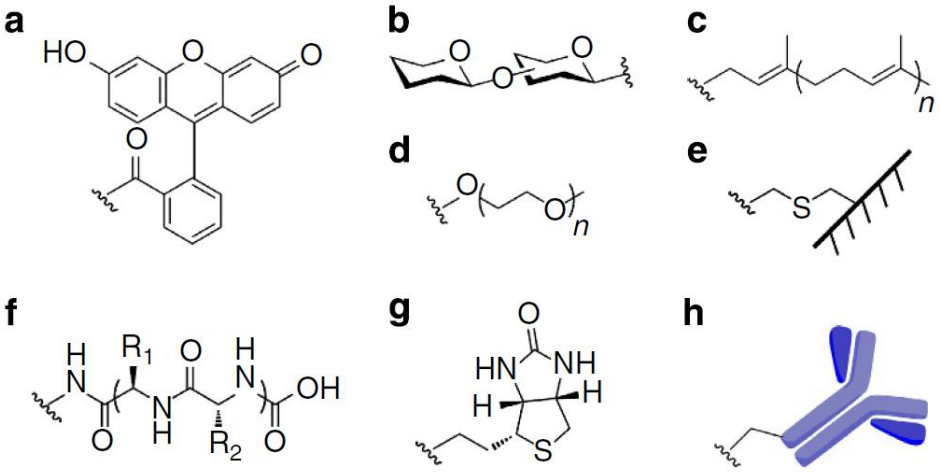


Figure 1.1 Structures of the functional units modified on the biomacromolecules through bioconjugation.[10] (a) fluorophore modification, (b) glycosylation, (c) prenylation, (d) PEGylation, (e) attachment to solid surfaces, (f) peptide conjugation, (g) biotinylation, and (h) antibody conjugation.

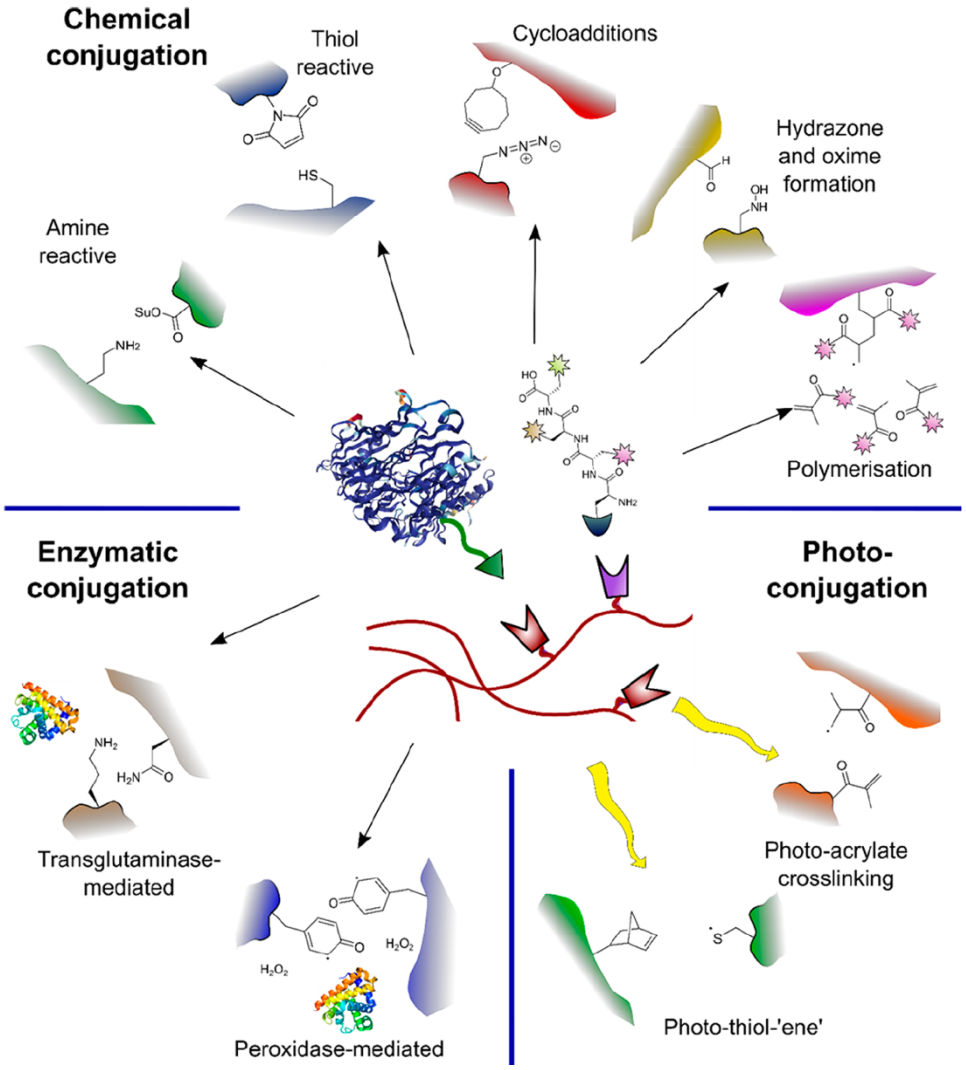


Figure 1.2 Key methods for the conjugation of functional units to biomacromolecules.[26] Chemical conjugation techniques are the widely studied and applied method for the modification. Active chemical groups, such as amine and thiol, and reactions, such as cycloaddition and hydrazone formation, are utilized as reactive handles for the covalent bond formation. Besides, enzymatic methods for the conjugation with polypeptides or oligonucleotides provide highly specific conjugation under mild condition. Also, photo-mediated conjugation can offer the spatial and temporal control for directing the conjugation reaction.

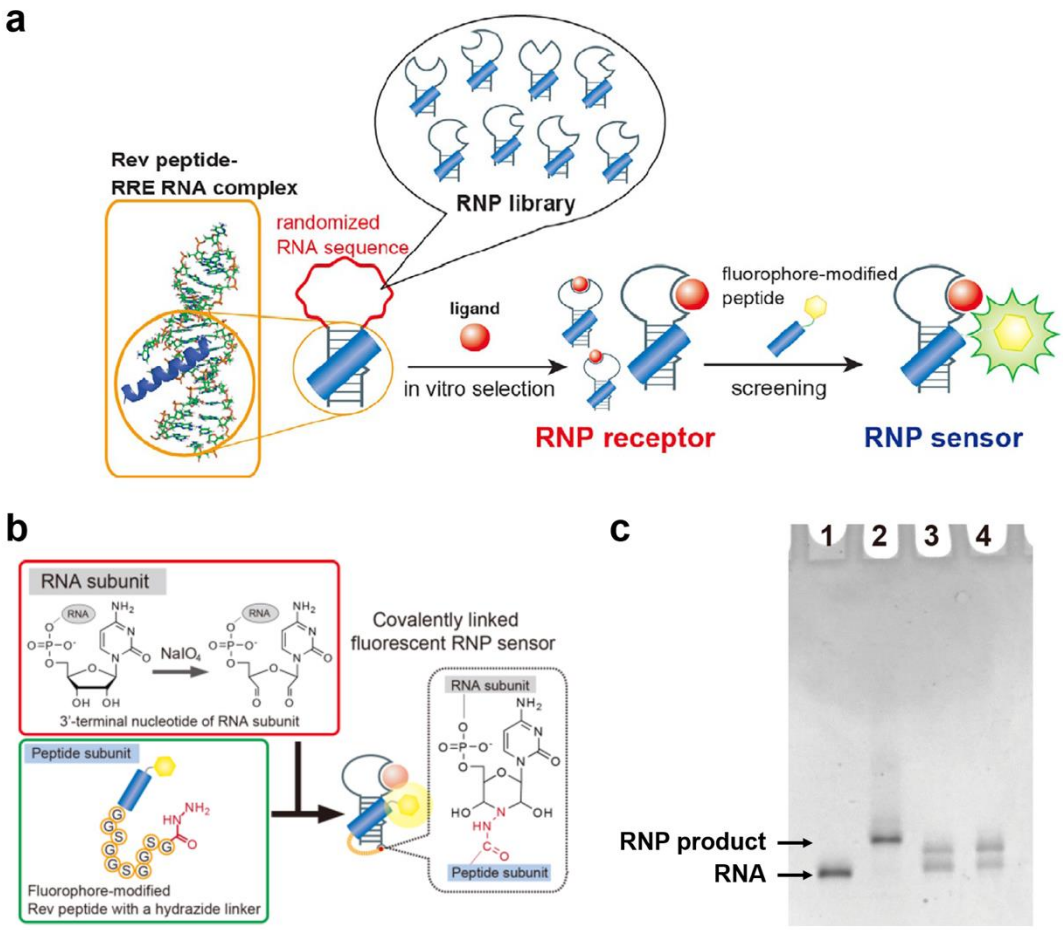


Figure 1.3 Ribonucleopeptide (RNP) framework (a) A scheme illustrates construction of fluorescent sensors by stepwise molding of RNP. [22] RNP receptors were selected from the RNA derived RNP library by applying *in vitro* selection. And fluorescent RNP sensors were constructed by complexing a fluorophore-modified peptide with the RNA subunit of RNP receptors. (b) Schematic illustration of the construction of covalently linked fluorescent RNP sensor.[20] 3'-Ribose of the RNA subunit was converted to 3'-dialdehyde. A flexible peptide linker (GGSGGSGGSG) with a carboxyhydrazine was introduced at the C-terminal of the fluorophore-modified Rev peptide. A covalently linked fluorescent RNP sensor was synthesized via hydrazine bond formation. (c) A denaturing PAGE (8 M Urea) analysis of the conjugation reaction products of RNP constructed from RNA and different peptidic moieties.[25] Lane 1: RNA; Lane 2: RNA reacted with the Rev peptide through a peptidic linker for 3 h; Lane 3: RNA with the peptidic linker reacted for 3 h; Lane 4: RNA with the peptidic linker reacted for 14 h.

1.2 Overview of biomacrmolecular assembilies

Among those biomolecules to be targeted for the modification, biomacromolecular assemblies are essential to introduce multiple signals and functions to facilitate delivery of functional units into cell and to arrange immobilization within the molecular level.[26] Biomacromolecular assemblies formed through self-assembly or polymerization of hundreds of individual units/monomers, such as polysaccharides, DNA scaffolds, and scaffold proteins, are attractive species.[27–29] The properties of individual biomolecular units make the biomacromolecular assemblies equipped with remarkable biocompatibility and exponential reactive moieties.[30,31] The increased molecular size and availability of modification handles make the biomacromolecular assemblies to be essential scaffolds to incorporate other functional units and realize spatial arrangements of units of interest with precise molecular design. As an example of biocompatibility, peptides conjugated on polysaccharide scaffolds are able to exert various effects in the cellular processes, including signaling, adhesion, and trafficking.[32] As a significant example of highly-ordered molecular modification, proteins are spatially arranged in nanometer-scale on DNA scaffolds through site-specific conjugation by taking advantage of the programmability of DNA nanostructures.[33] Therefore, the functionalization of biomacromolecular assemblies with functional units is of particular interest to generate novel biocompatible materials for assessing cellular functions and clinical discovery.

Over past decades, biomacromolecular assemblies, such as scaffold proteins, DNA nanostructures and polysaccharides, are gaining more and more attention, due to their availability for a wide variety of modifications including dyes, biotin and chemical handle.[26] Scaffold and anchoring proteins are very essential for the action of signaling regulatory protein.[34] They can bind over two signaling enzymes together to promote enzymatic reaction or signaling by proximity. By recombination of the highly modular scaffold domains, multidomain scaffold proteins were designed and used as protein-based

scaffold.[35] Artificial enzymatic pathway was constructed on the protein-based scaffold via the noncovalent binding of heterologous protein-protein interaction (Figure 1.4).[36–38]

For making a biological nanostructure which allows precise manipulation in nanometer scale, DNA nanostuctures are supposed to be the exact candidate due to the well understood and predictable nature of Watson–Crick base pairing.[39,40] Furthermore, the flexible and dynamic properties of DNA hybridization impart more potential for precise movement control to DNA scaffold, making it an attractive vehicle to load fluorophore, protein and liposome (Figure 1.5).[41–44] Noticeably, DNA scaffold consists of a long single strand of DNA and hundreds of short strands called "staple strands". Those short strands were produced through synthetic and semi-synthetic processes for the modification with chemical handles or functional groups. Therefore, the conjugation of functional molecules on DNA scaffold not only rely on DNA hybridization, but also conducted by the chemical functionality of the DNA building blocks which could be precisely inserted into DNA scaffold.[45]

Natural polymers such as polysaccharides are featured with inherent biocompatibility and biodegradability which are attractive and important to be exploited as a material for biomedical application. Especially, the polysaccharides, such as hyaluronic acid, gelatin, cellulose, dextran and chitosan, provide abundant chemical groups, especially free hydroxyl and carboxyl groups, to be functionalized as a natural polymeric scaffold.[46] Those scaffolds can be designed to delivery drug through physical encapsulation, chemical coupling and even performing as an active targeting ligand (Figure 1.6).[47–49]

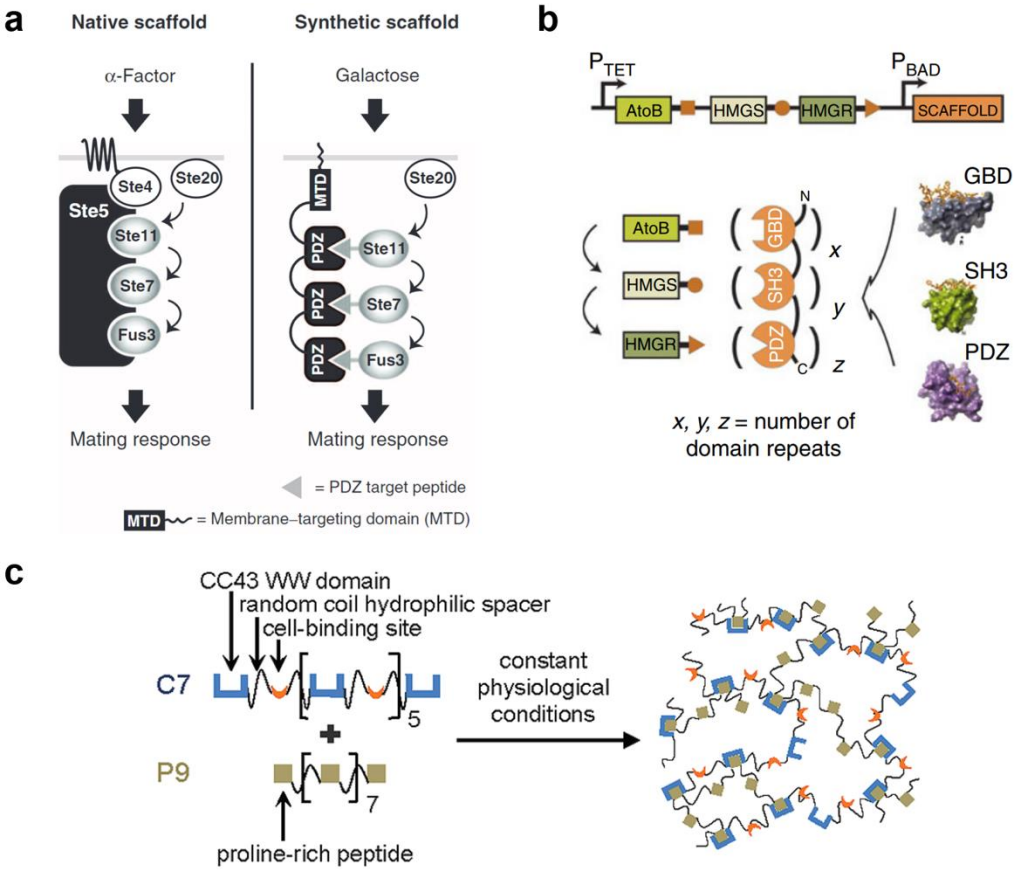


Figure 1.4. Examples of protein-based scaffold. (a) Synthetic scaffolds containing the second PDZ domain of the PSD95 protein was used to recruit Ste11, Ste7, and Fus3, each of which was tagged with a peptide that binds to the PDZ. The MTD localized the synthetic scaffolds to the plasma membrane to mimic protein complexes of Ste5.[36] (b) The synthetic scaffolds were constructed with three protein-protein interaction domains (GBD, SH3 and PDZ), providing modular control over metabolic pathway flux.[37] (c) Protein polymers C7 and P9 contain repeats of the CC43 WW domain and proline-rich peptide. Mixing C7 and P9 results in hydrogel formation.[38]

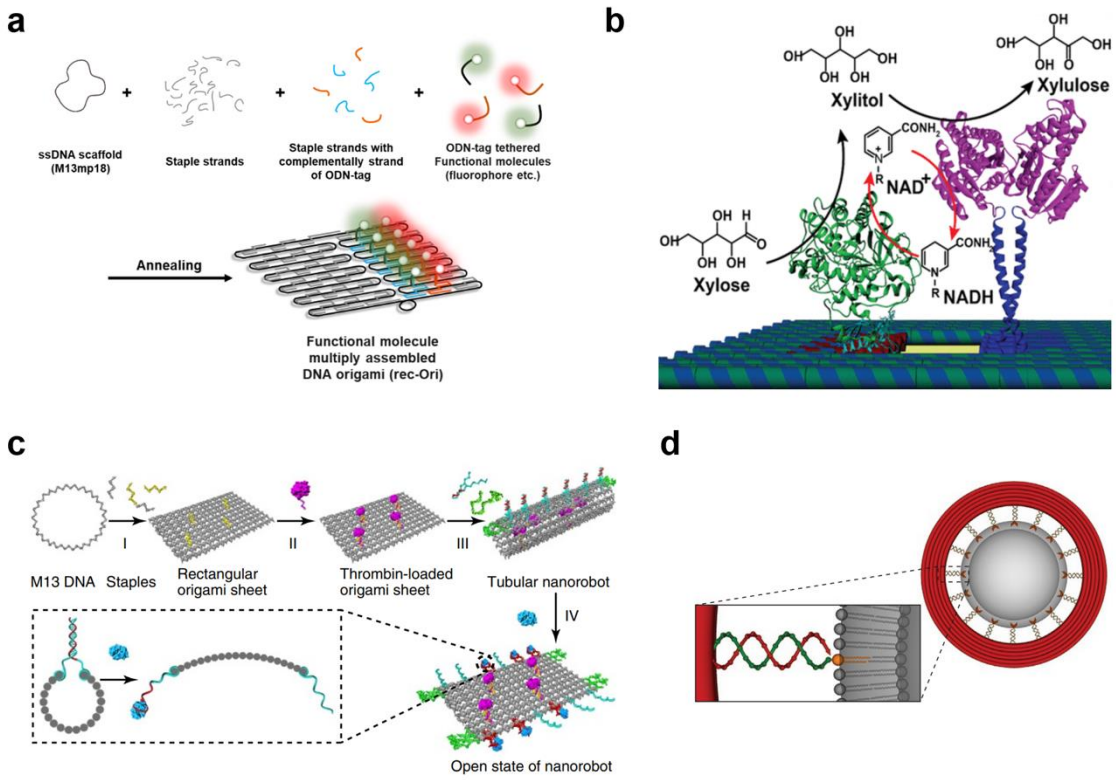


Figure 1.5 DNA scaffold for loading different molecules. (a) A scheme illustrates construction of a DNA origami scaffold (rec-Ori) assembled with multiple fluorophores modified with ODN-tag. Two types of fluorophores were assembled on a DNA origami scaffold for the real-time sensing of cellular pH changes.[43] (b) Enzymatic cascade reactions on DNA scaffold were conducted by the coassembled enzymes ZS-XR and G-XDH on the DNA scaffold.[44] (c) Schematic illustration of the construction of thrombin loaded nanorobot by DNA origami, and its reconfiguration into a rectangular DNA sheet in response to nucleolin binding. Thrombin is loaded onto the surface of the DNA sheet structure.[41] (d) A DNA-origami ring (red) carrying multiple single-stranded extensions (empty handles) was constructed. DNA anti-handles (oligonucleotides with complementary sequence to handles) are chemically conjugated to DOPE and incubated with the DNA ring to allow hybridization.[42]

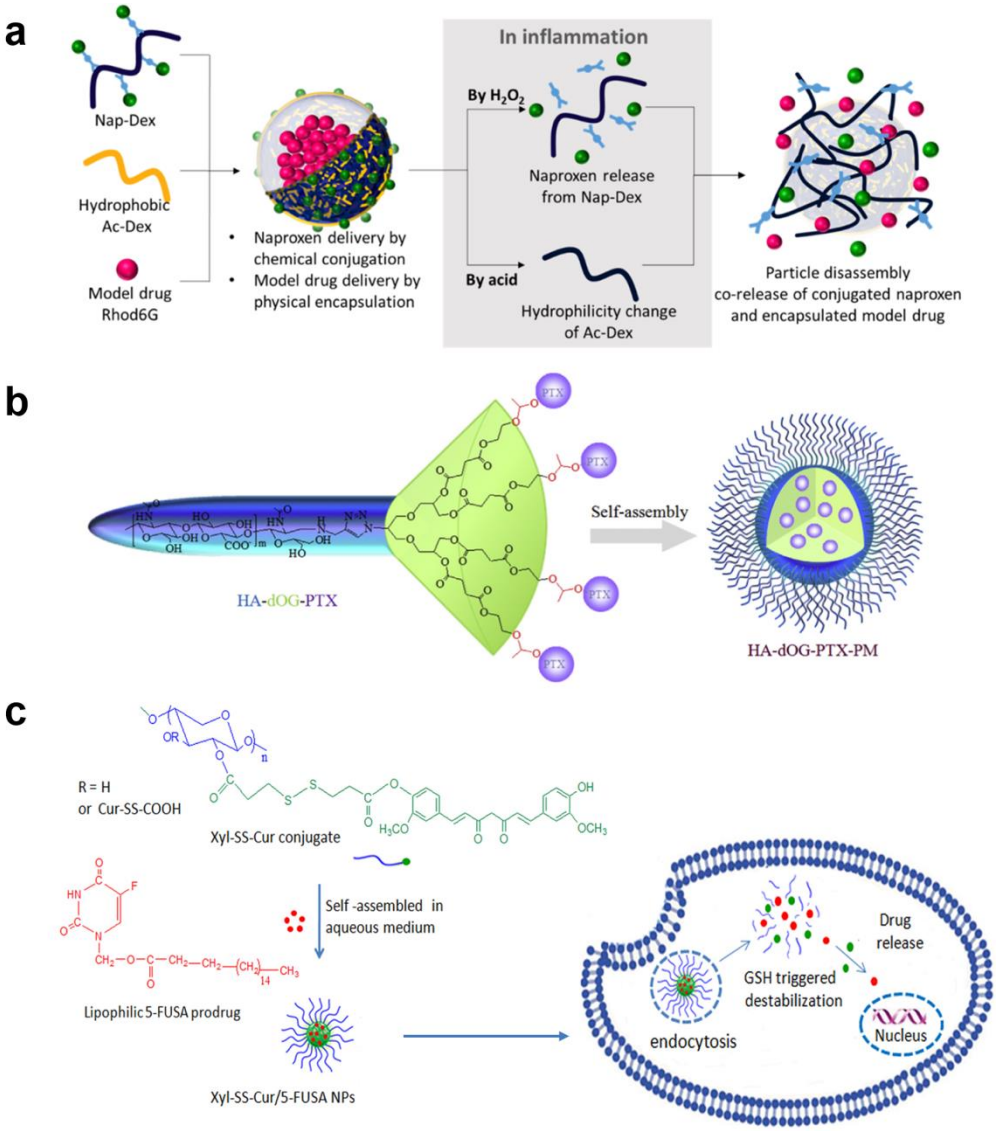


Figure 1.6 Polysaccharides for drug delivery. (a) The anti-inflammatory drug naproxen was modified with the ROS-responsive PBA and conjugated it onto dextran (Nap−Dex) for an inflammation-responsive drug release.[48] (b) Novel hyaluronic acid-shelled acid-activatable PTX prodrug micelles were developed from hyaluronic acid-b-dendritic oligoglycerol block copolymer (HA-dOG-PTX-PM) for effective targeting and treatment of xenografted human breast cancer *in vitro* and *in vivo*.[47] (c) Xylan was conjugated to Cur-SS-COOH to obtain Xyl-SS-Curcumin. The Xylan-SS-Curcumin conjugate were developed for the efficient delivery of 5-FU and curcumin in cancer therapy.[49]

1.3 Current strategies for chemical conjugation on biomacromolecular assemblies

Biomacromolecules have been acknowledged for abundance in reactive groups as nucleophiles for chemical functionalization through bioconjugation. For improving the selectivity of bioconjugation, different side chains of amino acids on protein, for instance, are able to be targeted selectively.[50,51] Furthermore, different reactive groups for bio-orthogonal reaction can also be modified on those side chains as a chemical handle and chemoselectivity can be utilized to conduct selective reactions (Figure 1.7).[52–54] Except relying on chemoselectivity, the structural complexity of biomacromolecules can play an essential role in intermolecular recognition.[55–57] Therefore, the specificity of bioconjugation can be achieved based on incorporating a recognition module to mediate specific interaction between the biomacromolecule and the functional unit to be incorporated.[15,58] The interaction between these two molecules can enhance both the selectivity and the reaction kinetics by virtue of proximity effect.[59] Consequently, recognition-driven conjugation can impart outstanding specificity and high efficiency to enable numerous bimolecular conjugation, including fluorophore labeling, protein glycosylation and DNA-protein conjugation.[60–63]

In the case of modifying biomacromolecular assemblies, bioconjugation is still a robust tool for the functionalization. Especially for polysaccharides and DNA scaffold, functional units can be also modified through the strategies I decribed above, chemoselective conjugation and recognition-driven conjugation.[15,64,65] Plenty of nucleophilic groups on polysaccharides allow the modification of different chemical handles for conducting chemoselective conjugation. For DNA scaffolds, owing to its remarkable addressability, the spatiotemporal organization of functional units on DNA scaffold could be realized by precise DNA hybridization and spacific DNA-protein conjuation.[66] In this chapter, only the covalent conjugation is focus, which is related to polysaccharides and DNA scaffolds.

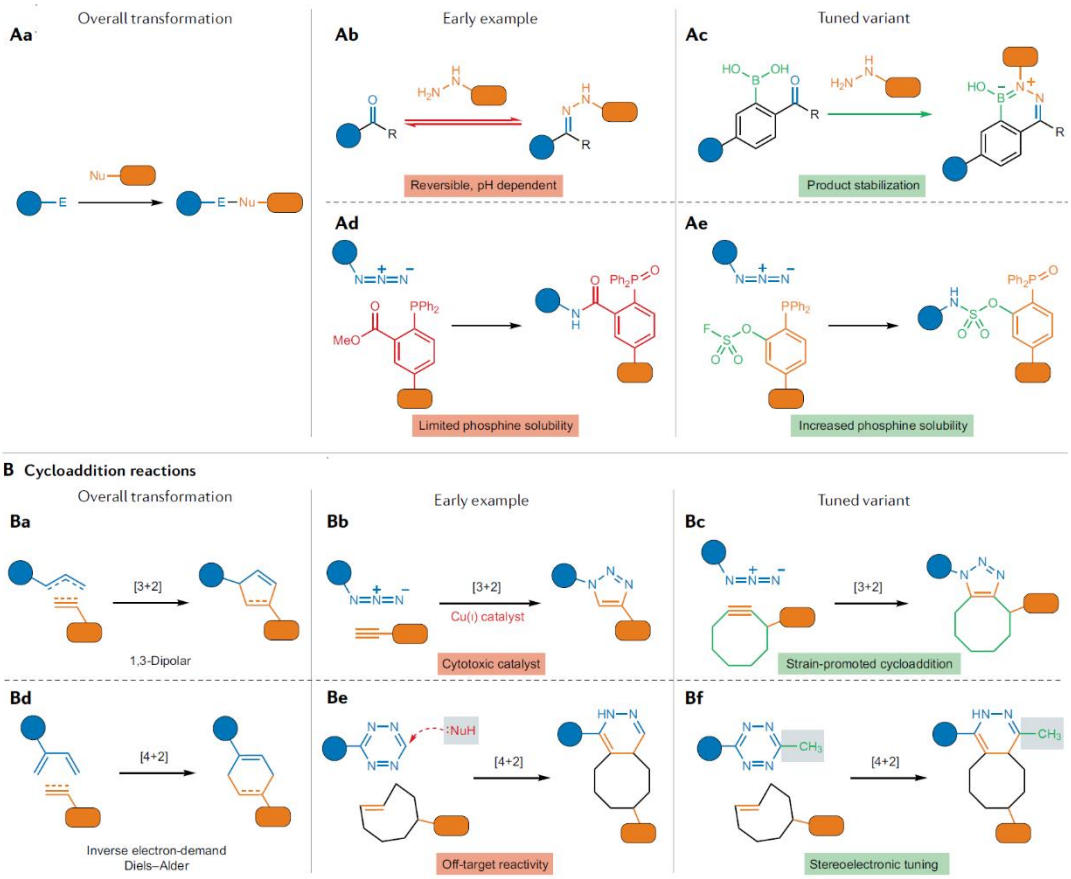


Figure 1.7 Two most common classes of bioorthogonal chemistry: polar reactions and cycloadditions.[54] Bioorthogonal chemistries largely fall into two categories: polar reactions (A) and cycloadditions (B). (A) Aldehydes and ketones were among the first reagents used as bioorthogonal labels. The transformation ultimately linked the two reactants through an amide bond. This variant, termed Staudinger ligation, was amenable to tagging azides in various complex environments, including live cells. (B) Cycloadditions are popular bioorthogonal transformations. One of the earliest exploited was [3+2] cycloaddition with azides. A key breakthrough was the recognition that the cycloaddition could be accelerated using Cu(I). Then, an entire family of strain-promoted azide−alkyne cycloadditions (SPAACs) minimize toxicity in living systems. One notable class comprising inverse electron-demand Diels−Alder (IEDDA) cycloadditions have been developed to address the need for fast-acting, biocompatible reagents.

1.3.1 Modifications on polysaccharides

As I introduced in **1.2**, polysaccharides can offer abundant nucleophiles for chemical modification.[67] However, the selection of the type of chemical modification is crucial considering the situation for modifying multiple functional groups. Thus, modification of chemical handles is essential to carry out orthogonal conjugation on polysaccharides. Unlike protein and nucleic acid, polysaccharides are not genetically encoded and cannot be studied using the standard molecular biology techniques. Polysaccharides were modified with a variety of bioorthogonal functionalities, with azides the most common, but terminal alkynes, strained alkynes and other bioorthogonal groups have also been used, in order to conjugate with polypeptides, lipids and proteins.[65]

As an example for conjugating functional units on polysaccharides via Cu-catalyzed azide-alkyne-cycloadition (CuAAC), Li *et al.* reported an activatable polymeric reporter (termed P-Dex) for dual-modal imaging of malignant breast cancer (Figure 1.8a).[68] For detecting uPA, an NIR dye was connected with the substrate of uPA through a self-immolative linker and emitted fluorescent signal after the cleavage of substrate. In this case, the uPA reporter was modified on dextran (6 kDa) by CuAAC. So that low molecular weight dextran could carry the report to passively accumulate in the tumor. Similarly, strain-promoted azide–alkyne cycloaddition (SPAACs) was used by Fiala *et al.* to conjugate both a monoamine transporter ligand and voltage sensor on dextran (6 kDa) (Figure 1.8b).[69] As a result, the ligand on dextran was able to conduct molecule-specific targeting of voltage sensor dyes by selectively binding the target membrane proteins. The usage of dextran could compensate for the lipophilicity of voltage sensitive dyes by dynamic encapsulation, thus limiting nonspecific binding of the probe to cell membranes while enabling diffusion through tissue.

Except the cycloaddition for molecular modification on polysaccharides, polypeptide could be modified on dextran (70 kDa) through the formation of carbamate.[70]

For this purpose, a pyrrolidine group was modified on the N-terminal of polyarginine peptide through the amino-PEG_8 linker. Cholesterol dextran (70 kDa) was reacted with *p*-nitrophenylchloroformate at the dextran moiety in activated condition. This condition facilitated the reaction with pyrrolidine-3-carboxamido-PEG peptides to finally produce peptide-dextran conjugation (Figure 1.9a). The final products were condensed to form nanoparticles for *in vivo* imaging. Compared with dextran, the amino group on chitosan enabled condensation reaction with carboxylic acid group directly. For example, Ryu *et al*., modified a cathepsin B substrate on the tumor-targeting glycol chitosan nanoparticles for potential tumor diagnosis (Figure 1.9b).[71]

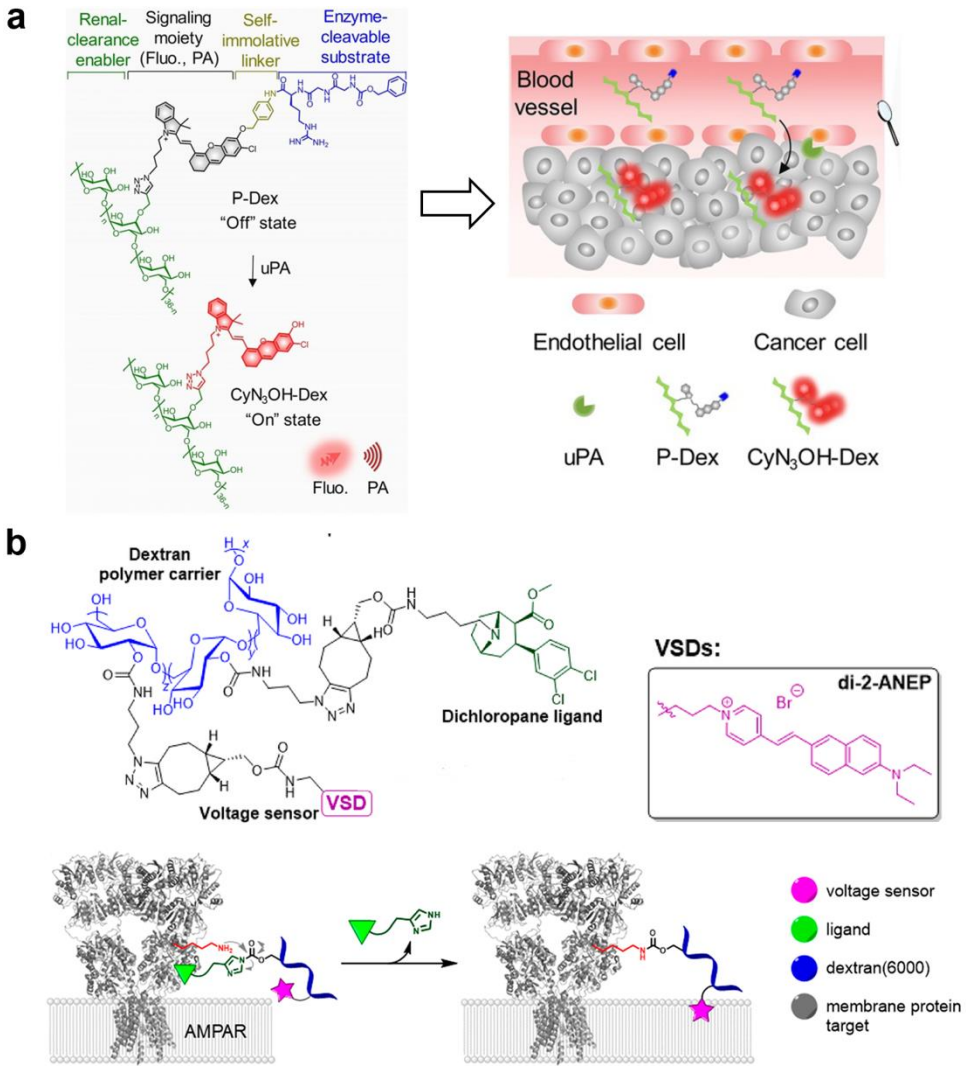


Figure 1.8 (a) CuAAC was used for loading uPA reporter on dextran (6 kDa) which is able to passively accumulate in the tumor.[68] b) Monoamine transporter ligand and voltage sensor were loaded on dextran by strain-promoted cycloaddition. The ligand on dextran enabled molecule-specific targeting of voltage sensor dyes by selectively binding the target membrane proteins.[69]

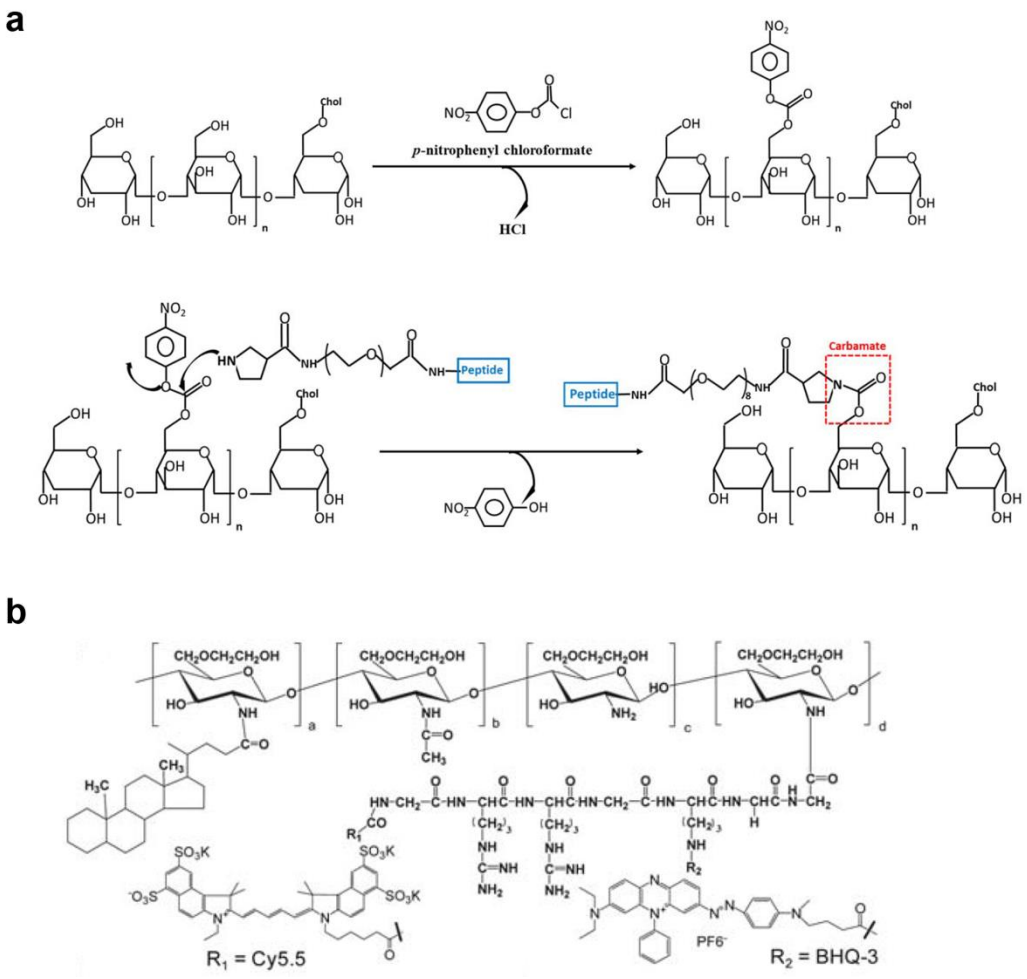


Figure 1.9 (a) Synthetic steps of peptide-dextran conjugates for forming nanoparticles. Glucosyl-OH groups of cholesterol dextran are activated with p-nitrophenyl chloroformate. Pyrrolidyl-PEG-peptide then displaces p-nitrophenol in DMSO to give stable the carbamate-linked peptide.[70] (b) Chemical structure of tumor-targeting glycol chitosan modified with the cathepsin B substrate.[71]

1.3.2 Assembly of functional units on DNA scaffold

Spatially addressable, finite sized DNA nanostructures were applied as the DNA scaffold. It was folded into various geometrical shapes with the help of hundreds of short oligodeoxynucleotides (ODN) serving as “staples” strands. The precise addressability on DNA scaffold is achieved by modifying a specific DNA strand and hybridizing the modified DNA strand into the DNA scaffold.[72–74] Therefore, modification of DNA is the first and essential step to conjugation biomolecules on DNA scaffold. Particularly, ODN or short DNA strand are available by the automatic chemical synthesis on the solid phase.[75] It enables the modification of chemical handle or functional group on the phosphate backbone of DNA, which would facilitate the chemoselective conjugation of DNA with biomolecules.[76] Especially, in the case with protein, the abundant nucleophiles on protein surface can be exploited to conjugate with DNA through homo- and hetero-bio-functional linkers after modifying DNA end with another nucleophilic group (Figure 1.10).[77]

With a proper mothed to achieve conjugation between DNA and target biomolecule, two basic methods can be used to modify the DNA scaffold: (1) DNA strand is modified with target biomolecule first and then the modified DNA strand is integrated on the DNA scaffold through hybridization, or (2) DNA strand is modified with a substrate or a chemical handle and integrated on DNA scaffold and the target biomolecule can directly react with the substrate or the chemical handle located on the DNA scaffold as the way of post-assembly modification (Figure 1.11).[64] Different from the first method, the second method not only requires high chemoselectivity to react the protein, for example, with the substrate or chemical handle on the DNA strand, but also enough efficiency for the conjugation reaction for assuring the assembly yield.

For the first type of methods, it can be also classified as two different ways: (i) the modified DNA strand together with other staple strands are hybridized for folding the

DNA scaffold or (ii) the modified DNA strand is hybridized after the folding of DNA scaffold (Figure 1.12a and b).[78,79] Hybridization of the modified DNA strand together with other staple strands during the folding process of DNA scaffold is normally used on the case when the stability and activity of modified ligand, such as fluorophores, cholesterols, peptide, streptavidin or other chemical handles, tolerate the annealing condition of DNA scaffold (Figure 1.12a and b). However, the bioactive molecules, especially proteins, may lose the activity in the annealing conditions for folding the DNA scaffold. The post-translational chemical modification was found to significantly impair the activity of enzymes.[80–82] Therefore, for preserving the activity of protein, protein is supposed to be chemically conjugated with DNA strand to serve as an anchor to hybridize with the corresponding complementary capture strand that is displayed on the surface of DNA nanoscaffold (Figure 1.12c and d).[80,83]

The modification reaction of short DNA does not require absolute selectivity and efficiency. However, specific conjugation of target molecules on the DNA scaffold relies on the high chemoselectivity of the conjugation with the chemical handle or the substrate labelled on DNA scaffold.[64,84,85] And orthogonality is necessary for the case of multiple types of target molecules. The Gothelf group performed 3 types of chemical reactions with single molecules on defined positions of DNA scaffold (Figure 1.13).[86] The reaction yields and selectivity were investigated by AFM and demonstrated the feasibility of post-assembly chemical modification on DNA scaffold with sufficient yields and good chemoselectivity. When the DNA scaffold is conjugated with protein or enzyme for its functionalization, the reaction efficiency and orthogonality at nano-molar scale are the significant premise to the feasibility in the post-assembly manner. C. Niemeyer group prepared the protein of interest (POI) by genetic fusion with SNAP-tag and Halo-tag.[87] Correspondingly, the substrates of SNAP-tag and Halo-tag, benzylguanine (BG) and chlorohexane (CH) groups, respectively, were incorporated on DNA scaffold as the

ligands for the site-specific coupling with matched protein tag. Thus, site-specific and orthogonal conjugation between POIs and DNA scaffold were performed by the chemoselectivity of SNAP-tag and Halo-tag to their substrate BG and CH (Figure 1.14). Importantly, the conjugation reaction took place through the genetically fused protein tag so that the bioactivity of protein could be completely preserved. However, owing to the concentration limitation of the usage of POIs, the yields for co-assembling several POIs under the low-concentration condition need to be further improved for reliable studies of enzyme cascades on DNA scaffold.

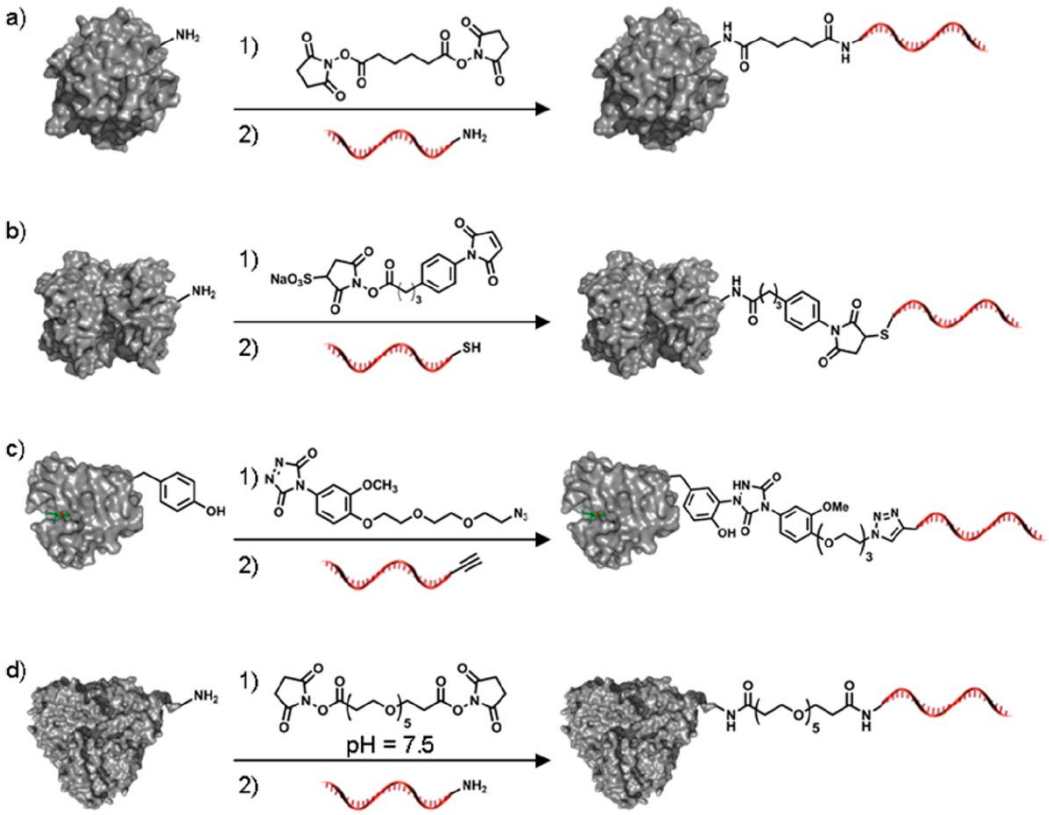


Figure 1.10 Examples of DNA-protein conjugation via bifunctional linker.[77] Here, the terminal of DNA sequence is firstly functionalized by nucleophilic groups, such as amino, alkyne and thiol. And it is further modified by bifunctional linker in order to react with the nucleophilic groups on protein surface.

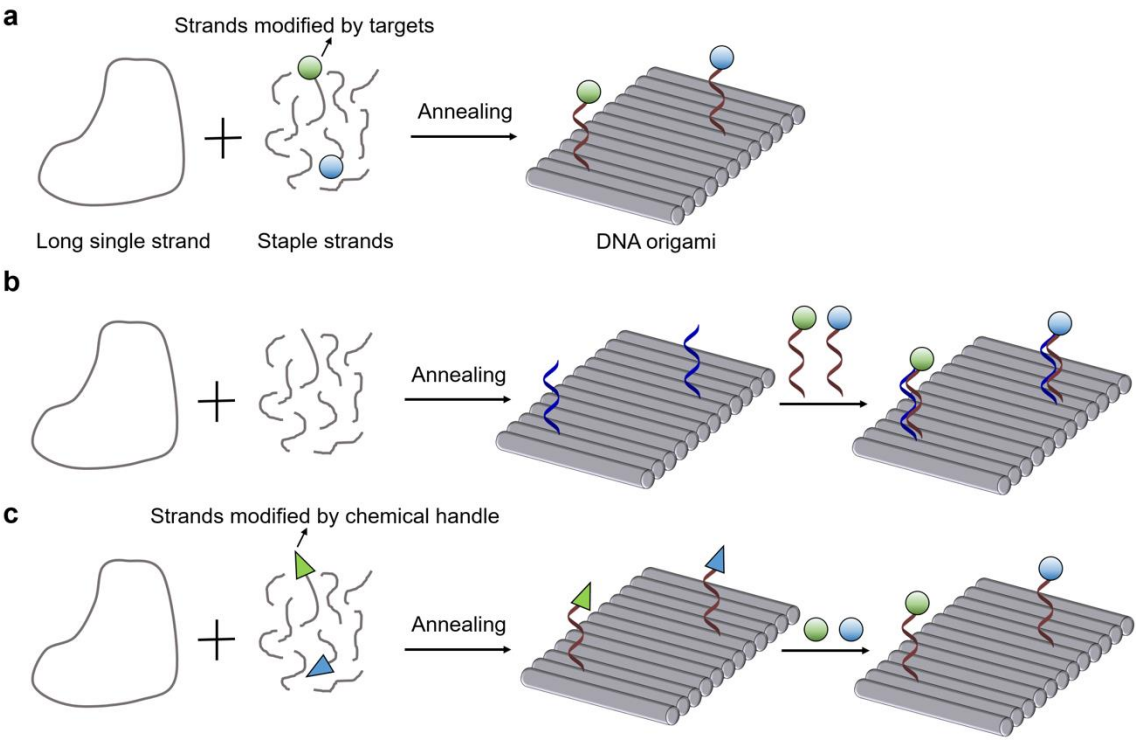


Figure 1.11 Methods for modification on the DNA scaffold. The first type of methods carry out the conjugation of target with DNA strand and (a) to integrate the modified DNA strand during the folding process of DNA scaffold or (b) to hybridize the modified DNA strand to the folded DNA scaffold.[64] (c) The second type of methods, post-assembly modification, the DNA strand modified with the specific substrate for target molecule is incorporated in the DNA scaffold, then the specific substrate on the DNA scaffold directly reacts with the target molecule to complete the conjugation.

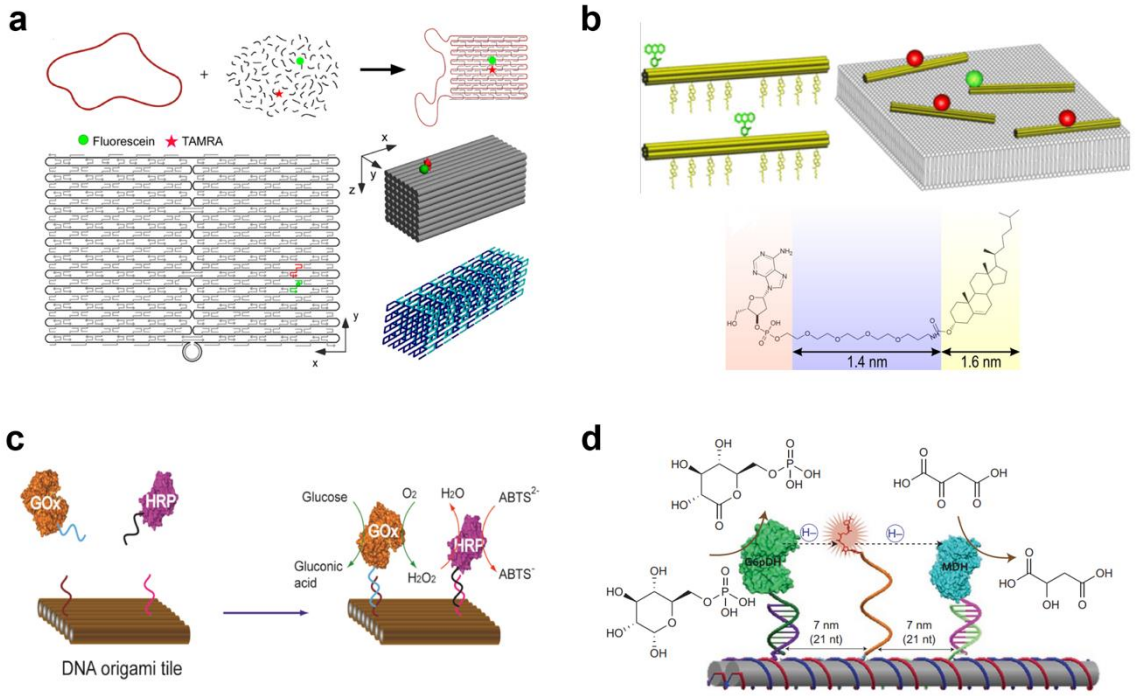


Figure 1.12 Functional molecules assembled on DNA scaffold. (a) FRET donor (fluorescein, green circle) and acceptor (TAMRA, red star) dyes on DNA strands are integrated into the formation of rectangular DNA origami and cuboid DNA origami for monitoring folding process of DNA origami.[78] (b) A cholesteryl-triethylene glycol-modified nucleotide used to produce the membrane anchors. Amphipathic stiff DNA origami nanoneedles in the form of six-helix bundles (6HB) were decorated with cholesteryl-triethylene glycol membrane anchors and fluorescent labels at defined positions for binding to freestanding lipid membranes.[79] (c) The DNA hybridization directed co-assembly of GOx and HRP on DNA origami tiles for controlling inter-enzyme distances.[80] Enzymes were conjugated with DNA strands via the reaction on the nucleophilic groups on enzyme surface. (d) Nanostructured complex consists of G6pDH and MDH organized on a DNA DX tile to explore the effect of substrate channeling inside enzyme cascades.[83] Two enzymes were conjugation with DNA strands and hybridized on DNA origami through DNA complementary base-pairing. In contrast, The NAD^+-modified single-stranded poly$(T)_{20}$ was directly folded into DNA origami.

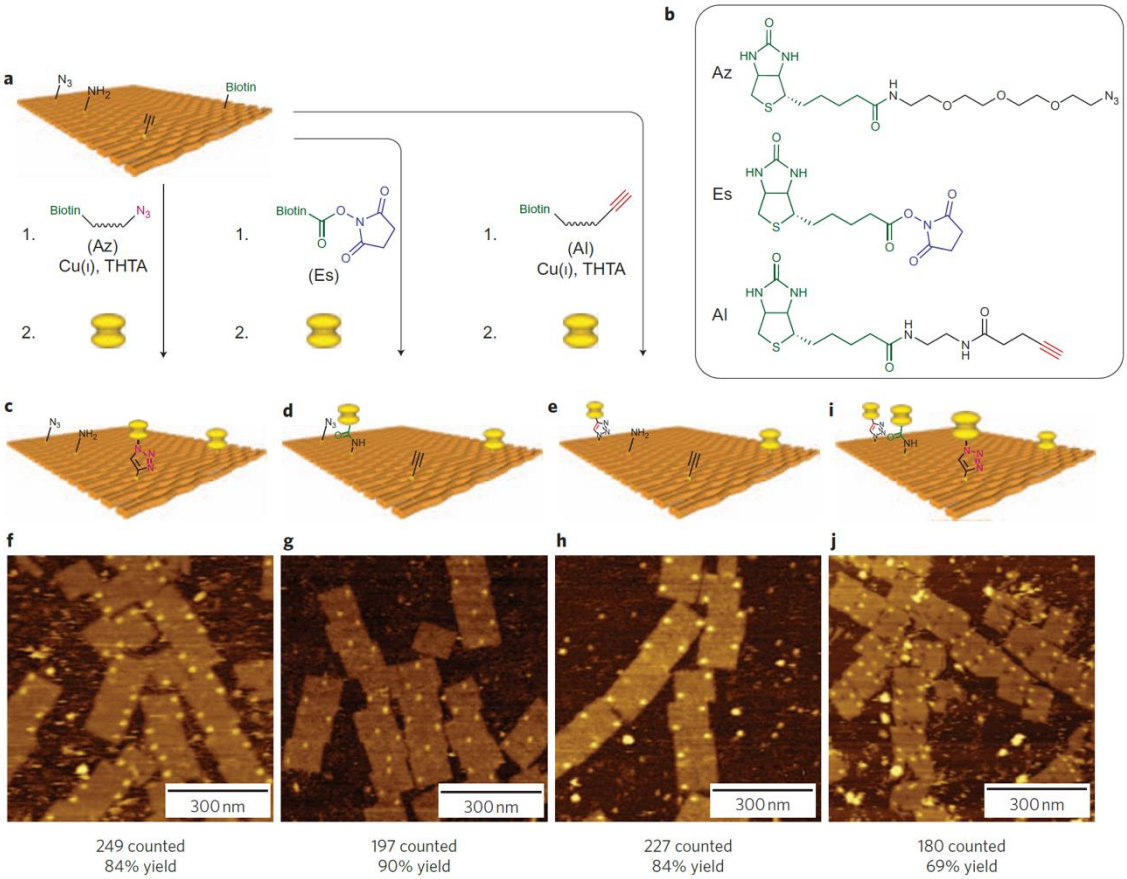


Figure 1.13 Single-molecular chemical conjugation on DNA scaffold.[86] (a)Model of the DNA scaffold incorporating a biotin reference and three functional groups (an alkyne, an amine and an azide) and their reactions with the complementary functionalities (yellow structure, streptavidin). (b) Chemical structure of the biotin-tethered functional groups azide, NHS and alkyne. The functional chemical handles were integrated on DNA scaffold. (c−e) Model of the expected product of the conjugations with azide, NHS and alkyne after incubation with streptavidin. (f−h) AFM images of the respective reaction products illustrated in (c−e), respectively. (i) Model of the expected product after three successive reactions with azide, NHS and alkyne. (j) AFM imaging of the co-assembled product.

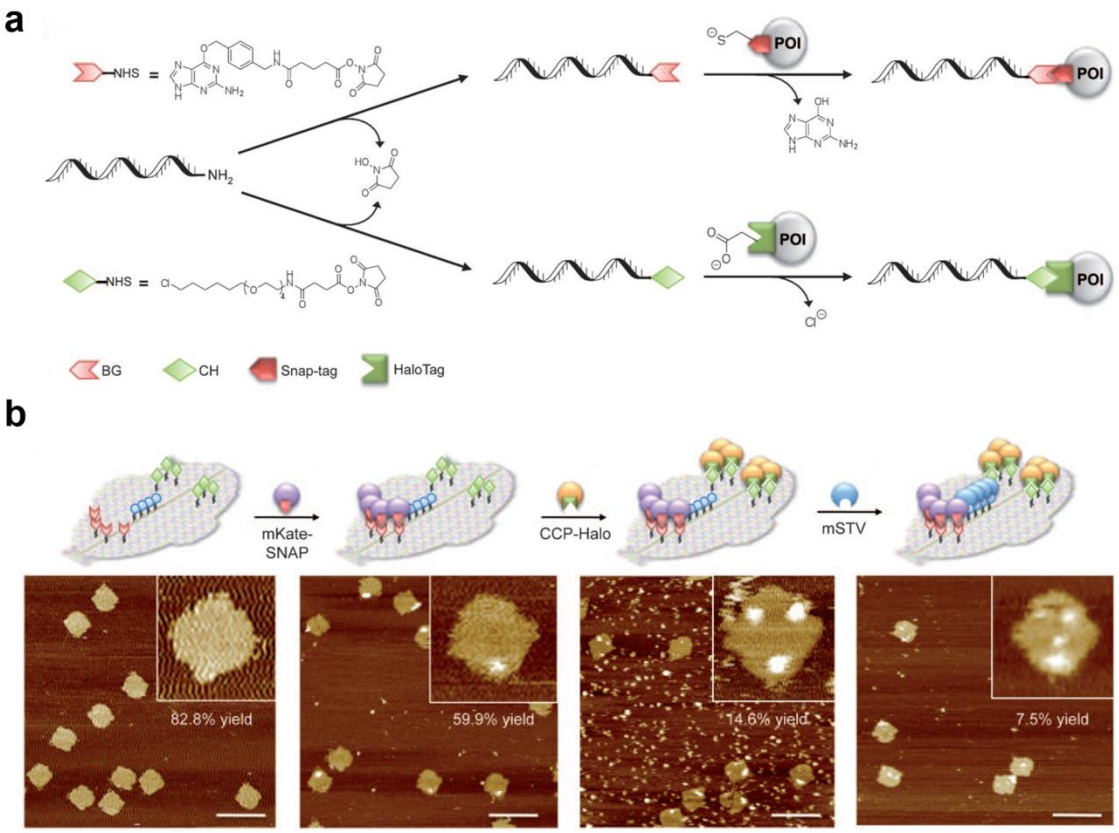


Figure 1.14 Conjugation POIs on DNA scaffold through protein tag.[87] (a) Schematic representation of chemical coupling of amino-modified DNA oligonucleotides with benzylguanine (BG) or chlorohexane (CH) groups as orthogonal tags were designed for chemoselective modification of proteins of interest (POI) which were genetically fused with SNAP-tag or Halo-tag, respectively, at the C-termini. (b) Consecutive orthogonal protein binding was demonstrated by incubation of POI with SNAP-tag (indicated by violet spheres), with Halo-tag (yellow spheres), and with streptavidin (blue spheres). The binding of these proteins led to decoration of the origami with features resembling a mouth, eyes, and a nose, respectively.

1.3.3 Recognition-driven conjugation of proteins on DNA scaffold

As a fundamental biomolecule in life, protein is featured by the complexity of conformation and the diversity of surface properties to direct versatile recognition and modification events.[12,50] By conjugated with DNA, the functional versatility of proteins of interest (POI) would be combined with high programmability of DNA to achieve the design of functional biomacromolecular assmblies in a highly specific manner.[77] As one of the important applications, DNA scaffold has emerged as an ideal nano-carrier for loading protein for precisely spatial localization.[88,89] The spatial programmability for POI promoted various studies on the DNA scaffold, such as enzyme cascades[89], substrate channeling[88] and drug delivery vehicle.[90] To locate POI on the DNA scaffold, the nucleophilic groups on protein surface would be the main targets for the conjugation with short DNA. The conjugated DNA can guide POI to be assembled on the desired position of DNA scaffold through DNA hybridization.[91] This strategy indeed facilitated DNA-protein conjugation, but it showed less co-assembly yields for locating multiple POIs on the DNA scaffold. And it cannot avoid the reduction of POI activity owing to the post-translation modification on the protein surface. [80–82] Therefore, new strategy should be exploited to improve the assembly yields and the selectivity without reducing the activity of POI.

For this purpose, recognition-driven DNA-protein conjugation is attractive for its high efficiency and remarkable specificity. The recognition-driven conjugation takes advantage of the high-affinity recognition and interaction conducted by protein (Figure 1.15).[14,58] The recognition-driven conjugation between POI and DNA is driven a noncanonical functional moiety incorporated on POI, which can mediate specific interaction through DNA base-pairing or DNA-protein binding.[15] The covalent bond formation is hence benefited from the resulting complex formation and triggered by the proximity effect to exhibit high reaction efficiency.[59] Besides the acceleration of

conjugation reaction, the molecular recognition step is a vital key to exert the reaction selectivity among different POIs to the designed site of DNA scaffold.

Affinity-directed DNA-protein conjugation utilizes metal-affinity probes, peptides or aptamers as a protein binder to modify the DNA template sequence and hence renders proximity effect by binding with POI to boost the efficiency of crosslink reaction between target DNA sequence and POI (Figure 1.16).[82,92,93] The strategy incarnates the concept of recognition-driven conjugation to improve the reaction specificity and yields between one type of POI and target DNA sequence. However, the selectivity of target ODN to various POIs remained to be optimized by developing different specific binders. Also, the conjugation reaction of post-translation modification on the nucleophilic groups of POI could cause negative impacts on the protein bearing highly reactive groups.

In our laboratory, we genetically fused POI with a modular adaptor (MA) for assembling multiple POIs on the DNA scaffold. MA consists of a self-ligating protein tag and a DNA binding domain.[94] On the condition that the substrate of protein tag is tethered to the MA binding DNA sequence, MA and the substrate-modified DNA binding sequence form DNA-MA complex to increase the effective molarity of substrate. Hence, the covalent linkage formation is expedited with fully retaining the activity of MA-fused enzyme. For instance, SNAP-tag was used as the protein tag and zif268 was used as the DNA binding protein to produce an MA, ZF-SNAP, to reach 97% assembly yield on the DNA scaffold which contains ODN with the zif268 binding sequence modified by the substrate of SNAP-tag (Figure 1.17a).[94] With the use of multiple POI-fused MAs and their respective DNA binding sequence modified with the specific substrate of each MA, spatial organization of POIs on the DNA scaffold was realized to construct an artificial enzyme cascade (Figure 1.17b and c).[44,95]

Particularly, besides the fast reaction kinetics, each component of MA, DNA binding protein and protein tag, shows prominent specificity for noncovalent and covalent

conjugation, respectively, to diversify the selectivity of MA. DNA binding domain is featured with selectivity to its specific DNA binding sequence. The protein tag shows remarkable chemoselectivity to its substrate. Previously SNAP-tag, CLIP-tag and Halo-tag were combined with different zinc finger proteins to create 3 unique MAs (Figure 1.18a).[96] Chemoselectvity of protein tag drove 3 MAs to orthogonally react with the specific substrates on the respective DNA binding sequences. However, it is difficult to increase the types of protein tag to expand the number of selective MA. To this end, a novel recognition-driven strategy was developed, in which the selectivity of MA crosslink reaction is exclusively governed by the distinct DNA sequence specificities of the DNA binding domain. CLIP-tag was exploited as protein tag of MA and combined with 3 types of DNA binding domains to provide 3 unique MAs (Figure 1.18b).[97] The resulting 3 MAs efficiently recognize and react with their own specific DNA binding sequences which were modified by the substrate of CLIP-tag, O^2-benzylcytosine (BC). The facts revealed that the specific crosslinking reaction of CLIP-tag was driven by the DNA recognition process because the reactivity of CLIP-tag to BC was kinetically favorable to realize the adjustment of reaction rate through DNA-MA binding process. Accordingly, recognition-driven CLIP-tag was created as MAs comprising CLIP-tag (MA-CLIP) to be sequence-selectively conjugated on DNA scaffold containing BC-modified DNA.

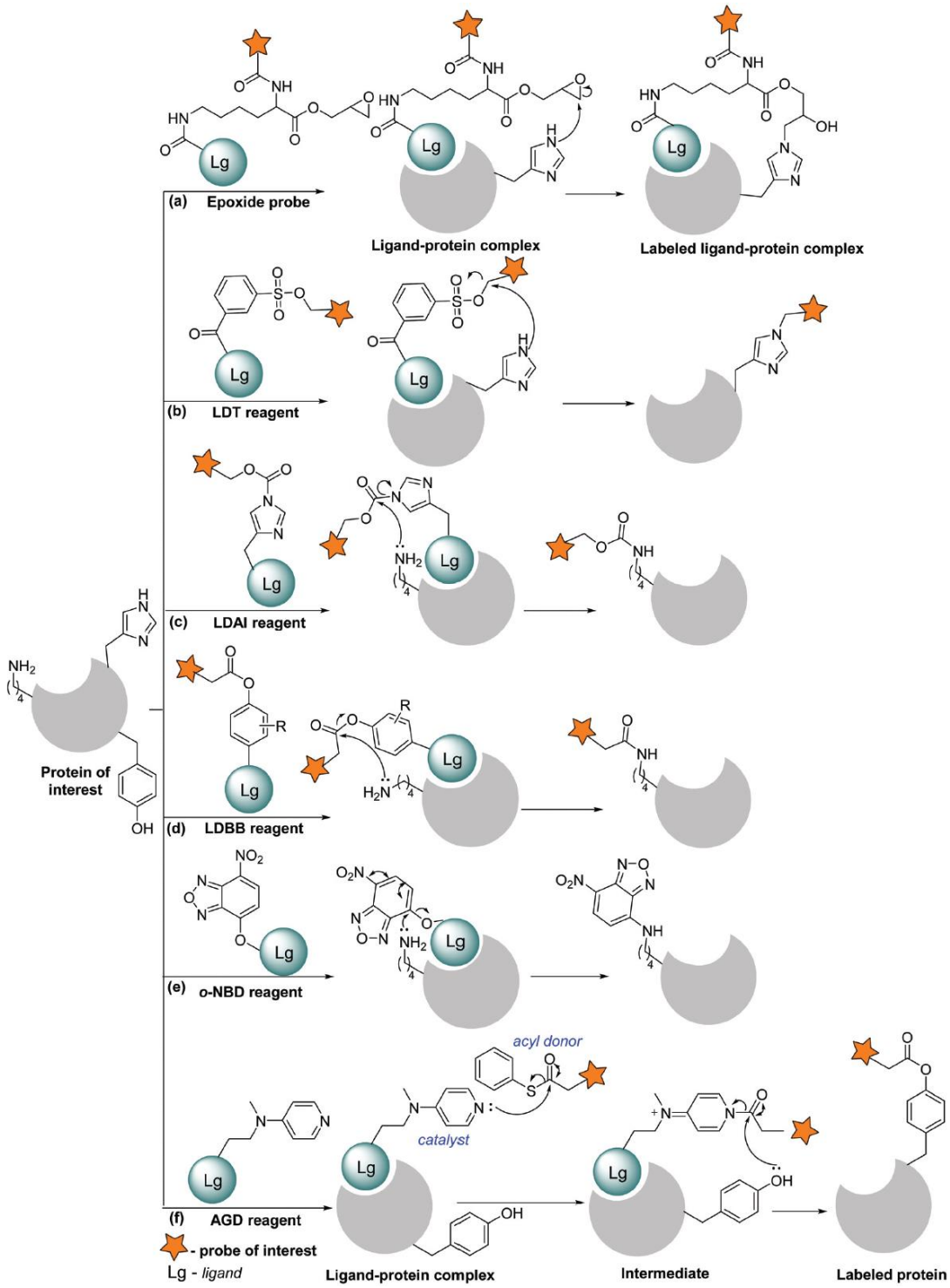


Figure 1.15. Ligand-directed modification of POI.[14] Ligand-directed modification are one of examples of recognition-driven reaction for protein labeling. For the principle of this design, the binding moiety on the target molecule drives the specificity to bind with POI which leads to proximity effect between the reactive moiety and nucleophilic group on protein. So that the reaction is accelerated by reactive moiety to complete the labeling process on protein.

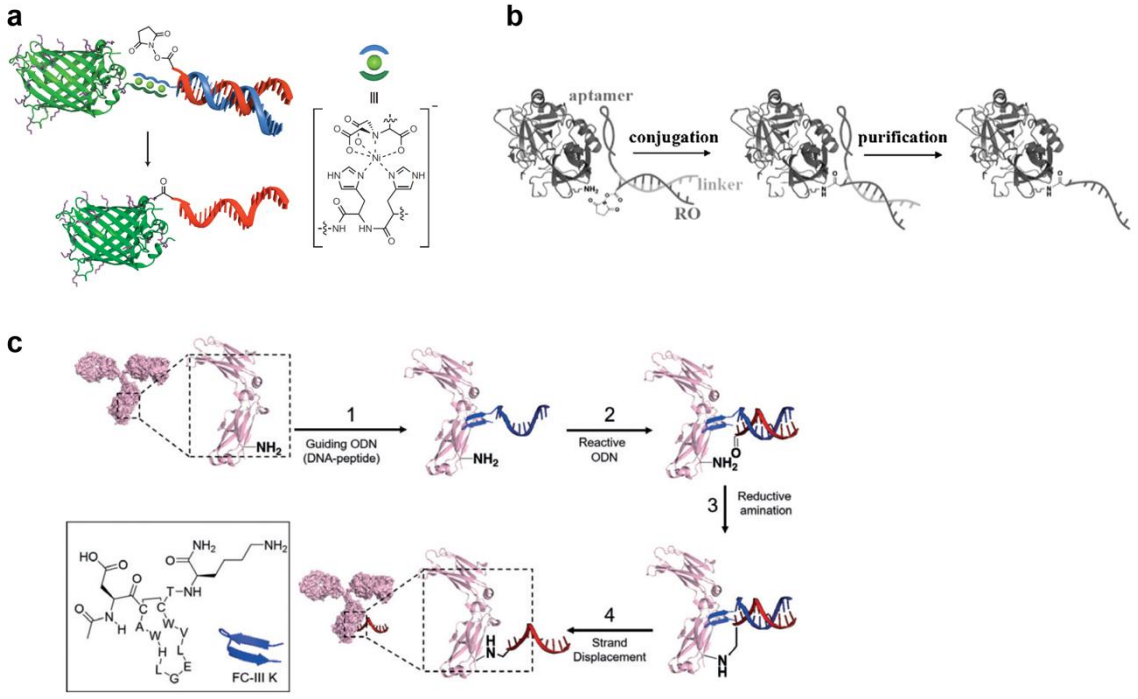


Figure 1.16. Affinity-directed DNA-protein conjugation. (a) Conjugation was driven by the coordination between the *tris*-(NTA) group on the ODN and His_6 tag on GFP through a Ni(II) complex.[82] The NHS ester on the complementary ODN was approached near the GFP by DNA hybridization to react with the lysine residues on the GFP mediated by proximity effect. (b) Aptamers are single-stranded oligonucleotides selected by SELEX to bind specifically to a target molecule, and they can be used as templates to introduce a reactive functionality on a protein.[92] (c) A protein-binding peptide conjugated with ODN functioned as a guiding moiety as the template. The aldehyde group on the complementary ODN reacted with protein lysine residues.[93]

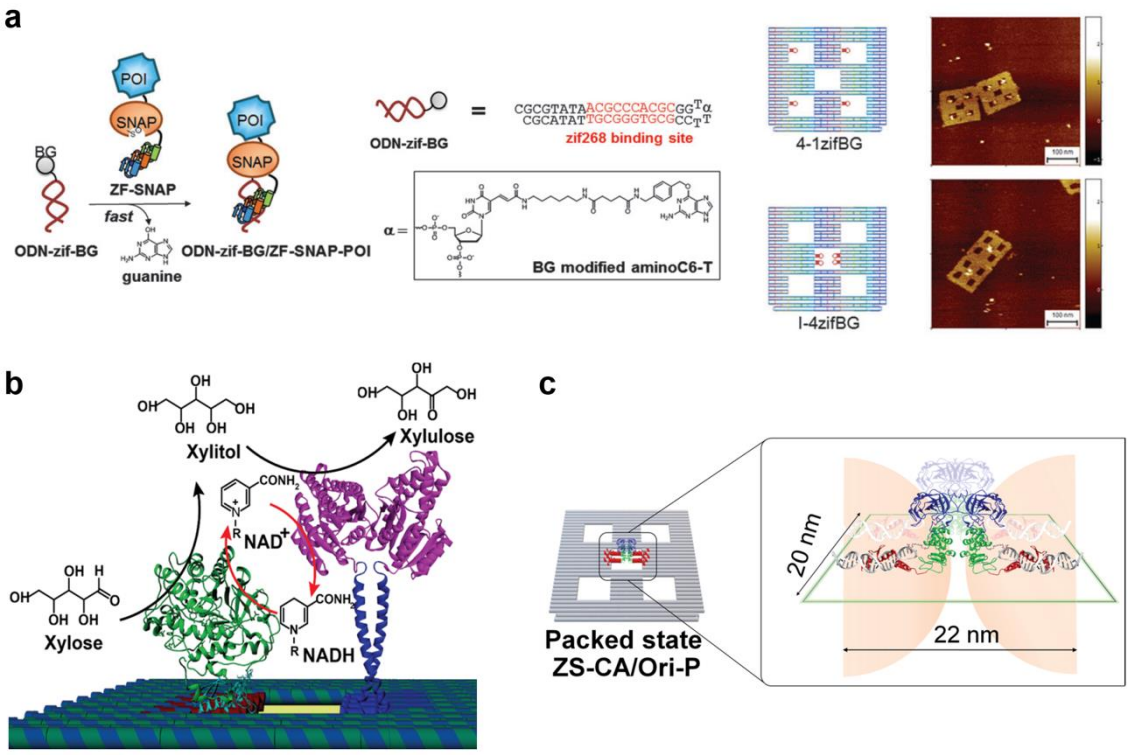


Figure 1.17. Modular adaptor for recognition-driven conjugation on DNA scaffold. (a) Modular adaptor (MA) consists of a DNA binding domain (zinc finger (ZF) protein) and a protein tag (SNAP-tag) and is fused with POI.[94] ZF domain assisted rapid formation of the covalent linkage by proximity effect between the SNAP-tag and a benzylguanine (BG) modified oligonucleotide (ODN-zif-BG). While the BG-modified ODN was folded into 5-well DNA origami, MA is able to be localized POI on the defined position. (b) Enzymatic cascade reactions including enzymes ZS-XR and G-XDH were assemble on the DNA scaffold conducted by ZF-SNAP fused with XR.[44] XDH was fused with GCN4 for attaching on the DNA scaffold. The distance between XR and XDH can be controlled on DNA scaffold. (c) Carbonic anhydrases fused with ZF-SNAP were assembled as the packed state on DNA scaffold which contained 4 ODN-zif-BG in the central well.[95] The model was used to study the acceleration of enzymatic reaction in the packed assembly of enzymes.

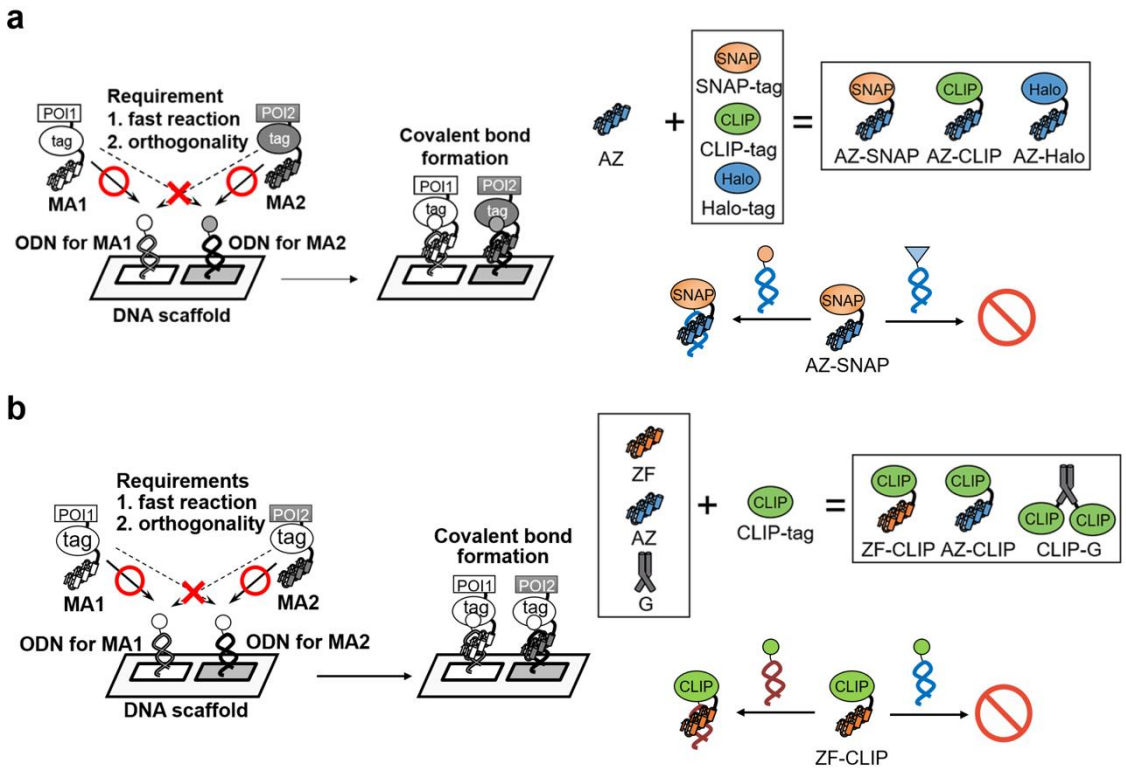


Figure 1.18. MA selectivity driven by chemoselectivity and sequence-selectivity. (a) Fast and orthogonal loading of POIs fused MAs are chemoselectively localized at the defined DNA addresses on the DNA scaffold.[96] DNA binding domain AZ were fused with SNAP-tag, CLIP-tag and Halo-tag, respectively. Orthogonal cross-linking reactions of MA were mainly driven by the chemo-selectivity of the protein tag. (b) MA guided POI to be sequence-selectively conjugated at the defined DNA addresses on the DNA scaffold. CLIP-tag were fused with DNA binding domain ZF, AZ and GCN4, respectively.[97] Orthogonal cross-linking reactions were mainly conducted by the affinity of DNA binding domain to its matched and unmatched DNA binding sequence.

1.4 Aim of the research in this thesis

In this thesis, the conjugation reactions were investigated for its application on biomacromolecular assemblies, such as DNA scaffold and polysaccharides. To study the kinetic influence on the selectivity of recognition-driven conjugation, I used modular adaptor comprising SNAP-tag (MA-SNAP) as a reaction model to explore the boundary condition for triggering sequence-specific reaction. Owing to the overly high reaction rate between SNAP-tag and benzylguanine (BG), the reaction between MA-SNAP and ODN modified by BG, failed to proceed in the sequence selective manner. By designing and synthesizing BG derivatives, benzyl inosine (BI) and 7-deaza-benzyl guanine (deBG), as new substrates of SNAP-tag, I tuned the reactivity of the substrate to SNAP-tag to be kinetically favorable for recognition-driven conjugation. As a result, it was found that BI enabled MA-SNAP to specifically react with BI-modified matched DNA sequence.

At the same time, BI showed remarkable specificity against CLIP-tag. In this case, chemoselectivity between SNAP-tag and CLIP-tag allowed orthogonal reaction to be directed between MA-SNAP and modular adaptor comprising CLIP-tag (MA-CLIP). With expanded diversity of MA, a series of MA-SNAP was sequence-selectively localized on the DNA scaffold. Combination of MA-SMAP and MA-CLIP shows potential to orthogonally locate four different types of POIs on DNA scaffold for more complicated design of enzyme cascade. These results showed the importance of studies on the recognition-driven DNA-protein conjugation to develop a method to orthogonally load biomolecules of interest on the DNA scaffold.

For another application of the DNA scaffold, I synthesized two FRET-based sensors with DNA handle for orthogonally detecting cathepsin B and D (CtB and CtD). Both cathepsin B and D sensors with DNA handle were used simultaneously to detect the activity of CtB and CtD, respectively. Particularly, conjugation of DNA handle to the cathepsin sensors improved their response to respective target enzyme by increasing the

solubility of hydrophobic FRET-based cathepsin sensors. Conjugation of a hydrophilic unit to the hydrophobic sensor would provide a new design strategy to improve the response of sensors, especially for the bulky and hydrophobic sensors. The DNA handle enabled the assembly of CtB sensor on the DNA scaffold with other sensors for detection of multiple signals associated with the action of CtB.

The FRET-based cathepsin sensor and a ratiometric fluorescent pH probe were assembled on dextran as another application of biomacromolecular assemblies. CtB is an attractive target for tumor-specific prodrug design and CtD could indicate the existence of active CtB. In the acidic environment of lysosomes, procathepsin B undergoes autocatalytic activation, leading to the formation of active CtB. Given the fact that dextran is transported into the cell through the macropinocytosis mechanism, coupling CtB or CtD sensor on dextran with pH sensor would provide real time information on the activation of CtB inside the lysosome. Characterization of the response of each sensor indicated that the cathepsin sensor and the pH probe on dextran would be utilized as orthogonal reporters inside the cell.

1.5 References

(1) Agafontsev, A. M.; Ravi, A.; Shumilova, T. A.; Oshchepkov, A. S.; Kataev, E. A. Molecular Receptors for Recognition and Sensing of Nucleotides. *Chem. - A Eur. J.* **2019**, *25*, 2684–2694.

(2) Cervoni, G. E.; Cheng, J. J.; Stackhouse, K. A.; Heimburg-Molinaro, J.; Cummings, R. D. O-Glycan Recognition and Function in Mice and Human Cancers. *Biochem. J.* **2020**, *477*, 1541–1564.

(3) Holm, C. K.; Paludan, S. R.; Fitzgerald, K. A. DNA Recognition in Immunity and Disease. *Curr. Opin. Immunol.* **2013**, *25*, 13–18.

(4) Filipczak, N.; Pan, J.; Yalamarty, S. S. K.; Torchilin, V. P. Recent Advancements in Liposome Technology. *Adv. Drug Deliv. Rev.* **2020**, *156*, 4–22.

(5) Gao, M.; Yu, F.; Lv, C.; Choo, J.; Chen, L. Fluorescent Chemical Probes for Accurate Tumor Diagnosis and Targeting Therapy. *Chem. Soc. Rev.* **2017**, *46*, 2237–2271.

(6) Srinivasarao, M.; Low, P. S. Ligand-Targeted Drug Delivery. *Chem. Rev.* **2017**, *117*, 12133–12164.

(7) Das, K.; Gabrielli, L.; Prins, L. J. Chemically Fueled Self-Assembly in Biology and Chemistry. *Angew. Chemie - Int. Ed.* **2021**, *60*, 20120–20143.

(8) Fitzgerald, P. R.; Paegel, B. M. DNA-Encoded Chemistry: Drug Discovery from a Few Good Reactions. *Chem. Rev.* **2021**, *121*, 7155–7177.

(9) Kong, G.; Xiong, M.; Liu, L.; Hu, L.; Meng, H. M.; Ke, G.; Zhang, X. B.; Tan, W. DNA Origami-Based Protein Networks: From Basic Construction to Emerging Applications. *Chem. Soc. Rev.* **2021**, *50*, 1846–1873.

(10) Spicer, C. D.; Davis, B. G. Selective Chemical Protein Modification. *Nat. Commun.* **2014**, *5*.

(11) Palaniappan, K. K.; Bertozzi, C. R. Chemical Glycoproteomics. *Chem. Rev.* **2016**,

116, 14277–14306.

(12) Boutureira, O.; Bernardes, G. J. L. Advances in Chemical Protein Modification. *Chem. Rev.* **2015**, *115*, 2174–2195.

(13) Sornay, C.; Vaur, V.; Wagner, A.; Chaubet, G. An Overview of Chemo- and Site-Selectivity Aspects in the Chemical Conjugation of Proteins. *R. Soc. Open Sci*. **2022**, *9*, 2115633.

(14) Reddy, N. C.; Kumar, M.; Molla, R.; Rai, V. Chemical Methods for Modification of Proteins. *Org. Biomol. Chem.* **2020**, *18*, 4669–4691.

(15) Mortensen, M. R.; Skovsgaard, M. B.; Gothelf, K. V. Considerations on Probe Design for Affinity-Guided Protein Conjugation. *ChemBioChem* **2019**, *20*, 2711–2728.

(16) Von Witting, E.; Hober, S.; Kanje, S. Affinity-Based Methods for Site-Specific Conjugation of Antibodies. *Bioconjug. Chem.* **2021**, *32*, 1515–1524.

(17) Bird, R. E.; Lemmel, S. A.; Yu, X.; Zhou, Q. A. Bioorthogonal Chemistry and Its Applications. *Bioconjug. Chem.* **2021**, *32*, 2457–2479.

(18) Matsumoto, T.; Tanaka, T.; Kondo, A. Enzyme-Mediated Methodologies for Protein Modification and Bioconjugate Synthesis. *Biotechnol. J.* **2012**, *7*, 1137–1146.

(19) Gauthier, M. A; Klok, H. A. Polymer–Protein Conjugates: an Enzymatic Activity Perspective. *Polym. Chem.*, **2010**,*1*, 1352–1373

(20) Nakano, S.; Fukuda, M.; Tamura, T.; Sakaguchi, R.; Nakata, E.; Morii, T. Simultaneous Detection of ATP and GTP by Covalently Linked Fluorescent Ribonucleopeptide Sensors. *J. Am. Chem. Soc.* **2013**, *135*, 3465–3473.

(21) Hagihara, M.; Hasegawa, T.; Sato, S; Yoshikawa, S.; Ohkubo, K.; Morii, T. Ribonucleopeptides: Functional RNA–Peptide Complexes. *Pept. Sci.* **2004**, *76*, 66–68

(22) Nakano, S.; Konishi, H.; Morii, T. Receptor-Based Fluorescent Sensors Constructed from Ribonucleopeptide, *Methods in Enzymology*, **2020**, *641*, 183–223.

(23) Nakano, S.; Tamura, T.; Das, R. K.; Nakata, E.; Chang, Y. T.; Morii, T. A Diversity-Oriented Library of Fluorophore-Modified Receptors Constructed from a Chemical Library of Synthetic Fluorophores. *ChemBioChem* **2017**, *18*, 2212–2216.

(24) Tamura, T.; Nakano, S.; Nakata, E.; Morii, T. Construction of a Library of Structurally Diverse Ribonucleopeptides with Catalytic Groups. *Bioorganic Med. Chem.* **2017**, *25*, 1881–1888.

(25) Formation, C. B.; Nakano, S.; Seko, T.; Zhang, Z.; Morii, T. RNA-Peptide Conjugation through an Efficient Covalent Bond Formation. *Appl. Sci.* **2020**, *10*, 8920.

(26) Spicer, C. D.; Pashuck, E. T.; Stevens, M. M. Achieving Controlled Biomolecule–Biomaterial Conjugation. *Chem. Rev.* **2018**, *118*, 7702–7743

(27) Peng, S.; Derrien, T. L.; Cui, J.; Xu, C.; Luo, D. From Cells to DNA Materials. *Mater. Today* **2012**, *15*, 190–194.

(28) Barua, E.; Deoghare, A. B.; Deb, P.; Lala, S. Das. Naturally Derived Biomaterials for Development of Composite Bone Scaffold: A Review. *IOP Conf. Ser. Mater. Sci. Eng.* **2018**, *377,* 12013.

(29) Tathe, A.; Ghodke, M.; Nikalje, A. P. A Brief Review: Biomaterials and Their Apllication. *Int. J. Pharm. Pharm. Sci.* **2010**, *2*, 19–23.

(30) Khan, F.; Tanaka, M. Designing Smart Biomaterials for Tissue Engineering. *Int. J. Mol. Sci.* **2018**, *19*, 1–14.

(31) Li, X.; Lv, H. F.; Zhao, R.; Ying, M. F.; Samuriwo, A. T.; Zhao, Y. Z. Recent Developments in Bio-Scaffold Materials as Delivery Strategies for Therapeutics for Endometrium Regeneration. *Mater. Today Bio* **2021**, *11*, 100101.

(32) Mohan, T.; Kleinschek, K. S.; Kargl, R. Polysaccharide Peptide Conjugates: Chemistry, Properties and Applications. *Carbohydr. Polym.* **2022**, *280*, 118875.

(33) Grossi, G.; Dalgaard Ebbesen Jepsen, M.; Kjems, J.; Andersen, E. S. Control of Enzyme Reactions by a Reconfigurable DNA Nanovault. *Nat. Commun.* **2017**, *8*, 1–8.

(34) Buday, L.; Tompa, P. Functional Classification of Scaffold Proteins and Related Molecules. *FEBS J.* **2010**, *277*, 4348–4355.

(35) Lemmens, L. J. M.; Ottmann, C.; Brunsveld, L. Conjugated Protein Domains as Engineered Scaffold Proteins. *Bioconjug. Chem.* **2020**, *31*, 1596–1603.

(36) Ryu, J.; Park, S. H. Simple Synthetic Protein Scaffolds Can Create Adjustable Artificial MAPK Circuits in Yeast and Mammalian Cells. *Sci. Signal.* **2015**, *8*, 1–11.

(37) Dueber, J. E.; Wu, G. C.; Malmirchegini, G. R.; Moon, T. S.; Petzold, C. J.; Ullal, A. V.; Prather, K. L. J.; Keasling, J. D. Synthetic Protein Scaffolds Provide Modular Control over Metabolic Flux. *Nat. Biotechnol.* **2009**, *27*, 753–759.

(38) Mulyasasmita, W.; Lee, J. S.; Heilshorn, S. C.; Molecular-Level Engineering of Protein Physical Hydrogels for Predictive Sol–Gel Phase Behavior. *Biomacromolecules* **2011**, *12*, 3406–3411

(39) Endo, M.; Yang, Y.; Sugiyama, H. DNA Origami Technology for Biomaterials Applications. *Biomater. Sci.* **2013**, *1*, 347–360.

(40) McCluskey, J. B.;Clark, D. S.; Glover, D. J. Functional Applications of Nucleic Acid–Protein Hybrid Nanostructures. *Trends Biotechnol.* **2020**, *38,* 976–989.

(41) Li, S.; Jiang, Q.; Liu, S.; Zhang, Y.; Tian, Y.; Song, C.; Wang, J.; Zou, Y.; Anderson, G. J.; Han, J. Y.; et al. A DNA Nanorobot Functions as a Cancer Therapeutic in Response to a Molecular Trigger in Vivo. *Nat. Biotechnol.* **2018**, *36*, 258–264.

(42) Yang, Y.; Wang, J.; Shigematsu, H.; Xu, W.; Shih, W. M.; Rothman, J. E.; Lin, C.

Self-Assembly of Size-Controlled Liposomes on DNA Nanotemplates. *Nat. Chem.* **2016**, *8*, 476–483.

(43) Nakata, E.; Hirose, H.; Gerelbaatar, K.; Arafiles, J. V. V.; Zhang, Z.; Futaki, S.; Morii, T. A Facile Combinatorial Approach to Construct a Ratiometric Fluorescent Sensor: Application for the Real-Time Sensing of Cellular PH Changes. *Chem. Sci.* **2021**, *12*, 8231–8240.

(44) Ngo, T. A.; Nakata, E.; Saimura, M.; Morii, T. Spatially Organized Enzymes Drive Cofactor-Coupled Cascade Reactions. *J. Am. Chem. Soc.* **2016**, *138*, 3012–3021.

(45) Yang, Y. R.; Liu, Y.; Yan, H. DNA Nanostructures as Programmable Biomacromolecular assembliess. *Bioconjug. Chem.* **2015**, *26*, 1381–1395.

(46) Sood, A.; Gupta, A.; Agrawal, G. Recent Advances in Polysaccharides Based Biomaterials for Drug Delivery and Tissue Engineering Applications. *Carbohydr. Polym. Technol. Appl.* **2021**, *2*, 100067.

(47) Zhong, Y.; Goltsche, K.; Cheng, L.; Xie, F.; Meng, F.; Deng, C.; Zhong, Z.; Haag, R. Hyaluronic Acid-Shelled Acid-Activatable Paclitaxel Prodrug Micelles Effectively Target and Treat CD44-Overexpressing Human Breast Tumor Xenografts in Vivo. *Biomaterials* **2016**, *84*, 250–261.

(48) Lee, S.; Stubelius, A.; Hamelmann, N.; Tran, V.; Almutairi, A. Inflammation-Responsive Drug-Conjugated Dextran Nanoparticles Enhance Anti-Inflammatory Drug Efficacy. *ACS Appl. Mater. Interfaces* **2018**, *10*, 40378–40387.

(49) Sauraj; Vinay kumar; Kumar, B.; Priyadarshi, R.; Deeba, F.; Kulshreshtha, A.; Kumar, A.; Agrawal, G.; Gopinath, P.; Negi, Y. S. Redox Responsive Xylan-SS-Curcumin Prodrug Nanoparticles for Dual Drug Delivery in Cancer Therapy. *Mater. Sci. Eng. C* **2020**, *107*, 110356.

(50) Carrico, I. S. Chemoselective Modification of Proteins: Hitting the Target. *Chem. Soc. Rev.* **2008**, *37*, 1423–1431.

(51) Foley, T. L.; Burkart, M. D. Site-Specific Protein Modification: Advances and Applications. *Curr. Opin. Chem. Biol.* **2007**, *11*, 12–19.

(52) Kenry; Liu, B. Bio-Orthogonal Click Chemistry for In Vivo Bioimaging. *Trends Chem.* **2019**, *1*, 763–778.

(53) Winssinger, N. Bioorthogonal Chemistry. *Chimia (Aarau).* **2018**, *72*, 755.

(54) Nguyen, S. S.; Prescher, J. A. Developing Bioorthogonal Probes to Span a Spectrum of Reactivities. *Nat. Rev. Chem.* **2020**, *4*, 476–489.

(55) Dey, B.; Thukral, S.; Krishnan, S.; Chakrobarty, M.; Gupta, S.; Manghani, C.; Rani, V. DNA-Protein Interactions: Methods for Detection and Analysis. *Molecular and Cellular Biochemistry*. **2012**, *89*, 279–299.

(56) Tse, W. C.; Boger, D. L. Sequence-Selective DNA Recognition: Natural Products and Nature's Lessons. *Chem. Biol.* **2004**, *11*, 1607–1617.

(57) Persch, E.; Dumele, O.; Diederich, F. Molecular Recognition in Chemical and Biological Systems. *Angew. Chemie - Int. Ed.* **2015**, *54*, 3290–3327.

(58) Amaike, K.; Tamura, T.; Hamachi, I. Recognition-Driven Chemical Labeling of Endogenous Proteins in Multi-Molecular Crowding in Live Cells. *Chem. Commun.* **2017**, *53*, 11972–11983.

(59) Shiraiwa, K.; Cheng, R.; Nonaka, H.; Tamura, T.; Hamachi, I. Chemical Tools for Endogenous Protein Labeling and Profiling. *Cell Chem. Biol.* **2020**, *27*, 970–985.

(60) Mortensen, M. R.; Skovsgaard, M. B.; Okholm, A. H.; Scavenius, C.; Dupont, D. M.; Rosen, C. B.; Enghild, J. J.; Kjems, J.; Gothelf, K. V. Small-Molecule Probes for Affinity-Guided Introduction of Biocompatible Handles on Metal-Binding Proteins. *Bioconjug. Chem.* **2018**, *29*, 3016–3025.

(61) Sau, S.; Petrovici, A.; Alsaab, H. O.; Bhise, K.; Iyer, A. K. PDL-1 Antibody Drug Conjugate for Selective Chemo-Guided Immune Modulation of Cancer. *Cancers (Basel).* **2019**, *11*, 1–12.

(62) Steiner, M.; Neri, D. Antibody-Radionuclide Conjugates for Cancer Therapy: Historical Considerations and New Trends. *Clin. Cancer Res.* **2011**, *17*, 6406–6416.

(63) Li, G.; Liu, Y.; Liu, Y.; Chen, L.; Wu, S.; Liu, Y.; Li, X. Photoaffinity Labeling of Small-Molecule-Binding Proteins by DNA-Templated Chemistry. *Angew. Chemie* **2013**, *125*, 9723–9728.

(64) Madsen, M.; Gothelf, K. V. Chemistries for DNA Nanotechnology. *Chem. Rev.* **2019,** *119*, 6384–6458.

(65) Sarkar, B.; Jayaraman, N. Glycoconjugations of Biomolecules by Chemical Methods. *Front. Chem.* **2020**, *8*, 1–27.

(66) Hong, F.; Zhang, F.; Liu, Y.; Yan, H. DNA Origami: Scaffolds for Creating Higher Order Structures *Chem. Rev.* **2017**, *117*, 12584–12640

(67) Tomasik, P. Chemical Modifications of Polysaccharides. *Chem. Funct. Prop. Food Saccharides* **2003**, *2013*, 123–130.

(68) Li, Q.; Li, S.; He, S.; Chen, W.; Cheng, P.; Zhang, Y.; Miao, Q.; Pu, K. An Activatable Polymeric Reporter for Near-Infrared Fluorescent and Photoacoustic Imaging of Invasive Cancer. *Angew. Chemie - Int. Ed.* **2020**, *59*, 7018–7023.

(69) Fiala, T.; Wang, J.; Dunn, M.; Šebej, P.; Choi, S. J.; Nwadibia, E. C.; Fialova, E.; Martinez, D. M.; Cheetham, C. E.; Fogle, K. J.; et al. Chemical Targeting of Voltage Sensitive Dyes to Specific Cells and Molecules in the Brain. *J. Am. Chem. Soc.* **2020**, *142*, 9285–9301.

(70) Melgar-Asensio, I.; Kandela, I.; Aird, F.; Darjatmoko, S. R.; de los Rios, C.; Sorenson, C. M.; Albert, D. M.; Sheibani, N.; Henkin, J. Extended Intravitreal Rabbit Eye Residence of Nanoparticles Conjugated with Cationic Arginine Peptides for Intraocular Drug Delivery: In Vivo Imaging. *Investig. Ophthalmol. Vis. Sci.* **2018**, *59*, 4071–4081.

(71) Ryu, J. H.; Kim, S. A.; Koo, H.; Yhee, J. Y.; Lee, A.; Na, J. H.; Youn, I.; Choi, K.;

Kwon, I. C.; Kim, B. S.; et al. Cathepsin B-Sensitive Nanoprobe for in Vivo Tumor Diagnosis. *J. Mater. Chem.* **2011**, *21*, 17631–17634.

(72) Nangreave, J.; Han, D.; Liu, Y.; Yan, H. DNA Origami: A History and Current Perspective. *Curr. Opin. Chem. Biol.* **2010**, *14*, 608–615.

(73) Endo, M.; Yang, Y.; Sugiyama, H. DNA Origami Technology for Biomaterials Applications. *Biomater. Sci.* **2013**, *1*, 347–360.

(74) Bandy, T. J.; Brewer, A.; Burns, J. R.; Marth, G.; Nguyen, T. N.; Stulz, E. DNA as Supramolecular Scaffold for Functional Molecules: Progress in DNA Nanotechnology. *Chem. Soc. Rev.* **2011**, *40*, 138–148.

(75) Kosuri, S.; Church, G. M. Large-Scale de Novo DNA Synthesis: Technologies and Applications. *Nat. Methods* **2014**, *11*, 499–507.

(76) Niemeyer, C. M. Semisynthetic DNA–Protein Conjugates for Biosensing and Nanofabrication. *Angew. Chem. Int. Ed.* **2010**, *49*, 1200 – 1216

(77) Trads, J. B.; Tørring, T.; Gothelf, K. V. Site-Selective Conjugation of Native Proteins with DNA. *Acc. Chem. Res.* **2017**, *50*, 1367–1374.

(78) Czogalla, A.; Kauert, D. J.; Seidel, R.; Schwille, P.; Petrov, E. P. DNA Origami Nanoneedles on Freestanding Lipid Membranes as a Tool to Observe Isotropic-Nematic Transition in Two Dimensions. *Nano Letters*. **2015**, *15*, 649–655.

(79) Wei, X.; Nangreave, J.; Jiang, S.; Yan, H.; Liu, Y. Mapping the Thermal Behavior of DNA Origami Nanostructures. *J. Am. Chem. Soc.* **2013**, *135*, 6165–6176.

(80) Fu, J.; Liu, M.; Liu, Y.; Woodbury, N. W.; Yan, H. Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *J. Am. Chem. Soc.* **2012**, *134*, 5516–5519.

(81) Wilner, O. I.; Weizmann, Y.; Gill, R.; Lioubashevski, O.; Freeman, R.; Willner, I. Enzyme Cascades Activated on Topologically Programmed DNA Scaffolds. *Nat. Nanotechnol.* **2009**, *4*, 249–254.

(82) Rosen, C. B.; Kodal, A. L. B.; Nielsen, J. S.; Schaffert, D. H.; Scavenius, C.; Okholm, A. H.; Voigt, N. V.; Enghild, J. J.; Kjems, J.; Tørring, T.; et al. Template-Directed Covalent Conjugation of DNA to Native Antibodies, Transferrin and Other Metal-Binding Proteins. *Nat. Chem.* **2014**, *6*, 804–809.

(83) Fu, J.; Yang, Y. R.; Johnson-Buck, A.; Liu, M.; Liu, Y.; Walter, N. G.; Woodbury, N. W.; Yan, H. Multi-Enzyme Complexes on DNA Scaffolds Capable of Substrate Channelling with an Artificial Swinging Arm. *Nat. Nanotechnol.* **2014**, *9*, 531–536.

(84) Linko, V.; Eerikäinen, M.; Kostiainen, M. A. A Modular DNA Origami-Based Enzyme Cascade Nanoreactor. *Chem. Commun.* **2015**, *51*, 5351–5354.

(85) Helmig, S.; Rotaru, A.; Arian, D.; Kovbasyuk, L.; Arnbjerg, J.; Ogilby, P. R.; Kjems, J.; Mokhir, A.; Besenbacher, F.; Gothelf, K. V. Single Molecule Atomic Force Microscopy Studies of Photosensitized Singlet Oxygen Behavior on a DNA Origami Template. *ACS Nano* **2010**, *4*, 7475–7480.

(86) Voigt, N. V.; Tørring, T.; Rotaru, A.; Jacobsen, M. F.; Ravnsbæk, J. B.; Subramani, R.; Mamdouh, W.; Kjems, J.; Mokhir, A.; Besenbacher, F.; et al. Single-Molecule Chemical Reactions on DNA Origami. *Nat. Nanotechnol.* **2010**, *5*, 200–203.

(87) Saccà, B.; Meyer, R.; Erkelenz, M.; Kiko, K.; Arndt, A.; Schroeder, H.; Rabe, K. S.; Niemeyer, C. M. Orthogonal Protein Decoration of DNA Origami. *Angew. Chemie - Int. Ed.* **2010**, *49*, 9378–9383.

(88) Ellis, G. A.; Klein, W. P.; Lasarte-Aragonés, G.; Thakur, M.; Walper, S. A.; Medintz, I. L. Artificial Multienzyme Scaffolds: Pursuing in Vitro Substrate Channeling with an Overview of Current Progress. *ACS Catal.* **2019**, *9*, 10812–10869.

(89) Rajendran, A.; Nakata, E.; Nakano, S.; Morii, T. Nucleic-Acid-Templated Enzyme Cascades. *ChemBioChem* **2017**, *18*, 696–716.

(90) Jiang, Q.; Liu, S.; Liu, J.; Wang, Z. G.; Ding, B. Rationally Designed DNA-Origami Nanomaterials for Drug Delivery In Vivo. *Adv. Mater.* **2019**, *31*, 1804785.

(91) Fu, J.; Yang, Y. R.; Dhakal, S.; Zhao, Z.; Liu, M.; Zhang, T.; Walter, N. G.; Yan, H. Assembly of Multienzyme Complexes on DNA Nanostructures. *Nat. Protoc.* **2016**, *11*, 2243–2273.

(92) Cui, C.; Zhang, H.; Wang, R.; Cansiz, S.; Pan, X.; Wan, S.; Hou, W.; Li, L.; Chen, M.; Liu, Y.; et al. Recognition-Then-Reaction Enables Site-Selective Bioconjugation to Proteins on Live-Cell Surfaces. *Angew. Chemie - Int. Ed.* **2017**, *56*, 11954–11957.

(93) Nielsen, T. B.; Thomsen, R. P.; Mortensen, M. R.; Kjems, J.; Nielsen, P. F.; Nielsen, T. E.; Kodal, A. L. B.; Cló, E.; Gothelf, K. V. Peptide-Directed DNA-Templated Protein Labelling for The Assembly of a Pseudo-IgM. *Angew. Chemie - Int. Ed.* **2019**, *58*, 9068–9072.

(94) Nakata, E.; Dinh, H.; Ngo, T. A.; Saimura, M.; Morii, T. A Modular Zinc Finger Adaptor Accelerates the Covalent Linkage of Proteins at Specific Locations on DNA Nanoscaffolds. *Chem. Commun.* **2015**, *51*, 1016–1019.

(95) Dinh, H.; Nakata, E.; Mutsuda-Zapater, K.; Saimura, M.; Kinoshita, M.; Morii, T. Enhanced Enzymatic Activity Exerted by a Packed Assembly of a Single Type of Enzyme. *Chem. Sci.* **2020**, *11*, 9088–9100.

(96) Nguyen, T. M.; Nakata, E.; Saimura, M.; Dinh, H.; Morii, T. Design of Modular Protein Tags for Orthogonal Covalent Bond Formation at Specific DNA Sequences. *J. Am. Chem. Soc.* **2017**, *139*, 8487–8496.

(97) Nguyen, T. M.; Nakata, E.; Zhang, Z.; Saimura, M.; Dinh, H.; Morii, T. Rational Design of a DNA Sequence-Specific Modular Protein Tag by Tuning the Alkylation Kinetics. *Chem. Sci.* **2019**, *10*, 9315–9325.

CHAPTER 2

Tuning the reactivity of protein tag to realize sequence-specificity of modular adaptor

2.1 Introduction

As an ideal reaction model for studying recognition-driven conjugation, modular adaptors (MAs) constitute both a sequence-specific DNA-binding domain (DBD) and a self-ligating protein-tag.[1–3] This approach couples the molecular recognition event (via DBD) to the chemoselective reaction between the protein tag and substrate on DNA to realize a far greater reactivity and selectivity than that exclusively rely on the chemoselectivity alone. As our previous discussion, crosslink formation between MA and substrate-modified DNA is a two-step process that DNA-protein complex is reversibly formed through the sequence-specific binding of DBD and next the self-ligation reaction of protein tag with its substrate is activated by the proximity (Figure 2.1a).[4–6] So it was speculated that the apparent rate constant of the MA reaction with substrate-modified DNA (k_{app}) (M^{-1} s^{-1}) depends on the parameters which are related with both the binding affinity of DNA-protein binding process and the reactivity of protein tag reaction, in the case of $k_{cov} \ll k_{off}$ (Figure 2.1c). And hence equilibrium dissociation constant (K_D) can result in different k_{app} values between DNA binding protein and different DNA binding sequences, in other words matched pair and unmatched pair based on DNA-protein specific binding. Sequence-selective recognition in the case of $k_{cov} \ll k_{off}$ was driven by different k_{app} values between matched pair and unmatched pair. But when MA consisting of SNAP-tag[7] (MA-SNAP) reacted with O^2-benzylguanine (BG) -modified DNA, overly high reaction rate of SNAP-tag with BG ($k_{cov} \gg k_{off}$) propelled k_{app} to be irrelevant with K_D (Figure 2.1b).[5] It revealed that the excessively high reactivity of SNAP-tag stymied the sequence-selectivity of MA-SNAP. Therefore, the kinetics speculation prompted us to tune the reactivity of SNAP-tag to fit it back into the case, $k_{cov} \ll k_{off}$, in order to attain the

sequence-selectivity of MA-SNAP.

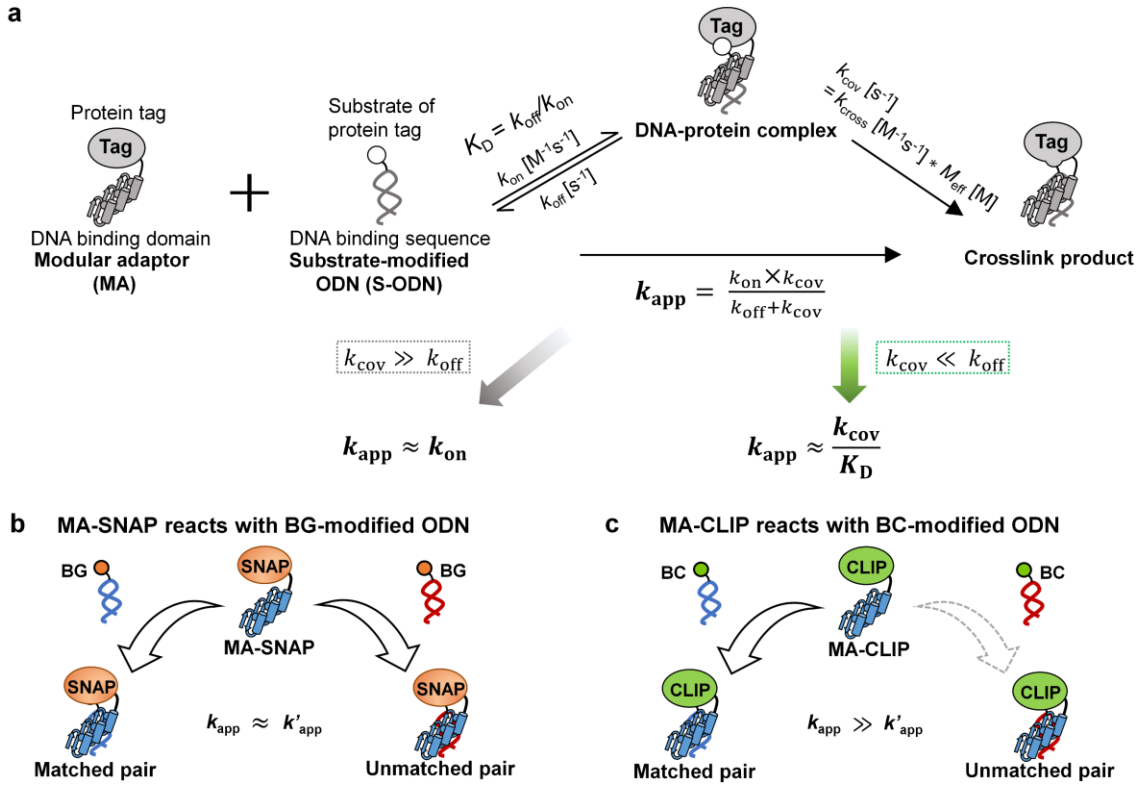
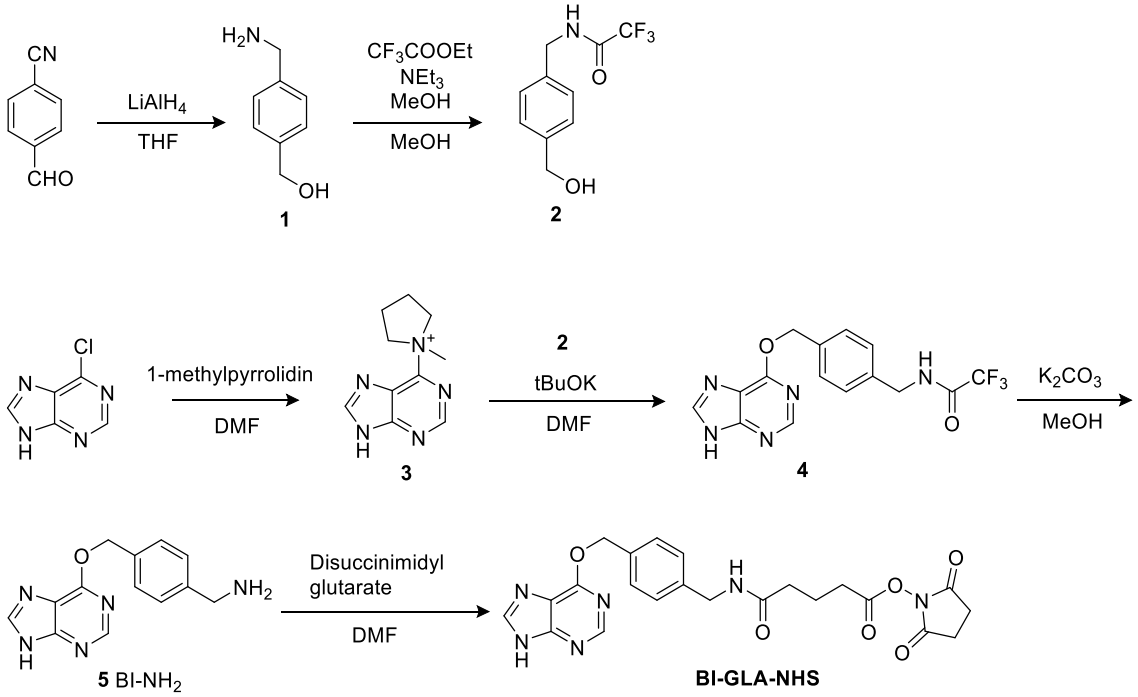


Figure 2.1 Reaction scheme of the DNA-protein conjugation mediated by modular adaptors.[4–6] (a) Schematic representation of the DNA recognition and cross-linking reaction between MA and target DNA. The complex formation between MA and DNA is governed by the equilibrium dissociation constant (K_D) (M) defined by $K_D = k_{off}/k_{on}$ and the kinetics of overall cross-linking reaction is defined by the apparent rate constant (k_{app}) (M^{-1} s^{-1}). The rate constant of covalent bond formation in close proximity (k_{cov}) (s^{-1}) is defined by the product of rate constant of intermolecular covalent formation (k_{cross}) (M^{-1} s^{-1}) and the effective concentration (M_{eff}) (M). (b) An illustration of the cross-linking of ZF-SNAP to substrate-modified oligodeoxynucleotides (ODN) the matched pair (ODN-ZF-BG) and the unmatched pair (ODN-AZ-BG). (c) An illustration of the cross-linking of ZF-CLIP to the matched pair (ODN-ZF-BC) and the unmatched pair (ODN-AZ-BC). In both (b) and (c), the apparent reaction rate constants for the matched and unmatched pairs are indicated as k_{app} and k'_{app}, respectively.

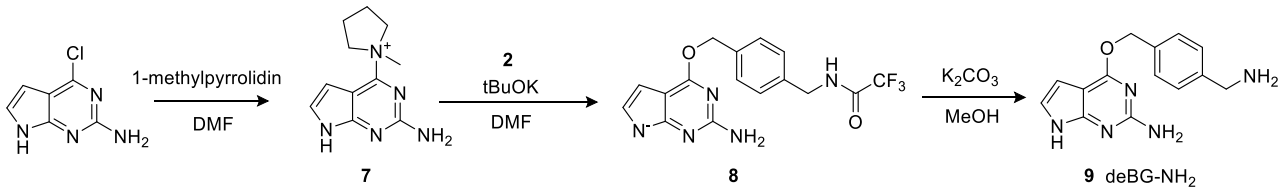
2.2 Results

2.2.1 Structure-based design and synthesis of SNAP-tag substrates

For tuning the reactivity of SNAP-tag, the modification of the substrate of SNAP-tag, BG, was aimed to be slowed down its reaction rate by reducing the binding affinity between SNAP-tag and its substrate.[7] The crystal structure of SNAP-tag with BG implied the key features of the substrate binding: the amino group at C2 and the N7 of BG played important roles in positioning BG for efficient alkyl transfer by forming hydrogen bonds interactions with the residues of SNAP-tag (Figure 2.2a and 2.2b).[8] The structural facts suggested the feasibility to reduce the efficiency of alkyl transfer through impairing the substrate binding of SNAP-tag. For this reason, design and synthesis of two BG derivatives, benzylinosine (BI) without the amino group at C2 (Scheme 2.1, Figure 2.2c) and 7-deaza-benzylguanine (deBG) were initiated by replacing the N7 as C7 (Scheme 2.2 and Figure 2.2d).[9,10] By modifying BI with a linker and an N-hydroxysuccinimide group (NHS) on the methylbenzene, it would facilitate BI to be modified on the DNA which was pre-modified with amino group (Scheme 2.1). NHS-tethered BI, BI-GLA-NHS, was characterized by ^{1}H and ^{13}C NMR (Figure 2.3).



Scheme 2.1. Reaction scheme of the synthesis of BI-GLA-NHS



Scheme 2.2 Reaction scheme of the synthesis of O^6-(4-aminomethyl-benzyl) 7-deazaguanine (deBG-NH_2)

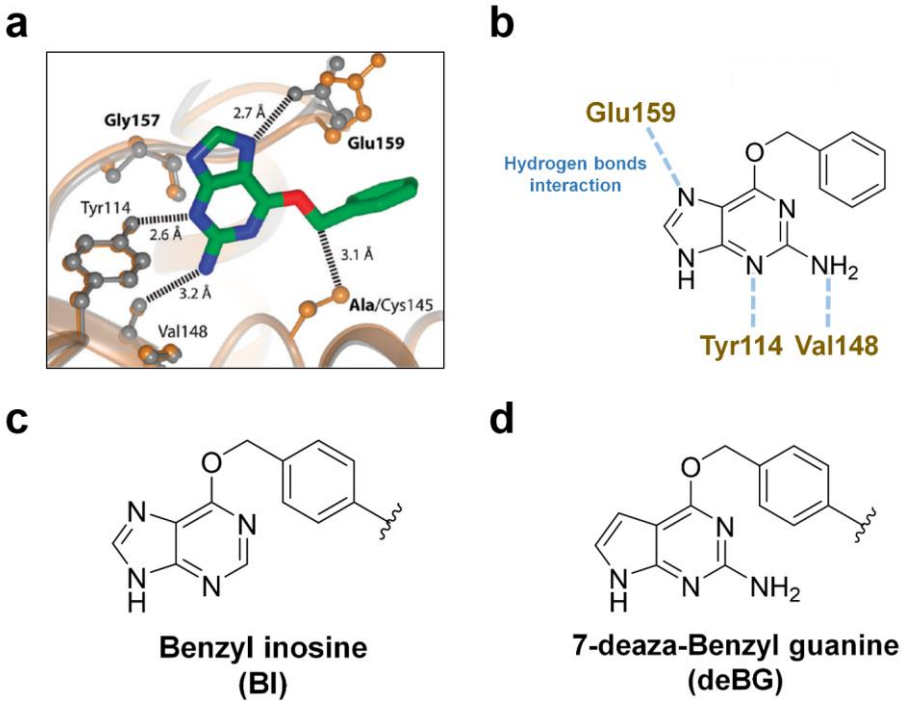


Figure 2.2 (a) Crystal structure of BG-binding pocket of SNAP-tag.[8–10] (b) Amino acid residues of SNAP-tag that directly interact with BG through hydrogen bonding interactions (PDB entry 1EH6). (c) Chemical structure of benzylinosine (BI). (d) Chemical structure of 7-deaza-benzylguanine (deBG).

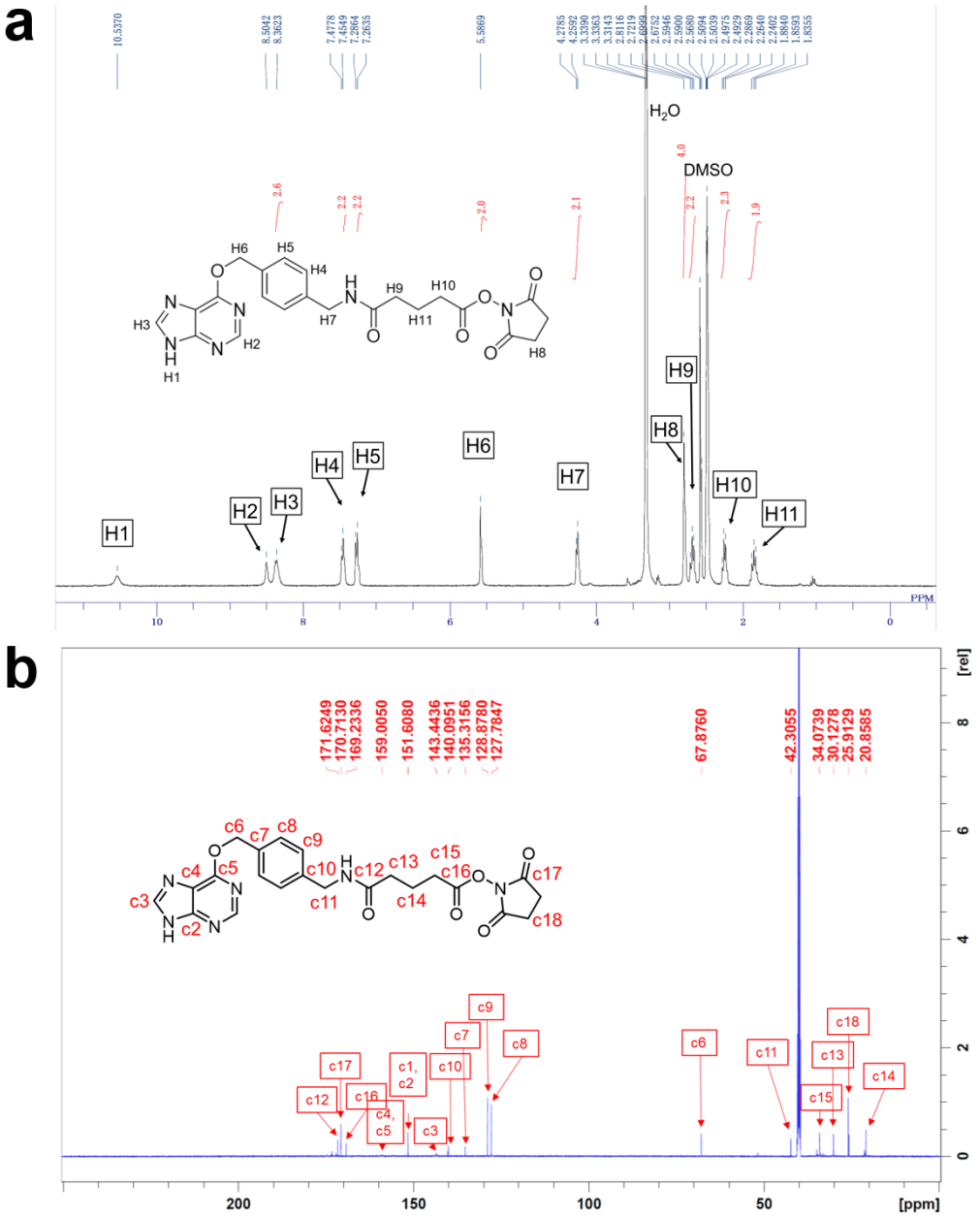


Figure 2.3 NMR spectra of BI-GLA-NHS: (a) ^{1}H NMR of BI-GLA-NHS, (b) ^{13}C NMR of BI-GLA-NHS. These peaks were identified by measuring $^{1}H/^{13}C$ heteronuclear single quantum coherence (HSQC) and $^{1}H/^{13}C$ heteronuclear multiple quantum coherence (HMBC).

2.2.2 The reactivity of protein tag with their substrates

In order to investigate the reactivity of BI and deBG with SNAP-tag containing MA, BI and deBG were conjugated with a fluorophore ATTO 488 through the condensation reaction between the amino group on the 4-(aminomethyl)-methylbenzene of BI and the NHS group on ATTO 488. The resulted substrates are denoted as ATTO 488-BI and ATTO 488-deBG, respectively (Figure 2.4a and 2.4b). The ATTO 488-modified BI and deBG were characterized by HPLC (Figure 2.4c and 2.4d). The absorbance peak of ATTO 488 can be observed on 498 nm from the UV spectrum. (Figure 2.4e and 2.4f)

For comparison, the parent substrate BG and the substrate of CLIP-tag, benzylcytosine (BC), was also modified with the same fluorophore to produce ATTO-BG and ATTO-BC. Cross-linking reactions between these fluorophore-modified substrates and protein tags (AZ-SNAP and AZ-CLIP) were measured by fluorescence polarization (FP, Figure 2.5). The results demonstrated that rate constants of BI (12 M^{-1} s^{-1}) and deBG (26 M^{-1} s^{-1}) were ~125 times smaller than that of BG (3227 M^{-1} s^{-1}) for the reaction with SNAP-tag and were very similiar to that of BC with CLIP-tag (46 M^{-1} s^{-1}), which not only indicated their possible application for the recognition-driven reaction between DNA and SNAP-tag, but also ensured their sufficient reactivity. Importantly, the rate constants of newly designed BI and deBG to CLIP-tag were much smaller than that of the original substrate BC (Table 2.1). For CLIP-tag, a small increase in the FP was observed with deBG when compared to BI (Figure 2.5b), indicating the higher selectivity of BI to SNAP-tag against CLIP-tag. These results taken together, BI is concluded to be a suitable substrate for SNAP-tag for the recognition-driven reaction while at the same time retaining the high specificity against CLIP-tag.

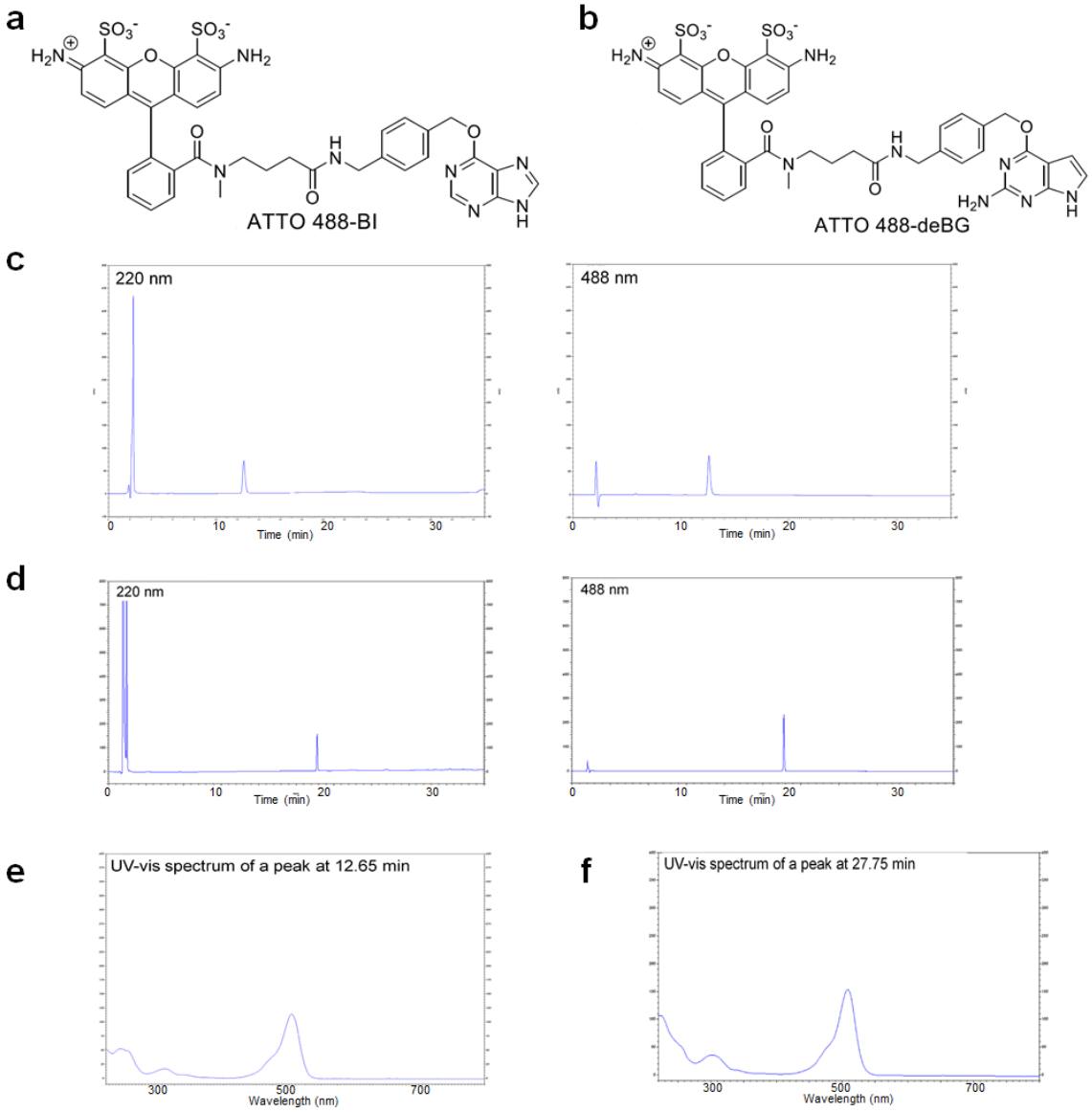


Figure 2.4 ATTO 488-modified BI and deBG. (a) Chemical structure of ATTO 488-BI. (b) Chemical structure of ATTO 488-deBG. (c) HPLC analysis of purified ATTO 488-BI. HPLC profiles were detected by UV at 220 nm and 488 nm. HPLC analysis was conducted on a ULTRON VX-ODS column (4.6 × 150 mm, eluent A: 0.05% TFA containing H2O, eluent B: 0.05% TFA in 80% acetonitrile). Sample was eluted with the gradient of eluent B increased from 20% to 50% in 35 min at a flow rate of 1.0 ml min^{-1}. (d) HPLC analysis of purified ATTO 488-deBG detected at 220 nm and 488 nm of UV. Sample was eluted with the gradient of eluent

B increased from 5% to 50% in 50 min at a flow rate of 1.0 ml min^{-1}. (e) UV-vis spectrum of a peak in (c) eluted at 12.65 min. (f) UV-vis spectrum of a peak in (d) eluted at 27.75 min.

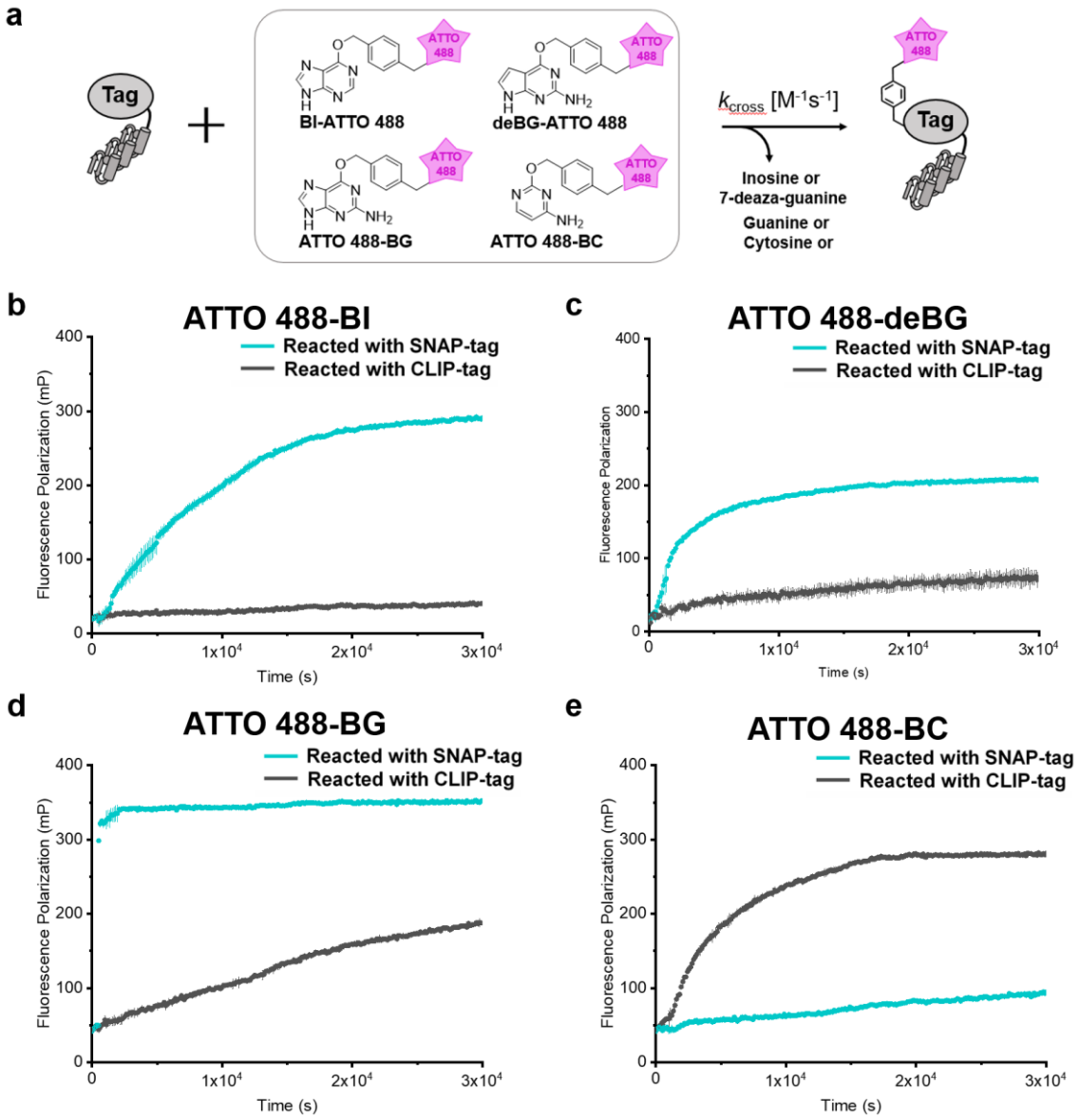


Figure 2.5 MA reacted with ATTO 488-labelled substrate. (b) A scheme illustrates covalent bond formation between protein tag derivatives with fluorophore modified substrates (ATTO 488-BI, ATTO 488-deBG, ATTO 488-BG or ATTO 488-BC). (c) The reaction time courses of SNAP-tag and CLIP-tag (10 μM) with (b) ATTO 488-BI, (c) ATTO 488-deBG, (d) ATTO 488-BG or (e) ATTO 488-BC (20 nM). The changes of FP were recorded every 60 s. The k_{cross} calculated from the FP curve are shown in Table 2.1.

Table 2.1 The rate constants of the cross-linking reaction (k_{cross})[a] (M^{-1} s^{-1}) between the fluorophore modified substrates and MAs.

Substrates	AZ-SNAP	AZ-CLIP
ATTO488-BI	12 ± 0.5	< 1
ATTO488-deBG	26 ± 1.0	< 1
ATTO488-BG	3227 ± 76	4 ± 0.1
ATTO488-BC	< 1	46 ± 1.2

[a] The k_{cross} values were estimated from the time-course analysis of reaction by fluorescence polarization (Figure 2.5).

2.2.3 Sequence-selective reaction performed by modular adaptor comprising SNAP-tag

As BI showed optimized reactivity to SNAP-tag and remarkable selectivity against CLIP-tag, it was used as the target substrate to investigate sequence-selectivity of MA-SNAP and potential chemoselectivity against MA comprising CLIP-tag (MA-CLIP). To examine the sequence-selective reaction of MA-SNAP, the developed SNAP-tag substrate BI was kept constant, while the sequence of the DNA was varied. The adapted oligodeoxynucleotides (ODNs) contained the target sequences for DNA-binding zinc finger proteins zif268 (ODN-ZF)[11] and AZP4 (ODN-AZ),[12] and the basic-leucine zipper class of protein GCN4 (ODN-AP).[13] As shown in Figure 2.6a, to modify ODN with the substrate, a unique amino-C6-T (T^R) was inserted into the T loop region of ODN-ZF and ODN-AZ, and two T^R to ODN-AP (because GCN4 is a dimeric protein). BI-modified ODNs named ODN-ZF-BI, ODN-AZ-BI and ODN-AP-2BI were prepared. With the modification of BI on DNA, the DNA-protein conjugation would be completed by the reaction of SNAP-tag with BI (Figure 2.6b). In the same manner, BG- and BC-modified ODNs (as shown in Figure 2.6a) were prepared. A series of

MA-SNAPs (ZF-SNAP, AZ-SNAP and SNAP-GCN4) were prepared by fusing a protein-tag with the respective DNA-binding protein through a GGSGGS linker (Figure 2.6c).[4] As the confrimation of the DNA-protein conjugation product, the cross-linking products between ZF-SNAP and ODN-ZF-BI were analyzed by LC-MS (Figure 2.7). By means of a liquid chromatography-mass spectrometry (LC-MS), it was observed that the molecular weight of the conjugation products (*m/z* = 45873 $[M]^+$, calcd.45873) was 13558 higher than that of ZF-SNAP (*m/z* = 32315 $[M]^+$, calcd. 32316), which was exactly the molecular weight of ODN-ZF-BI (13558, calcd. 13558).

Next, for checking the selectivity of MA-SNAP conducted by the modification of BI, the cross-linking reactions between BI-modified ODNs and all the combinations of MAs were performed for 30 min, and then the mixture was analyzed by denaturing polyacrylamide gel electrophoresis (PAGE). ^{32}P was labelled on the 5' end of BI-modified DNA for gel imaging in picomole scale. The slow migrating bands corresponding to the cross-linking products were observed only for the reactions of matched pairs of ODN-DBD (Figures 2.8a). Impressively, the reaction of matched pair still kept very high efficiency owing to the recognition-driven pricinple of MA, espeically 92% for ZF-SNAP reaction. In contrast, under the same experimental conditions, the reaction between MA-SNAP and its original substrate BG-modified ODN resulted cross-linking products with no sequence selectivity (Figures 2.8b). These results indicated that the newly developed substrate BI-modified ODN displayed an unprecedented higher degree of selectivity to MA-SNAPs when combined with all three DNA-binding proteins.

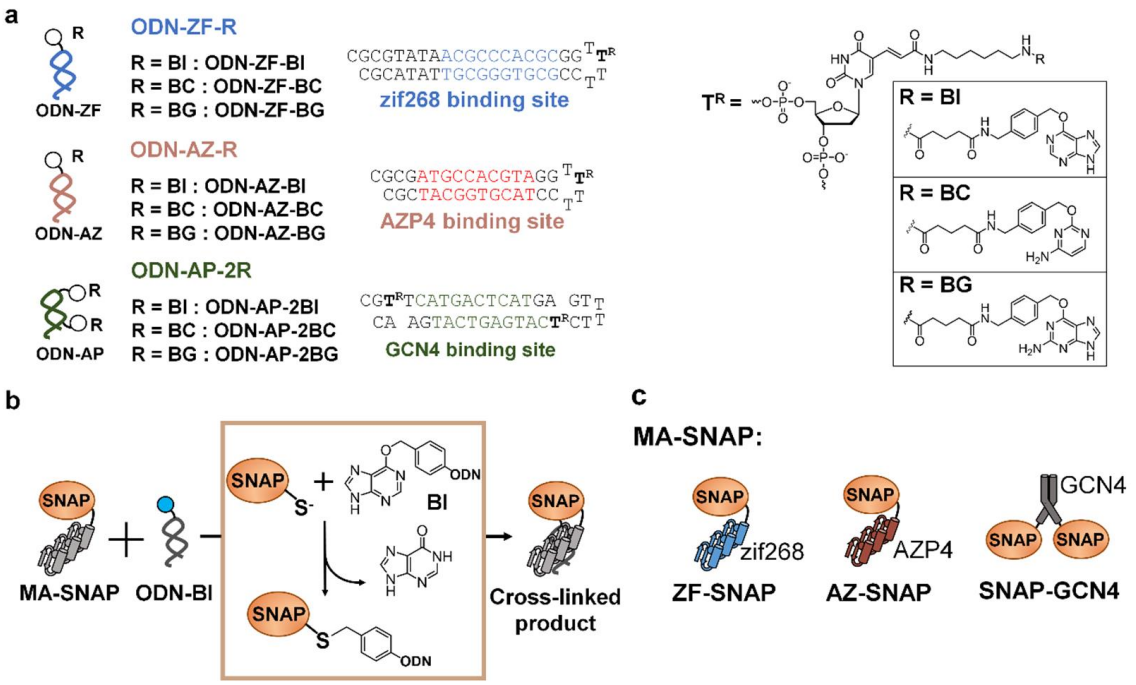


Figure 2.6 MA and substrate-modified ODN. a) Nucleotide sequences of ODNs modified with the substrates BI or BC and chemical structures of the substrates-modified T, denoted as "T^R" (right). (b) A scheme representing the cross-linking reaction between MA-SNAP and BI-modified ODN. (c) Illustration of MA-SNAPs with various DBD used in this study.

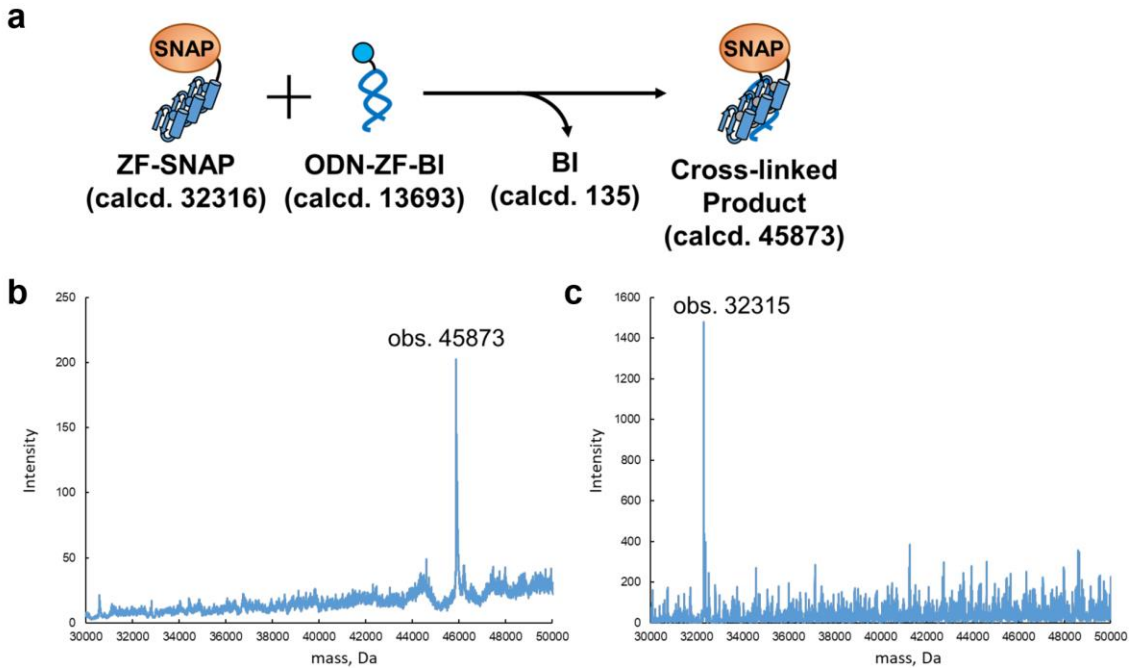


Figure 2.7 LC-MS analysis of the cross-linking product. (a) A scheme illustrates a covalent bond formation between ZF-SNAP with ODN-ZF-BI. (b-c) MS spectra of (b) cross-linked product (ZF-SNAP and ODN-ZF-BI) and (c) ZF-SNAP. The molecular weights of them were determined on a liquid chromatography-mass spectrometry (LC-MS) with Exion LC HPLC system (SCIEX, USA) equipped with YMC-Triart Bio C4 column (50 x 2.1 mm, S-5 μm, 30 nm, YMC, JAPAN) and QTOF X500R MS system (SCIEX, USA). The experiment was conducted as follow: The injection volume was 20 or 40 μL. The flow rate was 0.3 $mL \cdot min^{-1}$. The mobile phases consisted of (A) 0.1% formic acid in water and (B) 0.1% formic acid in acetonitrile.

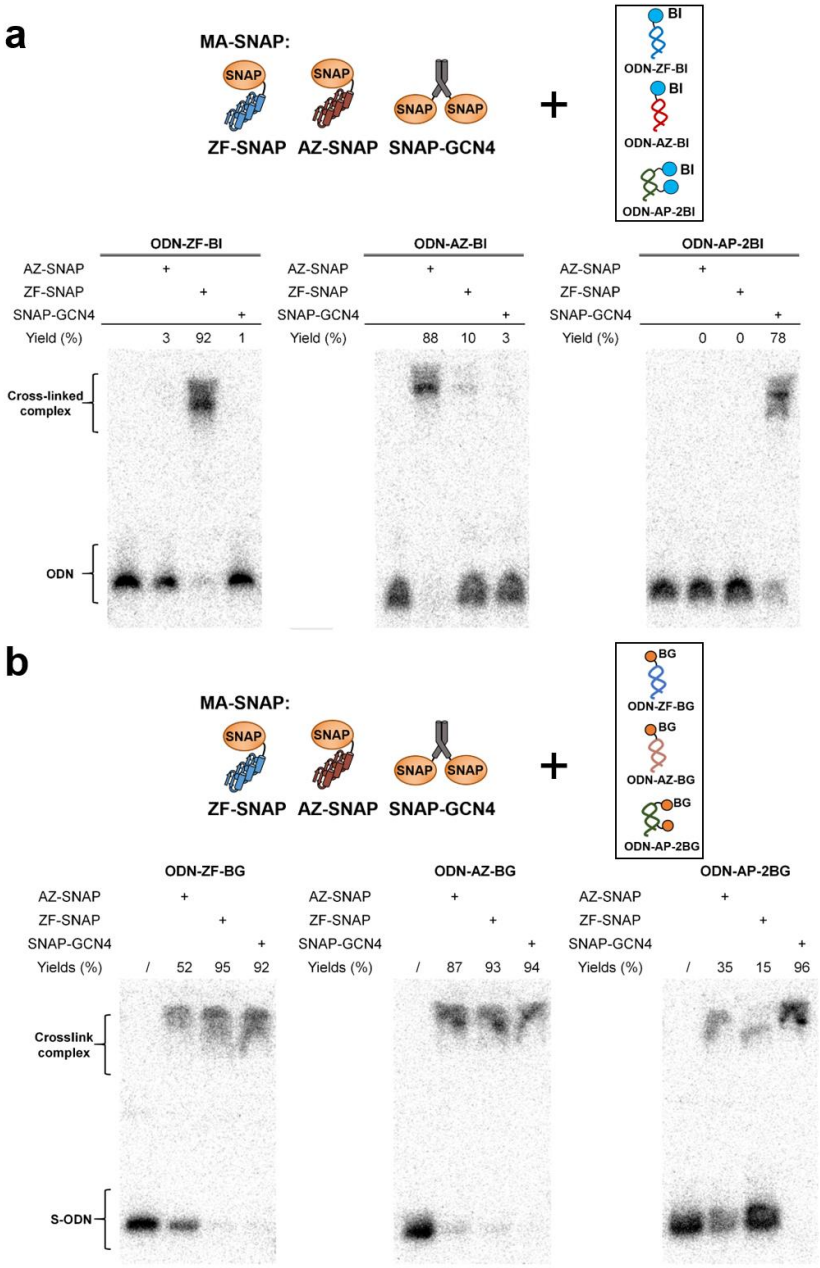


Figure 2.8 Cross-linking reaction of MA-SNAP with BI- or BG-modified DNA. (a) Autoradiograms show denaturing PAGE analyses of the cross-linking reactions between MA-SNAPs and BI-modified ODNs: ODN-ZF-BI, ODN-AZ-BI, and ODN-AP-2BI. Each 5′-^{32}P-end-labeled BI-modified ODN (0.5 nM) was incubated for 30 min with respective MA-SNAP (150 nM). (b) Denaturing PAGE analyses of the crosslinking reaction of MA-SNAP and BG-modified ODN (ODN-ZF-BG, ODN-AZ-BG, and ODN-AP-2BG). Each 5’-^{32}P-end-labeled BG-modified ODN (0.5 nM) were incubated for 30 min at room temperature with each MA-SNAP (50 nM), respectively.

2.2.4 Analysis of the kinetic parameters of the reaction with modular adaptor

For the sequence-selectivity of MA, it is speculated that if the kinetic parameter of the reaction, k_{cov} and k_{off}, can be quantitatively determined, the comparison of k_{cov} and k_{off} would directly conclude the answer of the selectivity of MA. Therefore, kinetic aspects of sequence-selective reaction were then studied by taking ODN-ZF-BI and ZF-SNAP pair as a model system. The second-order rate constant k_{app} was estimated from the linear relationship between the pseudo-first-order rate constant k_{obs} (s^{-1}) and the concentration of MA. Time course profiles for the reaction of ODN-ZF-BI with various concentrations of ZF-SNAP (5–40 nM) were monitored to estimate k_{obs} at each concentration (Figure 2.9a-d). And the values k_{obs} at each concentration were plotted in a dose-dependent manner to calculate k_{app} by linear curve-fitting. The k_{app} for this reaction was deduced to be 4.4×10^5 M^{-1} s^{-1} (Figure 2.9e).

With k_{app} value in hand, the complex formation by the same pair of MA and ODN was then investigated estimate k_{on} and k_{off}. For this purpose, the ODN-ZF was labelled with Alexa Fluor 488. The stopped-flow FP analyses of the non-covalent complex formation between the DBD (i.e, zif268) and ODN-ZF provided the k_{obs} at defined concentrations (Figures 2.10b and 2.10c). The k_{on} and k_{off} of the reversible complex formation were determined to be 8.9×10^6 M^{-1} s^{-1} and 0.64 s^{-1}, respectively (Table 2.2). In a parallel experiment, k_{off} was estimated to be 0.49 s^{-1} for the dissociation of complex by the addition of competitor DNA, which is in the same order of magnitude with the value obtained for the dissociation of original complex (0.64 s^{-1}) (Figure 2.10e and Table 2.2). The estimated K_D of 72 nM given by k_{off}/k_{on} is comparable with the previously reported value of 56 nM estimated by the steady-state FP titration.[4] From the above-mentioned kinetic parameters (k_{app}, k_{on} and k_{off}) and $k_{app} = (k_{on}\, k_{cov})/(k_{off} + k_{cov})$, the k_{cov} is estimated to be 0.032 s^{-1}. As a result, the k_{off} (0.64 s^{-1}) was one order

of magnitude higher than the k_{cov} (0.032 s^{-1}). Thus, the reactivity of BI to SNAP-tag was effectively reduced to meet the required criterion of $k_{cov} \ll k_{off}$ for conducting sequence-selective reaction.

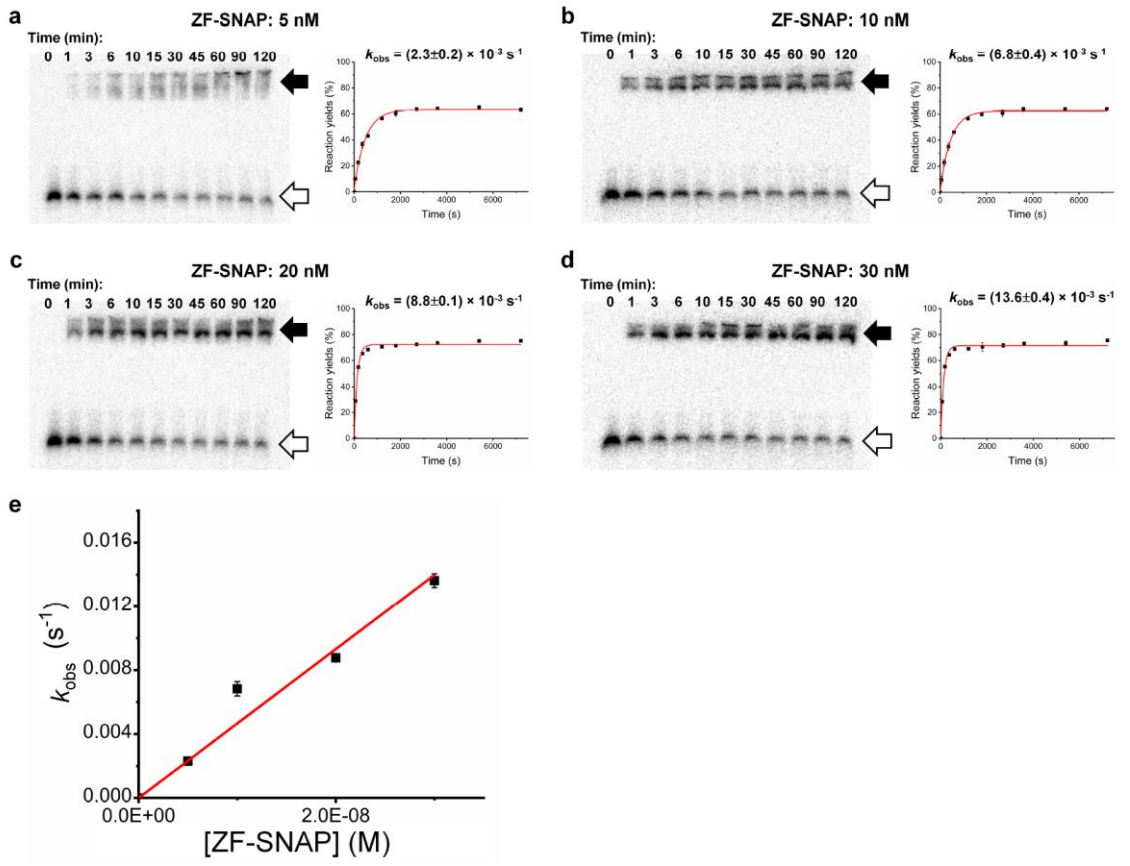


Figure 2.9 Autoradiogram show denaturing PAGE analyses of the samples from the reactions between 5'-^{32}P-end- labeled ODN-ZF-BI (0.5 nM) reacted with ZF-SNAP[(a) 5 nM, (b) 10 nM, (c) 20 nM and (d) 30 nM]. Open and filled arrows denote free ODN and MA-bound ODN, respectively. Time-course plots for the crosslinking reaction were fitted to a reaction model assuming first-order kinetics to obtain the second-order rate constants (k_{obs}). (e) The dependence of k_{obs} upon the concentration of ZF-SNAP. The second order rate constants k_{app} ($M^{-1}s^{-1}$) between ZF-SNAP and ODN-ZF-BI was determined as (4.4 ± 0.5) × 105 $M^{-1}s^{-1}$ by the linear-curve fitting.

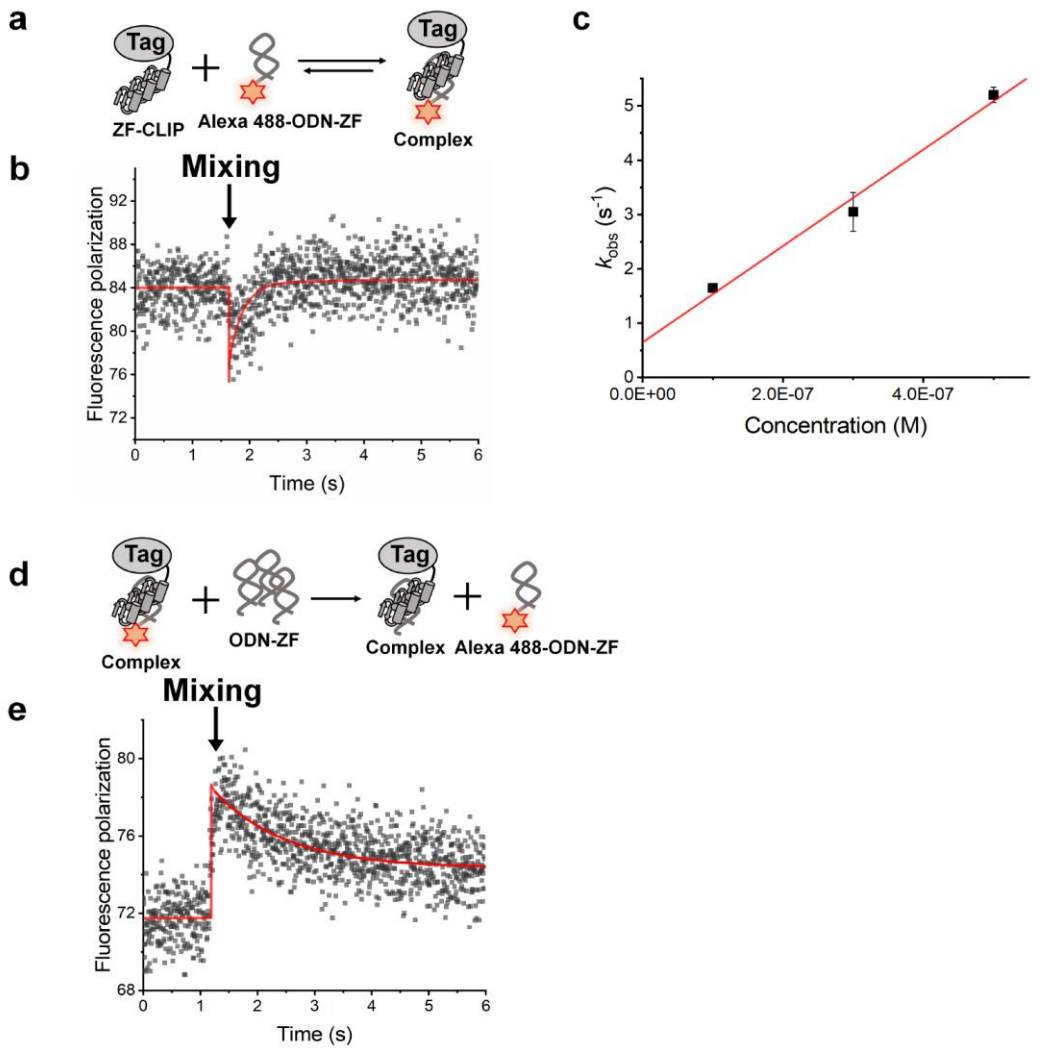


Figure 2.10 Stopped flow analyses of fluorescence polarization (FP) for the (a-c) association and (d, e) dissociation processes between zif268 included in MA and 5'-end Alexa Fluor 488-labeled ODN-ZF (Alexa 488-ODN-ZF). Determined kinetic parameters were shown in Table 2.2. (a) Schematic illustration of the reaction which takes place post mixing by stopped flow: Alexa 488-ODN-ZF binds to zif268 included in ZF-CLIP. (b) Real time FP change observed upon mixing ZF-CLIP with Alexa 488-ODN-ZF. The black arrow indicates the mixing process of ZF-CLIP (500 nM) solution and Alexa 488-ODN-ZF solution (5 nM). (c) Rate constants (k_{obs}, s^{-1}) determined for the reaction plotted against the concentration of ZF-CLIP. (d) Schematic illustration of the reaction which takes place post mixing by stopped flow: Alexa 488-ODN-ZF pre-bound with zif268 included in ZF-CLIP was mixed with an excess of non-labelled ODN-ZF. Dissociating Alexa 488-ODN-ZF was replaced by non-labelled ODN-ZF,

thereby preventing the back reaction of Alexa 488-ODN-ZF. Alexa 488-ODN-ZF (5 nM) was pre-bound with ZF-CLIP (500 nM) and the solution was mixed with non-labelled ODN-ZF (2.5 μM) solution. The k_{off} (s^{-1}) between zif268 and ODN-ZF was calculated by a single exponential function. (e) Real time FP change observed upon mixing Alexa 488-ODN-ZF pre-bound with ZF-CLIP and ODN-ZF.

Table 2.2 Binding parameters between zif268 and ODN-ZF determined by fluorescence polarization (FP) measurements.

Method for measurement	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	K_D (nM)
Time resolved FP (association process, Figure 2.10a-c)	8.6×10^6	0.64	72
Time resolved FP (dissociation process, Figure 2.10d-e)		0.49	
FP[4] (steady state)			56

2.2.5 Characterization of sequence-selectivity through the analyses of apparent reaction rate

Table 2.3 Nucleotide sequences of BI modified ODN-ZF derivatives.[14,15]

Oligo name	Sequence (from 5′ to 3′) T^{BI} = BI modified T
ODN-ZF-BI	CGCGTATAACGCCCACGCGCGTT^{BI}TTCGCGCGTGGGCGTTATACGC
ODN-ZF(G/T)-BI	CGCGTATAACGCC**A**ACGCGCGTT^{BI}TTCGCGCGT**T**GGCGTTATACGC
ODN-ZF(GC/CT)-BI	CGCGTATAAC**AG**CCACGCGCGTT^{BI}TTCGCGCGTGG**CT**GTTATACGC

As analysed by kinetic parameters, BI-modified DNA enabled MA-SNAP to react sequence-selectively, which could be kinetically described as $k_{app} = k_{cat}/K_D$. For the reaction performed by MA-SNAP, k_{cat} would be invariable. Hence, reaction rate k_{app} was supposed to be proportionally regulated by the binding affinity K_D between DBD and DNA binding sequence. The sequence-selective reaction was further investigated by using a consensus binding sequence of zif268 and its mutated sequences (GCGTGGGCGT, where the underlines indicate the mutated bases in the consensus

binding sequence of zif268, Table 2.3).[14,15] Compared to the wild-type ODN-ZF (56 nM), the K_D of G/T-mutant ODN-ZF(G/T) (1.1 × 10^3 nM) and GC/CT-mutant ODN-ZF(GC/CT) (> 10^4 nM) for zif268 was increased about one and two order of magnitude, respectively.[4] The k_{app} of ZF-SNAP with each BI-modified DNA was calculated by incubating DNA and MA in a time-course manner (Figure 2.11). The k_{app} of ZF-SNAP to G/T-mutated ODN-ZF(G/T)-BI (3.3 ×10^4 M^{-1} s^{-1}) and GC/CT-mutated ODN-ZF(GC/CT)-BI (5.3 ×10^2 M^{-1} s^{-1}) were drastically reduced when compared to the wild-type (4.4 ×10^5 M^{-1} s^{-1}), as listed in Table 2.4. These observations are consistent with the corresponding differences between their K_D values (Table 2.4). These results further indicated that the observed k_{app} for a given pair of ZF-SNAP and BI-modified ODN actually reflected the K_D of the pair according to the criterion of $k_{cov} \ll k_{off}$, and demonstrated that the selectivity of MA-SNAP to BI-modified DNA sequence was driven by the specific sequence recognition.

Table 2.4 Kinetic parameters (K_D) for binding between ZF-SNAP and BI-modified DNA, (k_{cross}) for self-ligating reaction of SNAP with BI, parameters (k_{app}) for the crosslink reaction between MA and BI-modified DNA. The measurement results of k_{app} on PAGE was described on Figure 2.11.

MAs	Substrate-modified ODN	K_D (nM)	$K_D/K_{D(ODN-ZF)}$	k_{app} ($M^{-1}s^{-1}$)	$k_{app(ODN-ZF-BI)}/k_{app}$
ZF-SNAP	ODN-ZF-BI	56 ± 7.0	1	(4.4 ± 0.5) ×10^5	1
	ODN-ZF(G/T)-BI	(1.1 ± 0.2) × 10^3	20	(3.3 ± 0.1) ×10^4	13
	ODN-ZF(GC/CT)-BI	> 10^4	>200	(5.2 ± 0.1) ×10^2	846

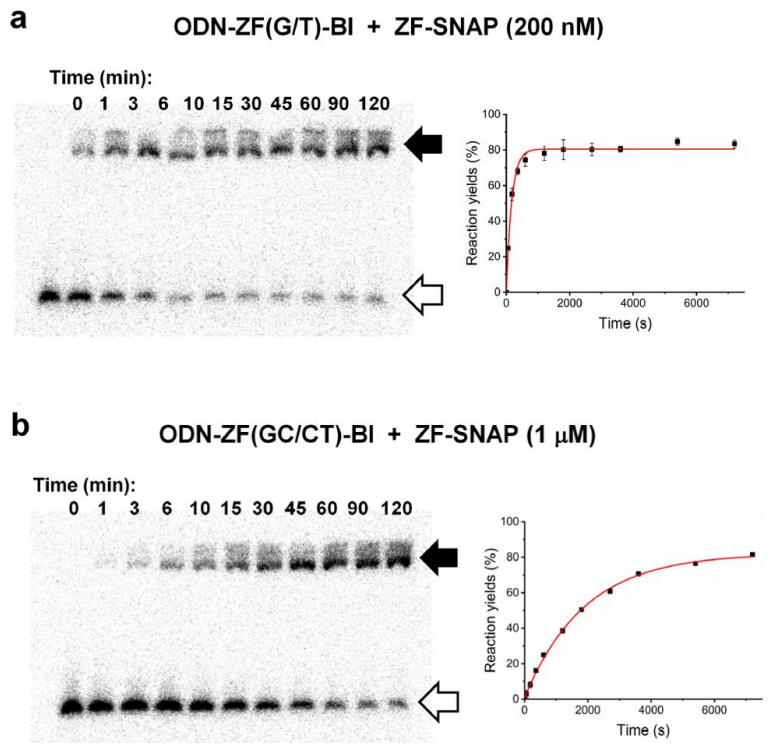


Figure 2.11 Autoradiograms show denaturing PAGE analyses of the samples from the reactions between 5'-^{32}P end-labeled ODNs with MAs. (a) ODN-ZF(G/T)-BI (0.5 nM) with ZF-SNAP (200 nM), (b) ODN-ZF(GC/CT)-BI (0.5 nM) with ZF-SNAP (1 μM), open and filled arrows denote free and MA bound ODNs, respectively. Time-course plots for the crosslinking reaction of substrate modified ODN with MA to obtain k_{app}, that are listed in Table 2.4.

2.3 Conclusion

In this study, two BG derivatives, BI and deBG, were designed by altering the interaction sites with SNAP-tag. These substrates were chosen as the candidates because impairing the intermolecular interaction between SNAP-tag and its substrate would reduce the rate constant of the reaction between them. For further characterization of sequence-selectivity of MA, BI was modified on DNA to elucidate the kinetic details of sequence-selective modification by MA-SNAP. As a result, sequence-selectivity of MA-SNAP with BI-modified DNA was demonstrated from three aspects: sequence-selective conjugation among multiple DNA sequences; calculation of all kinetic parameters to conclude $k_{cov} \ll k_{off}$; proportional relationship between k_{app} and K_D. Those results strongly proved the fact that k_{cov} of MA reaction is the vital key and boundary condition for conducting recognition-driven conjugation between MA and DNA.

2.4 Materials and methods

2.4.1 Materials

Oligonucleotides for substrate modification were purchased from Japan Bio Services Co., LTD (Saitama, Japan). Alexa Fluor 488-labelled oligonucleotides and ATTO 488 carboxylic acid succinimidyl ester were purchased from Invitrogen (Carlsbad, CA, USA). Ultrafree-MC-DV centrifugal filters were purchased from Merck Millipore (Billerica, MA, USA). A Cosmosil 5C18-MS II column (4.6 i.d. × 150 mm) and high-performance liquid chromatography (HPLC)-grade acetonitrile were purchased from Nacalai Tesque Inc. (Kyoto, Japan). ULTRON VX-ODS packed columns were purchased from Shinwa Chemical Industries (Kyoto, Japan). Reagents for chemical synthesis, HPLC-grade trifluoroacetic acid, gel electrophoresis-grade acrylamide,

bis(acrylamide), phenol, and all other chemicals and reagents were purchased from FUJIFILM Wako Pure Chemical Industries, Ltd. (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) or Nacalai Tesque, Inc. ^{1}H NMR and ^{13}C NMR spectra were recorded on JEOL JNM-ECP100 FT NMR (JEOL, Tokyo, Japan), JEOL JNM-ECA600 FT spectrometer (JEOL, Tokyo, Japan) or Bruker AVANCE 600 MHz NMR spectrometer (Billerica, MA, USA) equipped with a cryogenic probe, and chemical shifts were expressed in parts per million (ppm) and the coupling constants were calculated in Hz. ESI-TOF-MS spectra were recorded on a JEOL JMS-T100LP spectrometer (JEOL, Tokyo, Japan). Matrix-assisted laser ionization time-of-flight mass spectrometry (MALDI-TOF MS) spectra were recorded on an AXIMA-Confidence (Shimadzu, Kyoto, Japan) using sinapinic acid (SA) or 3-hydroxypicolinic acid (3-HPA) as a matrix. The liquid chromatography-mass spectrometry (LC-MS) were recorded on QTOF X500R MS system (SCIEX, USA) with Exion LC HPLC system (SCIEX, USA). Phosphate buffer (PB) was prepared by mixing 20 mM Na_2HPO_4 and 20 mM NaH_2PO_4 (pH 7.0). DNA origami buffer (pH 8.0) contained 40 mM Tris-HCl, 20 mM acetic acid, and 12.5 mM $MgCl_2$.

2.4.2 Synthesis of benzylinosine and 7-deaza-benzylguanine[7,10]

*Synthesis of **1***

To a solution of cyanobenzylaldehyde (1 g, 7.4 mmol) in anhydrous THF (10 mL), a solution of $LiAlH_4$ (2.83 g, 74.5 mmol) in anhydrous THF (30 mL) was added slowly at 0℃. After 24 h of refluxing at 90℃, the color of suspension changed from purple to green. The reaction was cooled to room temperature and quenched with H_2O (2 mL), 15% NaOH (4 mL) and H_2O (6 mL). The resulting precipitate was filtered by celite, washed with ethyl acetate (50 mL, 3 times) , and filtrate was dried in *vacuo*. The

residue was added in 1 M HCl (50 mL) and washed with ethyl acetate (50 mL). The aqueous layer was concentrated in *vacuo* to give **1** (0.404 g, 32%) as a white solid. **1**: ^{1}H NMR (600 MHz; D_2O): δ 7.29 (s, 4H, Ar), 4.49 (s, 2H, -Ar-CH_2-OH), 4.02(s, 2H, -Ar-CH_2-NH_2). ESI-TOF-MS *m/z* = 138.04 $[M+H]^+$ (calcd. for $C_8H_{12}NO$, 138.09).

*Synthesis of **2***

To a solution of **1** (1.29 g, 9.4 mmol) and triethylamine (1.16 mL. 8.3 mmol) in methanol (15 mL) was added trifluoroacetic acid ethyl ester (1.35 mL, 11.3 mmol) at room temperature. After 17 h of stirring, the reaction mixture was concentrated in *vacuo*. The residue was extracted with ethyl acetate (50 mL) and washed with saturated brine (50 mL, 3 times), and was dried over Mg_2SO_4. After removal of the solvents in *vacuo*, the crude product was dissolved in a mixture of chloroform and methanol, and purified by a silica gel column chromatography (a gradient of 0% to 7% methanol in chloroform) to give **2** (0.906 g, 38%) as a white solid. **2**: ^{1}H NMR (600 MHz; $CDCl_3$): δ 7.26 (dd, 4H, J_H = 7.8 Hz, Ar-H), 7.18 (s, 1H, -NH-CO-CF_3), 6.52 (s, 1H, -OH), 4.63 (d, 2H, J_H = 7.2 Hz, -Ar-CH_2-OH), 4.46 (d, 2H, J_H = 6.6 Hz, -Ar-CH_2-NHCO-). ESI-TOF-MS *m/z* = 256.06 $[M+Na]^+$ (calcd. for $C_{10}H_{10}NO_2F_3Na$, 256.06).

*Synthesis of **3***

1-Methylpyrrolidin (1.54 mL, 14.5 mmol) was added to a suspension of 6-chloro-5*H*-purine (1.0 g, 6.5 mmol) in N,N-dimethylformamide (DMF, 40 mL). The color of suspension was changed from yellow to white and finally turned to milky white. The reaction mixture was stirred for 25 h at room temperature. Next, acetone (25 mL) was added and the precipitate was filtered to afford **3** (1.266 g, 96%) as a white solid. ^{1}H NMR (300 MHz; DMSO-d6): δ 9.02 (d, 1H, J_H = 8.8 Hz, 5*H*-purine *H*8), 8.95 (d,

1H, J_H = 8.8 Hz, 5*H*-purine *H*5), 4.70-4.64 (m, 2H, -N^+-C$\underline{H_2}$-), 4.16-4.12 (m, 2H, -N^+-C$\underline{H_2}$-), 3.73 (s, 3H, -N^+-C$\underline{H_3}$), 2.28 (t, 2H, J_H = 2.1 Hz, -C$\underline{H_2}$-), 2.08(d, 2H, J_H = 1.5 Hz, -C$\underline{H_2}$-). ESI-TOF-MS *m/z* = 204.11 $[M]^+$ (calcd. for $C_{10}H_{14}N_5$, 204.12).

*Synthesis of **4***

The compound **2** (0.77 g, 3.0 mmol) was added into a solution of NaH (0.22 g, 9.0 mmol) in anhydrous DMF (20 mL) under the N_2 atmosphere at room temperature. To this solution, **3** (0.61 g, 3.0 mmol) and potassium-t-butoxide (1.0 g, 9.0 mmol) were added and the reaction mixture was stirred for 16 h at room temperature. The yellowish solution was concentrated in *vacuo*. The crude residue was dissolved in methanol, adsorbed on silica gel, and purified by a column chromatography with a gradient of 0% to 9% methanol in ethyl acetate to give **4** (0.67 g, 64%) as a white solid. **4**: ^{1}H NMR (300 MHz; DMSO-d6): δ 8.50 (d, 1H, 2*H*-purine *H*8), 8.38 (s, 1H, 2*H*-purine *H*5), 7.54 (m, 2H, Ar-$\underline{H}$), 7.42 (t, 2H, J_H = 2.9 Hz, Ar-$\underline{H}$), 5.61 (s, 2H, -O-C$\underline{H_2}$-Ar-), 3.94 (s, 2H, -Ar-C$\underline{H_2}$-NHCO-). ESI-TOF-MS *m/z* = 374.07 $[M+Na]^+$ (calcd. for $C_{15}H_{12}N_5O_2F_3Na$, 374.08).

*Synthesis of BI-NH_2 (**5**)*

The compound **4** (670 mg, 1.9 mmol) was suspended in a solution containing methanol (63 mL) and H_2O (7 ml). After addition of K_2CO_3 (1.3 g, 9.5 mmol), the reaction mixture was refluxed for 6 h. The solvent was removed in *vacuo* and the crude product was purified by a silica gel column chromatography (MeOH: $CHCl_3$: Et_3N = 1: 5: 0.05) to give **5** as a white solid (0.456 mg, 94% yield). **5**: ^{1}H NMR (300 MHz; DMSO-d6): δ 10.02 (s, 2H, -NH_2), 8.51 (d, 1H, J_H = 1.4 Hz, 5*H*-purine *H*8), 8.95 (s, 1H, 5*H*-purine *H*5), 7.51 (t, 2H, J_H = 3.0 Hz, Ar-$\underline{H}$), 7.30 (t, 2H, J_H = 4.0 Hz, Ar-$\underline{H}$), 5.59 (d,

2H, J_H = 4.4 Hz, -O-CH_2-Ar-), 4.40 (s, 2H, -Ar-CH_2-NHCO-). ESI-TOF-MS *m/z* = 256.11 $[M+H]^+$ (calcd. for $C_{13}H_{15}N_5O$, 256.12).

Synthesis of ***BI-GLA-NHS***

To a DMF solution (10 mL) containing disuccinimidyl glutarate (640 mg, 1.96 mmol, 1 eq.) and triethylamine (272 μl, 1.96 mmol, 1 eq.) as added a DMF solution (10 mL) of **5** (0.150 g, 0.59 mmol, 0.3 eq.). After 1 h, **5** (0.75g, 0.23 mmol, 0.15 eq.) was added dropwise to the reaction mixture. Formation of **6** and byproducts were monitored by TLC. After reacted for 1.5 h and 3 h, 0.09 eq. and 0.6 eq., respectively, of **5** were added to complete the reaction. After reacting for 4 h at room temperature, the solution was concentrated and purified by a silica gel column chromatography under the gradient of 2% to 10% MeOH in chloroform. Concentration of the fractions containing **5** gave a white solid, which was re-dissolved in chloroform and washed by water. During the washing procedure, precipitates were observed in the aqueous layer. The aqueous layer was filter to isolate the white solid. The organic layer was combined with the filtered solid, then concentrated to give **BI-GLA-NHS** (96 mg, 35%). **BI-GLA-NHS**: ^{1}H NMR (300 MHz; DMSO-d6): δ 10.54 (s, 1H, 5*H*-purine *H*9), 8.50 (s, 1H, 5*H*-purine *H*3), 8.36 (s, 1H, 5*H*-purine *H*8), 7.47 (d, 2H, J_H = 6.9 Hz, Ar-H), 7.27 (d, 2H, J_H = 6.9 Hz, Ar-H), 5.59 (s, 2H, -Ar-CH_2-O-), 4.27 (d, 2H, J_H = 5.8 Hz, -Ar-CH_2-NHCO-), 2.81 (s, 4H, NHS-H), 2.70 (t, 2H, J_H = 7.0 Hz, -NHCO-CH_2-), 2.26 (t, 2H, J_H = 7.1 Hz, -OCO-CH_2-), 1.86 (q, 2H, J_H = 7.3 Hz, -CH_2-). ^{13}C NMR (150 MHz; DMSO-d6): δ 171.6, 170.7, 169.2, 159.0, 151.6, 143.4, 140.1, 135.3, 128.8, 127.8, 67.9, 42.3, 34.1, 30.1, 25.9, 20.9. ESI-TOF-MS *m/z* = 489.14 $[M+Na]^+$ (calcd. for $C_{22}H_{22}N_6O_6Na$, 489.15). The ^{1}H NMR and ^{13}C NMR spectra were shown as Figure S12 and Figure S13.

*Synthesis of **7***

1-Methylpyrrolidin (0.7 mL, 6.6 mmol) was added to a suspension of 4-chloro-7*H*-pyrrolo[2,3-*d*]pyridmidin-2-amine (0.5 g, 3.0 mmol) in DMF (10 mL). The reaction mixture was stirred for 25 h at 100℃. After cooling to the room temperature, acetone (25 mL) was added and the resulting precipitate was filtered as a white solid. The crude was purified by a reverse phase column chromatography (ODS, H_2O/MeOH: a linear gradient from 9: 1 to 1: 9 (*v/v*)) to give **7** in 28% yield (180 mg, 0.83 mmol). **7**: ^{1}H NMR (300 MHz; CD_3OD): δ 6.75 (d, 1H, J_H = 3.6 Hz, C8-H), 6.45 (d, 1H, J_H = 3.6 Hz, C7-H), 3.82-3.78 (m, 2H, -N^+-CH_2-), 3.49-3.44 (m, 2H, -N^+-CH_2-), 2.97 (s, 3H, -N^+-CH_3), 2.16 (m, 2H, -CH_2-), 1.76 (m, 2H, -CH_2-). ESI-TOF-MS *m/z* = 218.13 $[M]^+$ (calcd. for $C_{11}H_{16}N_5$, 218.14).

*Synthesis of **8***

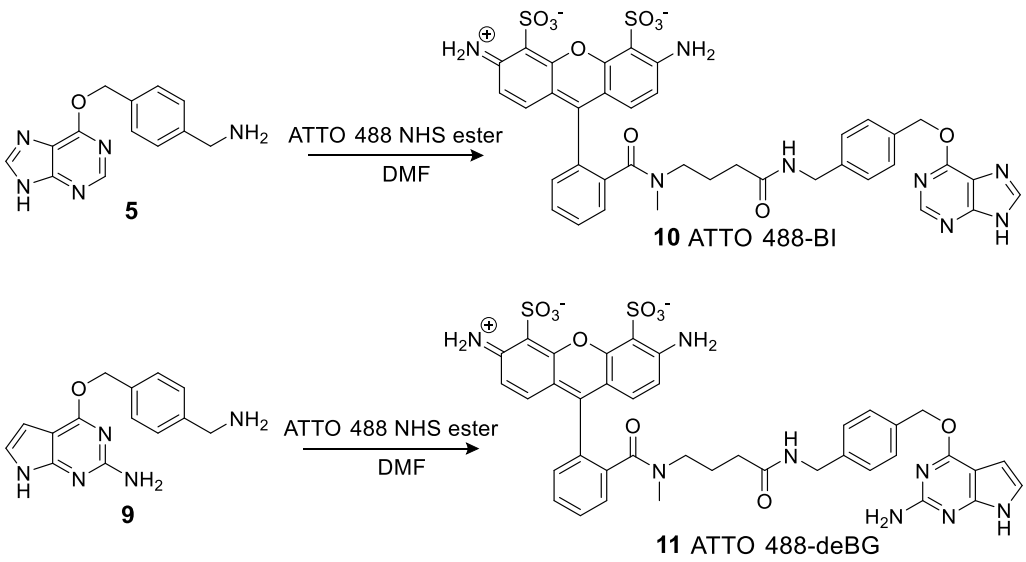
The compound **7** (50 mg, 0.23 mmol) was added to a solution of **2** (0.10 g, 0.42 mmol) in anhydrous DMF (20 mL) under the N_2 atmosphere at room temperature. The potassium-t-butoxide (0.15 g, 1.3 mmol) was added to the solution and stirred for 10 h at room temperature. The crude was dissolved in methanol, adsorbed on silica gel, and purified by a column chromatography ($CHCl_3$/MeOH: a linear gradient from 0% to 10% MeOH) to give **8** in 61% yield (50 mg, 0.14 mmol). **8**: ^{1}H NMR (300 MHz; CD_3OD): δ 7.73-7.70 (m, Ar-H), 7.55-7.46 (m, Ar-H), 6.81-6.79 (m, C8-H), 6.34-6.28 (m, C7-H), 5.50 (s, 2H, -O-CH_2-Ar-), 4.46 (s, 2H, -Ar-CH_2-NHCO-). ESI-TOF-MS *m/z* = 364.07 $[M-H]^-$ (calcd. for $C_{16}H_{13}F_3N_5O_2$, 364.10).

*Synthesis of deBG-NH_2 (**9**)*

The compound **8** (50 mg, 0.14 mmol) was suspended in a solution of methanol

(10 mL) and H_2O (2 mL). After addition of K_2CO_3 (100 mg, 0.72 mmol), the reaction mixture was refluxed for 3 h. The solvent was removed in *vacuo* and the residue was purified by a column chromatography (amino-bonded silica, $CHCl_3$/MeOH: a linear gradient from 0% to 10% MeOH) to give **9** in 27% yield (10 mg, 37 μmol). **9**: ^{1}H NMR (300 MHz; CD_3OD): δ7.44 (m, 2H, C2-NH_2), δ7.36-7.30 (m, 4H, Ar-H), 6.79-6.78 (m, 2H, C8-H), 6.32-6.27 (m, 2H, C7-H), 5.47 (m, 2H, -O-CH_2-Ar-), 4.58 (m, 2H, -Ar-CH_2-NHCO-), 3.80 (m, 2H, -CH_2-NH_2). ESI-TOF-MS *m/z* = 270.12 $[M+H]^+$ (calcd. for $C_{14}H_{15}N_5O$, 270.13).

2.4.3 Synthesis of a fluorophore-modified substrate of protein tag



Scheme 3. Synthetic scheme of ATTO 488-labeled BI (**10**) and deBG (**11**)

The compound **5** (0.2 mg, 0.75 μmol) was dissolved in DMF (35 μL) and mixed with a DMF solution (100 μL) of ATTO 488-NHS (0.1 mg, 0.15 μmol). The reaction mixture was incubated for 24 h at 30°C. The reaction was monitored by means of HPLC (phase A: 0.05% TFA containing H_2O and phase B: 50 % of 0.05% TFA containing acetonitrile). Upon completion of the reaction, the reaction mixture was diluted by using water (365 μL) and purified by HPLC with a gradient of 5% to 65%

phase B in 40 min. The solution containing **10** was lyophilized to give **10** as yellow powder. The compound **11** was prepared as the same way as **10** by using **9** as the starting material. **10**: ESI-TOF-MS, *m/z* = 825.12 $[M]^-$ (calcd. for $C_{38}H_{33}N_8O_{10}S_2$, 825.18). **11**: ESI-TOF-MS, *m/z* = 839.25 $[M]^-$ (calcd. for $C_{39}H_{35}N_8O_{10}S_2$, 839.19) HPLC chart for checking the purity of **10 and 11** were shown in Figure 2.4.

2.4.4 Preparation of substrate-modified ODNs

Coupling reaction between the amino group on ODN (50 μM) with the succinimidyl derivative of protein-tag substrate (for example BI-GLA-NHS, 3 mM) was carried out for 12 h at room temperature in a phosphate buffer (50 mM, pH 8.0) containing 50% dimethyl sulfoxide (DMSO). The product was purified by reversed-phase HPLC on a Cosmosil 5C18-MS II column (4.6 × 150 mm, eluted with 100 mM triethylammonium acetate buffer, pH 7.0, with a linear gradient over 30 min from 10 to 25% acetonitrile at a flow rate of 1.0 ml min^{-1}) and characterized by MALDI-TOF-MS spectrometry using 3-HPA as matrix.

2.4.5 Measurement of fluorescent polarization

FP assay was conducted by using a plate reader Infinite 200 PRO FPlex (Tecan Group Ltd., Mannedorf, Switzerland) to determine the rate constant of intermolecular covalent bond formation (k_{cross}, Figure 2.5). The reaction between protein-tag (10 μM) and substrate (20 nM) was carried out in 1× origami buffer (pH 8.0) supplemented with 100 μM mercaptoethanol, 1 mM $ZnCl_2$, 0.02% Tween 20, 200 nM BSA, 100 nM calf thymus DNA and 200 mM NaCl at room temperature. Reaction mixture was incubated in a 96-well microplate inside the plate reader for 10 min and the parallel (I||) and perpendicular (I⊥) fluorescence intensities were recorded every minute.

The acquired time-dependent FP intensities were fitted to a reaction model assuming the first-order kinetics for calculating the second-order rate constants.

2.4.6 ^{32}P-labelling for characterization of cross-linking reaction

The 5′ end of ODN was ^{32}P-labelled as reported method.[16] To study the covalent bond formation between the substrate and protein-tag in DNA-MA complex, ^{32}P-labelled ODN containing protein-tag substrate (0.5 nM) was incubated with MA in 1× origami buffer (pH 8.0) supplemented with 100 μM mercaptoethanol, 1 mM $ZnCl_2$, 0.02% Tween 20, 200 nM BSA, 100 nM calf thymus DNA and 200 mM NaCl at room temperature. The cross-linking reaction was quenched by adding SDS and formamide to the reaction solution, and the resultant mixture was heated at 80°C for 3 min and rapidly cooled down on ice. The aliquots were analyzed by denaturing PAGE containing 8 M urea, and the gel images were recorded by using Storm 860 Molecular Imager (Amersham). The kinetic data reflected from the bands of gel were fitted to a reaction model assuming the first-order kinetics, and hence the second-order rate constants were determined depended on the concentration of MA.

2.4.7 Stopped flow analyses of fluorescence polarization

Real time observation of the association process of zif268 to ODN-ZF was analysed by time-dependent FP change with the rapid mixing method by stopped-flow (RX2000 Rapid Mixing Stopped-Flow Unit, Applied Photophysics, Leatherhead Surrey, UK). The parallel and perpendicular fluorescence intensities were collected using 473 nm laser excitation and a 500 nm filter with polarization filter in the emission path. The fluorescence intensities were recorded every 0.005 s during the mixing of ZF-CLIP with Alexa 488-ODN-ZF and FP value was calculated as shown in Figure 2.10(b). The black

arrow in Figure 2.10(b) indicates that ZF-CLIP solution was mixed with Alexa 488-ODN-ZF solution. The FP signal of Alexa 488-ODN-ZF increases upon binding to ZF-CLIP. The final concentration of Alexa 488-ODN-ZF was 5 nM and the concentrations of ZF-CLIP were100 nM, 300 nM and 500 nM.

The kinetic parameter k_{obs} (s^{-1}) was obtained by fitting to a reaction model assuming the first-order kinetics (eq. 2.1).

$$P = (P_{max} - P_0)(1 - e^{-kobst}) + P_0 \qquad \text{(Eq. 2.1)}$$

P, P_{max} and P_0 represent the fluorescence polarization at the certain time, the maximum fluorescence polarization, and the fluorescence polarization in the absence of modular adaptor, respectively.

The k_{obs} values from FP assay were plotted in a dose-dependent manner as shown in Figure 2.10(c) and k_{on} ($M^{-1}s^{-1}$) and k_{off} (s^{-1}) were determined by following eq. 2.2.

$$k_{obs} = k_{on} \times [MA] + k_{off} \qquad \text{(Eq. 2.2)}$$

where MA represents modular adaptor. The calculated value of k_{on} and k_{off} were 8.6×10^6 M^{-1} s^{-1} and 0.64 s^{-1}, respectively. And from k_{on} and k_{off}, K_D (= k_{off}/k_{on}) value between zif268 and ODN-ZF can be calculated out as 72 nM, which was consistent with the determined K_D value (56 nM), which we reported previously (as shown Table 2.2).[4]

2.5 References

(1) Nakata, E.; Dinh, H.; Ngo, T. A.; Saimura, M.; Morii, T. A Modular Zinc Finger Adaptor Accelerates the Covalent Linkage of Proteins at Specific Locations on DNA Nanoscaffolds. *Chem. Commun.* **2015**, *51*, 1016–1019.

(2) Ngo, T. A.; Nakata, E.; Saimura, M.; Morii, T. Spatially Organized Enzymes Drive Cofactor-Coupled Cascade Reactions. *J. Am. Chem. Soc.* **2016**, *138*, 3012–3021.

(3) Nakata, E.; Dinh, H.; Nguyen, T. M.; Morii, T. DNA Binding Adaptors to Assemble Proteins of Interest on DNA Scaffold, *Methods Enzymol.*, **2019**, *617*, 287–322.

(4) Nguyen, T. M.; Nakata, E.; Zhang, Z.; Saimura, M.; Dinh, H.; Morii, T. Rational Design of a DNA Sequence-Specific Modular Protein Tag by Tuning the Alkylation Kinetics. *Chem. Sci.* **2019**, *10*, 9315–9325.

(5) Nguyen, T. M.; Nakata, E.; Saimura, M.; Dinh, H.; Morii, T. Design of Modular Protein Tags for Orthogonal Covalent Bond Formation at Specific DNA Sequences. *J. Am. Chem. Soc.* **2017**, *139*, 8487–8496.

(6) Schreiber, G.; Haran, G.; Zhou, H. X. Fundamental Aspects of Protein−Protein Association Kinetics. *Chem. Rev.* **2009**, *109*, 839–860.

(7) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A General Method for the Covalent Labeling of Fusion Proteins with Small Molecules in Vivo. *Nat. Biotechnol.* **2003**, *21*, 86–89.

(8) Mollwitz, B.; Brunk, E.; Schmitt, S.; Pojer, F.; Bannwarth, M.; Schiltz, M.; Rothlisberger, U.; Johnsson, K. Directed Evolution of the Suicide Protein O 6-Alkylguanine-DNA Alkyltransferase for Increased Reactivity Results in an Alkylated Protein with Exceptional Stability. *Biochemistry* **2012**, *51*, 986–994.

(9) Spratt, T. E.; De Los Santos, H. Reaction of O6-Alkylguanine-DNA Alkyltransferase with O6-Methylguanine Analogues: Evidence That the Oxygen of O6-Methylguanine Is Protonated by the Protein To Effect Methyl Transfer. *Biochemistry* **1992**, *31*, 3688–3694.

(10) Ohyanagi, T.; Shima, T.; Okada, Y.; Tsukasaki, Y.; Komatsuzaki, A.; Tsuboi, S.; Jin, T. Compact and Stable SNAP Ligand-Conjugated Quantum Dots as a Fluorescent Probe for Single-Molecule Imaging of Dynein Motor Protein. *Chem. Commun.* **2015**, *51*, 14836–14839.

(11) Pavletich, N. P.; Pabo, C. Zinc Finger-DNA Recognition : Crystal Structure of a Zif268-DNA Complex at 2 *Science* **1991**, *252*, 809–818.

(12) Sera, T.; Uranga, C. Rational Design of Artificial Zinc-Finger Proteins Using a Nondegenerate Recognition Code Table. *Biochemistry* **2002**, *41*, 7074–7081.

(13) Hope, I. A.; Mahadevan, S.; Struhl, K. Structural and Functional Characterization of the Short Acidic Transcriptional Activation Region of Yeast GCN4 Protein. *Nature* **1988**, *333*, 635–640.

(14) Kim, J. S.; Pabo, C. O. Getting a Handhold on DNA: Design of Poly-Zinc Finger Proteins with Femtomolar Dissociation Constants. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 2812–2817.

(15) Geertz, M.; Shore, D.; Maerkl, S. J. Massively Parallel Measurements of Molecular Interaction Kinetics on a Microfluidic Platform. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 16540–16545.

(16) Nakata, E.; Liew, F. F.; Uwatoko, C.; Kiyonaka, S.; Mori, Y.; Katsuda, Y.; Endo, M.; Sugiyama, H.; Morii, T. Zinc-Finger Proteins for Site-Specific Protein Positioning on DNA-Origami Structures. *Angew. Chemie - Int. Ed.* **2012**, *51*, 2421–2424.

CHAPTER 3

Expanding selectivity of modular adaptor by taking advantage of sequence-specificity and chemoselectivity

3.1 Introduction

As I introduced, modular adaptor (MA), consists of both a sequence-specific DNA-binding domain (DBD) and a self-ligating protein-tag, was used for recognition driven DNA-protein conjugation (Figure 3.1a).[1–3] However, recognition-driven conjugation can only be performed in the case of $k_{cov} \ll k_{off}$, which the specificity of reaction would be conducted by DBD to realize DNA sequence-specific reaction. CLIP-tag [4] shows moderate reactivity to its specific substrate, O^2-benzylcytosine (BC).[5] So the sequence-selectivity of MA containing CLIP-tag (MA-CLIP) is directed by the K_D between DBD and DNA binding sequence. In contrast, overly high reactivity of SNAP-tag with its original substrate, O^2-benzylguanine (BG), cannot reach $k_{cov} \ll k_{off}$, so that it failed in producing sequence-selective MA.[6] Only if the reactivity of SNAP-tag was tuned by using benzylinosine (BI) as the substrate of SNAP-tag for the modification of DNA binding sequence, MA consisting of SNAP-tag [7] (MA-SNAP) would demonstrate sequence-selectivity. As a result, two series of sequence-selective MA were exploited to facilitate multiple enzyme localization on DNA scaffold. Given that 3 types of DBD, zinc finger proteins zif268 (ODN-ZF)[8] and AZP4 (ODN-AZ)[9] and the basic-leucine zipper class of protein GCN4,[10] are included, three different proteins of interests (POIs) are able to be site-specifically assembled on DNA scaffold.

To enable more diversity of POI-assembly on DNA scaffold, chemoselectivity among different protein tags emerged as a robust strategy to separate different series of sequence-selective MA (Figure 3.1b-d). SNAP-tag, CLIP-tag and Halo-tag [11] were previously combined with different zinc finger proteins to create 3 unique MAs.[6] And chemoselectvity of protein tag drove 3 MAs to orthogonally react with the specific substrates on DNA binding sequences. But the specificity of CLIP-tag against BG was not good enough to realize chemoselectivity between SNAP-tag and CLIP-tag. Therefore, the chemoselectivity of CLIP-tag with BI was aimed to be explored. This

chemoseletivity of CLIP-tag against BI would be the premise to realize selectivity between two series of sequence-selective MA.

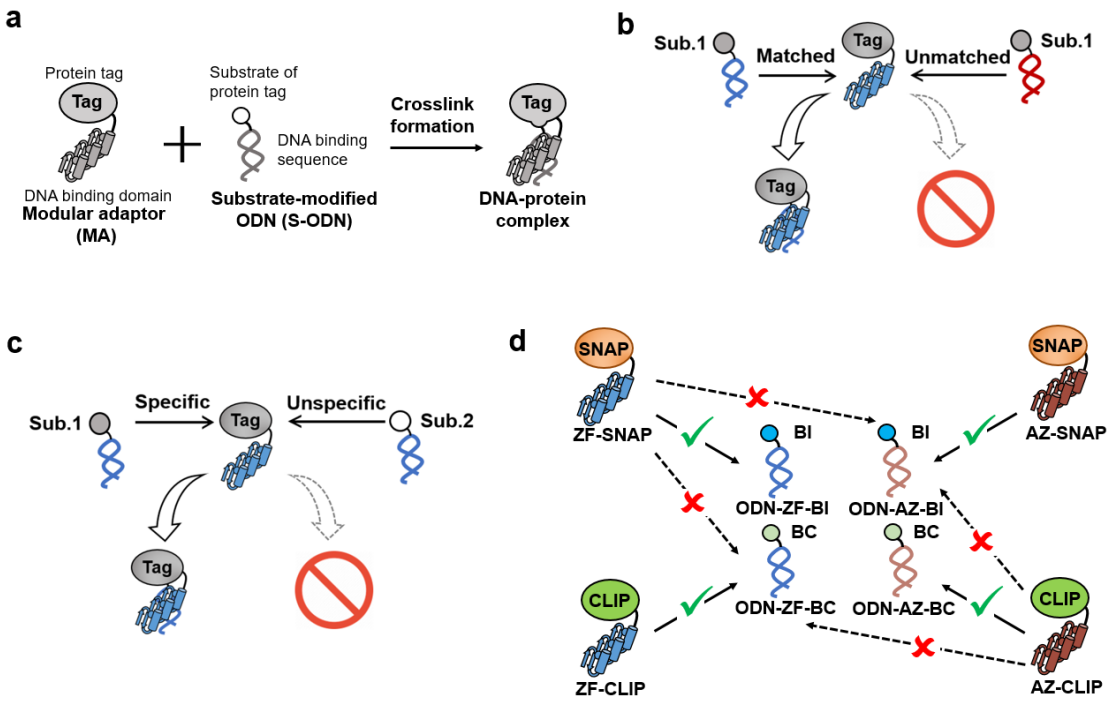


Figure 3.1 Illustration of recognition-driven conjugation mediated by MA. (a) Reaction scheme between MAs and substrate-modified DNA binding sequence. (b) Orthogonal cross-linking reactions of MAs conducted by DNA sequence specificities of zinc finger proteins. (c) Orthogonal cross-linking reactions of MAs driven by the chemoselectivity of the protein tag. (d) The new strategy combined chemoselectivity of the protein tag and sequence-selectivity of DNA binding domain to expand the selectivity of modular adaptors.

3.2 Results and discussion

3.2.1 Selective reactions by a series of modular adaptors consisting of SNAP- or CLIP-tag

As it was introduced in 2.2.2, BI showed remarkable selectivity against CLIP-tag, while enabling sequence-selectivity of MA-SNAP. Coupled with BC for distinguishing CLIP-tag, chemoselectivity between SNAP-tag and CLIP-tag is able to be used for diversifying Mas. For this purpose, MA reaction with SNAP-tag or CLIP-tag was further examined to show that k_{cov} also exerted an impact on the specificity of MAs through chemoselectivity. BI-modified ODN-ZF (ODN-ZF-BI) reacting with ZF-CLIP showed the sufficiently low reactivity of BI to CLIP-tag, causing neglectable formation of the cross-linking by this pair when ZF was used as DBD (Figure 3.2c). Likewise, k_{app} of the cross-linking formation between ZF-SNAP and ODN-ZF-BC was three orders of magnitude lower than that of ZF-CLIP with the same ODN (Figure 3.2a and 3.2b, Table 3.1). These results verified that the pair of newly synthesized BI-modified ODN and MA-SNAP exerts an excellent specificity against the pair of BC-modified ODN and MA-CLIP.

The observed orthogonality between the BI/SNAP-tag and the BC/CLIP-tag pairs prompted me to systematically investigate the reactions by six types of unique MAs, i.e. each three types of MA-SNAP and MA-CLIP, to evaluate their selectivity for DNA-protein conjugation (Figure 3.3). Each of the ODNs was reacted with the six types of MAs. BI-modified ODNs exhibited sequence-selective modifications with high reaction yield only for the matched pair of MA-SNAP. Especially, the reaction yield of ZF-SNAP to ODN-ZF-BI and AZ-SNAP to ODN-AZ-BI exceeded 90% (Figures 3.3b-c). All the matched pairs of MA-SNAP and BI-modified ODN showed nearly perfect specificity against MA-CLIP, confirming the high selectivity towards the six different MAs. Furthermore, BC-modified ODNs reacted sequence-selectively with

their respective matched MA-CLIP (Figure 3.4). Unlike the case of BI-modified ODN, ODN-ZF-BC formed a cross-linking with ZF-SNAP as reported result. Because the reactivity of BC to SNAP-tag was not low enough to avoid the side reaction between BC-modified ODN and MA-SNAP, while application of lower concentration of MA (80 nM) was preferred to avoid the cross-linking reaction between this pair. Under such conditions, combination of the sequence-specificity of DNA-binding protein and the chemoselectivity of protein-tag with its substrate allowed us to achieve the selective DNA-protein cross-linking reaction within the six MAs constructed.

Table 3.1 Kinetic parameters (k_{app}) for the crosslink reaction between modular adaptor and matched substrate-modified ODN. The measurement results of k_{app} on PAGE was described on Figure 3.2.

MAs	Substrate-modified ODN	k_{app} ($M^{-1}s^{-1}$)	$k_{app(ODN\text{-}ZF\text{-}BI)}/k_{app}$
ZF-SNAP	ODN-ZF-BI	$(4.4 \pm 0.5) \times 10^5$	1
	ODN-ZF-BC	$(8.0 \pm 0.6) \times 10^2$	550
ZF-CLIP	ODN-ZF-BC	$(1.3 \pm 0.1) \times 10^5$	4
	ODN-ZF-BI	n.d.	n.d

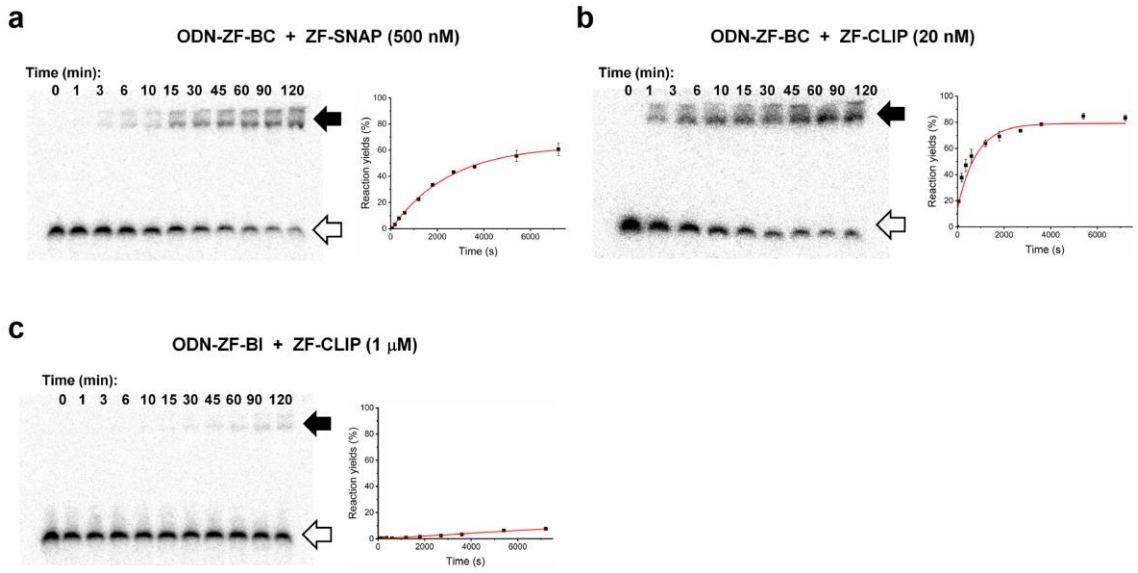


Figure 3.2 Autoradiograms show denaturing PAGE analyses of the samples from the reactions between 5'-^{32}P end-labeled ODNs with MAs. (a) ODN-ZF-BC (0.5 nM) with ZF-SNAP (500 nM), (b) ODN-ZF-BC (0.5 nM) with ZF-CLIP (20 nM). (c) ODN-ZF-BI (0.5 nM) with ZF-CLIP (1 μM). Open and filled arrows denote free and MA bound ODNs, respectively. Time-course plots for the crosslinking reaction of substrate modified ODN with MA to obtain the rate constants (k_{app}) ($M^{-1}s^{-1}$), that are listed in Table 3.1.

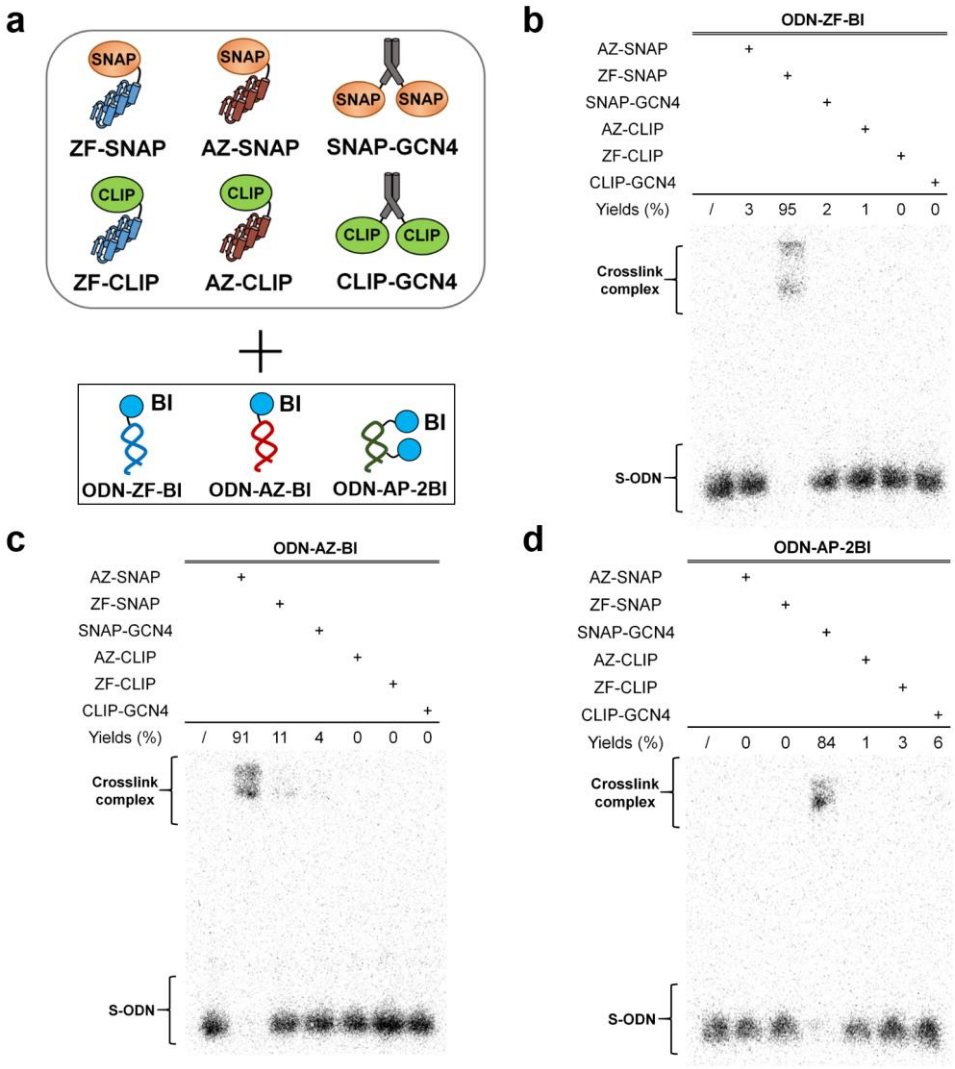


Figure 3.3 Selective crosslink reactions between BI-modified ODNs and 6 types of MAs. (a) An illustration of 6 types of MAs reacting with BI-modified DNA binding sequences. (b-d) Cross-linking reactions of the MAs with substrate-modified ODNs. All six types of MAs (each 150 nM) was incubated with 0.5 nM 5'-^{32}P-end-labelled ODN-ZF-BI (b), ODN-AZ-BI (c), and ODN-AP-2BI (d) for 30 min at room temperature. The samples after reaction were analyzed by denaturing PAGE.

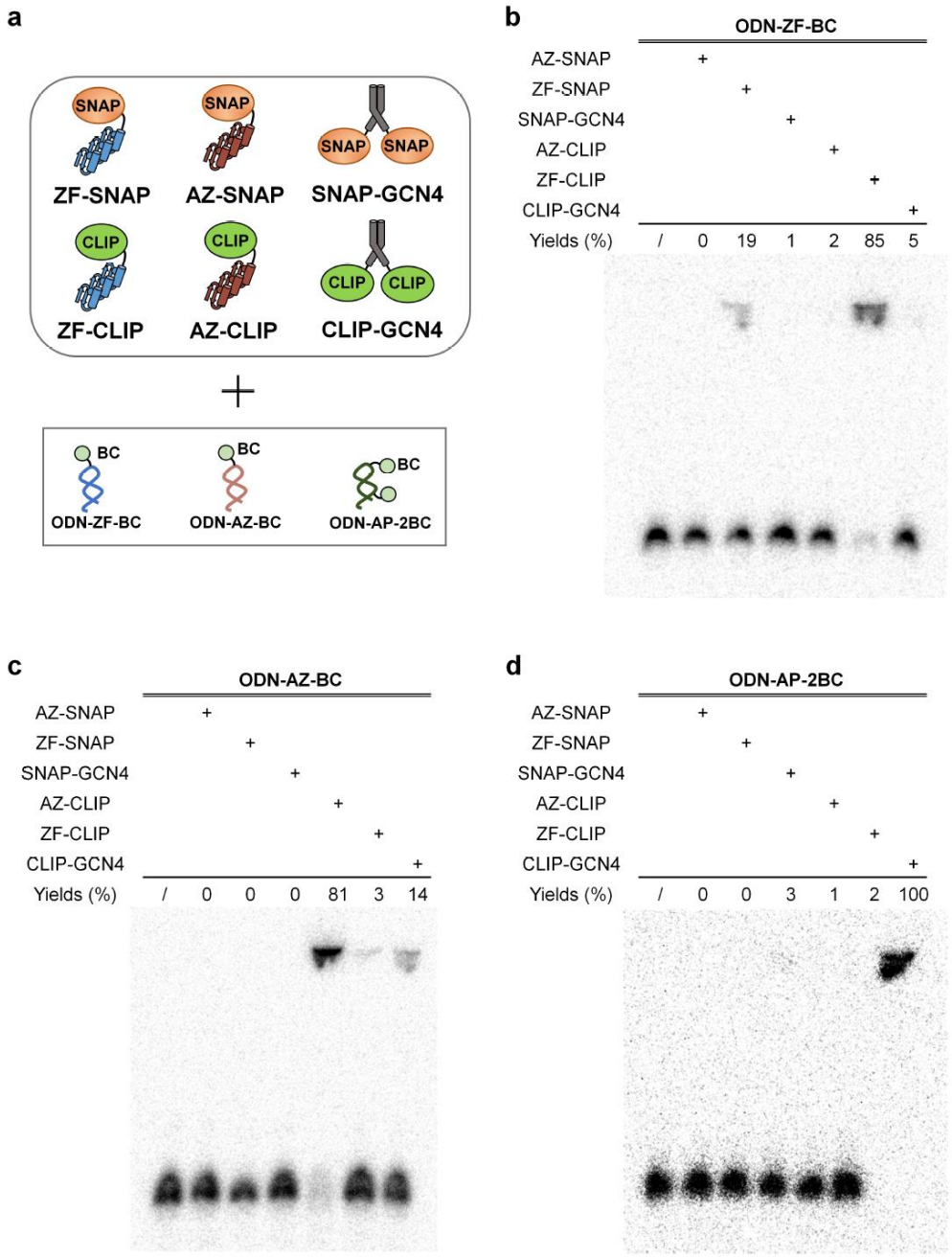


Figure 3.4 Selective crosslink reactions between BC-modified ODNs and 6 types of modular adaptors. (a) An illustration of the optimized modular adaptors and BC-modified DNA binding sequences. 3 MA-SNAPs (80 nM) and 3 MA-CLIPs (80 nM) were incubated for 30 min at room temperature with respective 5′-^{32}P-end-labeled BC-modified ODNs. The PAGE analysis of the reactions are shown in (b) ODN-ZF-BC, (c) ODN-AZ-BC, and (d) ODN-AP-2BC.

3.2.2 Sequence-selective conjugation on DNA scaffolds

Based on the quantitative analyses on the sequence-selectivity of MA-SNAP, a DNA scaffold was used next to highlight the recognition-driven conjugation of MA-SNAP to BI-modified DNA sequences. For this purpose, a DNA origami scaffold containing three cavities was prepared.[2] Each cavity contained a unique BI-modified binding site: ODN-AP-2BI in cavity I, ODN-ZF-BI in cavity II and ODN-AZ-BI in cavity III (Figure 3.5a, Table 3.2). A single type of MA-SNAP (250 nM) was first reacted individually with the DNA scaffold (5 nM) for 30 min. Evaluation of the reaction products by atomic force microscopy (AFM) imaging indicated that each MA-SNAP specifically located in the designated cavity (Figure 3.5). It was also verified from dimensions estimation of AFM image that height change triggered by ZF-SNAP reaction was observed on the well of DNA scaffold marked by ODN-ZF-BI (Figure 3.6). The cross-linking yield by SNAP-GCN4, ZF-SNAP, and AZ-SNAP at the respective sequence-matched sites reached 86% or more, and the yield at unmatched sites were less than 8% (Figures 3.5b-d). Simultaneous reactions by all three MAs on the DNA scaffold provided a co-assembly yield of 85% after 30 min of incubation (Figure 3.5e). These results clearly indicated that the sequence-selective binding of MA-SNAP drove the specific cross-linking reaction of this MA at the target site on DNA scaffold.

Table 3.2 Nucleotide sequences of staple strands containing zif268, AZP4 and GCN4 binding sites with BI modified T^{R} (R = BI : T^{BI}). The GCN4, zif268 and AZP4 binding sites on the staple strands were colored in red, blue and green, respectively.

DNA scaffold	Oligo name	Sequence (from 5′ to 3′)
3 well DNA scaffold (cavity I = AP2BI, cavity II = ZF-BI, cavity III = AZ-BI)	8g-AP-2BI	TAATACTGT^{BI}TCATGAGTCATGAGTTTTCT^{BI}CATGACTCATGAAC CAGGCAAGGCAAAGACATCCAATA
	11g-ZF-BI	AACAGGTCCTTACGCCCACGCGCGTT^{BI}TTCGCGCGTGGGCGTAAGG AACCAGACCGGAAGATTCGAGC
	24d-AZ-BI	GGACAGATCTTATGCCACGTAGCGTT^{BI}TTCGCTACGTGGCATAAGA AATTGTGTCGAAATCTGTATCAT

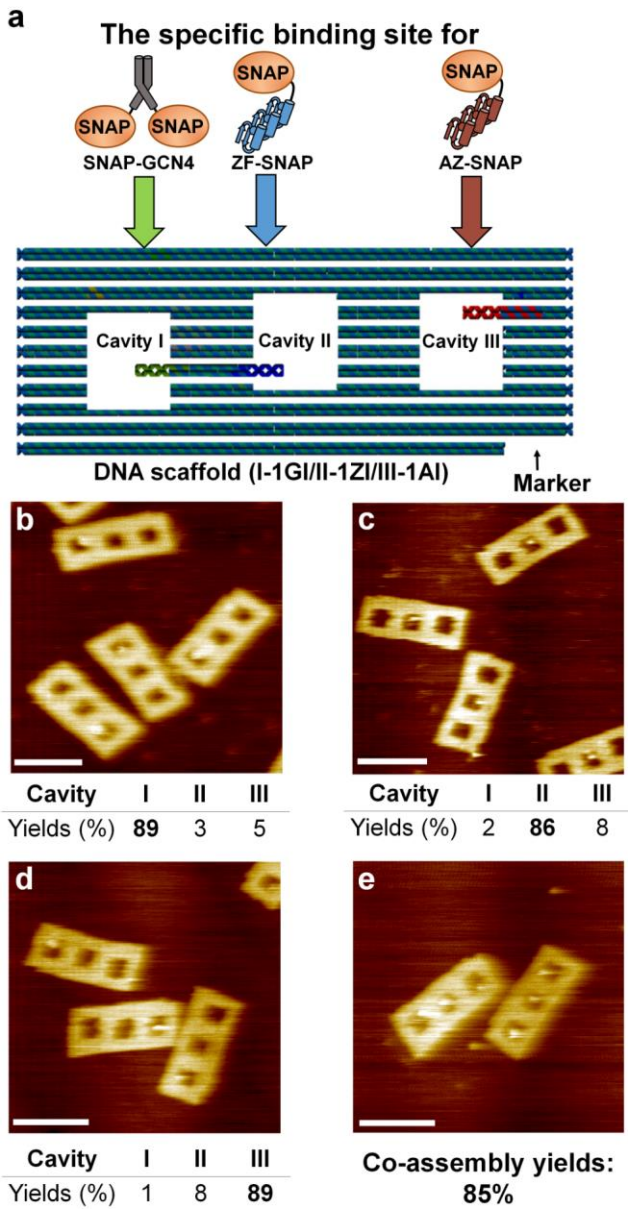


Figure 3.5 Three types of MA-SNAPs site-selectively reacted at the programmed positions on DNA scaffold. (a) Illustration of the DNA scaffold. Each cavity was designed to contain a specific binding site for one of the MA-SNAPs: SNAP-GCN4 in cavity I, ZF-SNAP in cavity II, and AZ-SNAP in cavity III. Representative AFM images of the DNA scaffold reacted with individual MAs: (b) SNAP-GCN4, (c) ZF-SNAP, and (d) AZ-SNAP. (e) A representative AFM image of the DNA scaffold with co-assembly of three types of MA-SNAPs. The reaction mixture was purified by size exclusion chromatography and then analyzed by AFM. Scale bars: 100 nm.

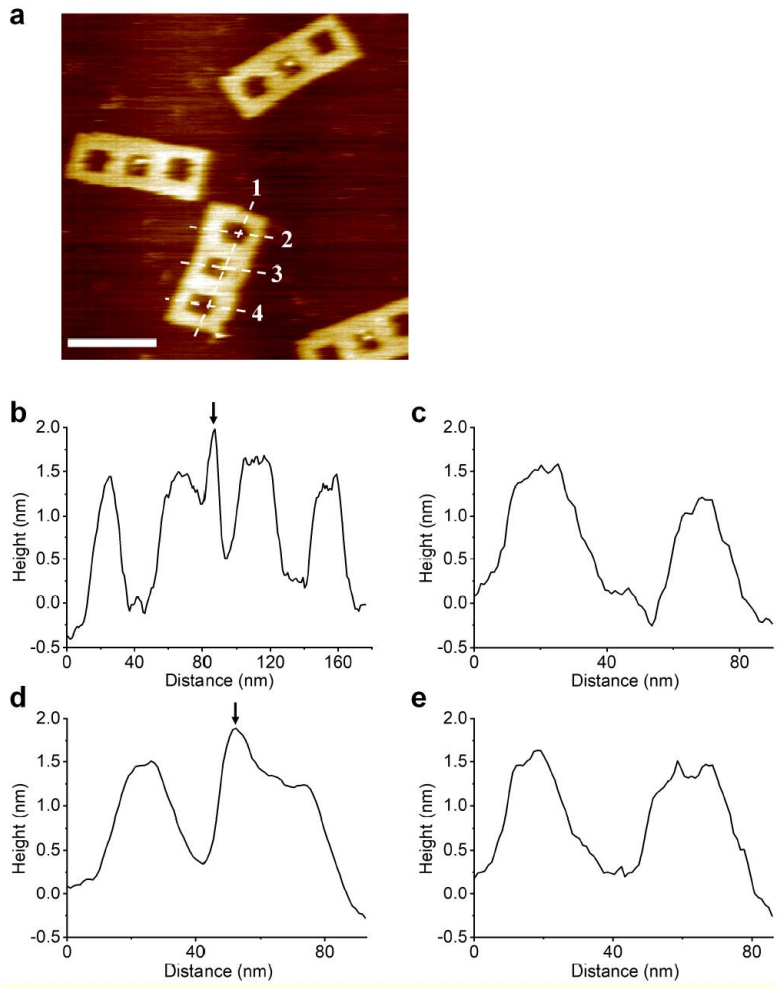


Figure 3.6 Estimation of the dimensions from the AFM image of DNA scaffold in Figure 3.5c. (a) Representative AFM image of the DNA scaffold reacted with ZF-SNAP shown in Figure 3.5c. Four dashed lines marked different positions on the selected DNA scaffold for the height analysis. The height analyses at positions 1, 2, 3 and 4 were shown on (b), (c), (d) and (e), respectively. The black arrows indicate ZF-SNAP selectively assembled on cavity II containing ODN-ZF-BI.

3.2.3 Chemoselective reaction on DNA scaffolds

The cross-reactivity of MA-SNAP and MA-CLIP that share the same DNA-binding protein was further examined to the substrates BI, BC and BG. The same DNA scaffold was prepared in a way to display the same DNA sequence in all three cavities but modified with the above-mentioned three different substrates: ODN-ZF-BI, ODN-ZF-BC and ODN-ZF-BG at cavities I, II and III, respectively (Figure 3.7a, Table 3.3). After the reaction of ZF-SNAP or ZF-CLIP with the DNA scaffold for 30 min, the recorded AFM images revealed that ZF-SNAP was located to cavity III containing BG with higher yield (94%) than cavity I containing BI (81%) and cavity II containing BC (8%) (Figure 3.7b). As expected, ZF-SNAP showed higher reactivity to the BG- than to the BI-modified sequence, with clear discrimination against the BC-modified one. Whereas, the ZF-CLIP reacted at cavity II bearing ODN-ZF-BC in 84% yield and at cavity III bearing ODN-ZF-BG in 16% yield, but hardly any reaction was observed at cavity I bearing ODN-ZF-BI (Figure 3.7c). The time-course reaction of ZF-SNAP with this DNA scaffold clearly showed the preferential modification at cavities I and III by ZF-SNAP (Figure 3.7d). These results indicated that combination of the chemoselectivity of protein-tag with the recognition-driven modification reaction further increased the number of orthogonal sets of MAs.

Table 3.3 Nucleotide sequences of staple strands containing zif268, AZP4 and GCN4 binding sites with BI, BG or BC modified T^{R} (R = BI : T^{BI}, R = BG : T^{BG}, R = BC : T^{BC}). The GCN4, zif268 and AZP4 binding sites were written in red, blue and green, respectively.

DNA scaffold	Oligo name	Sequence (from 5′ to 3′)
3 well DNA scaffold (cavity I = ZF-BI, cavity II = ZF-BC, cavity III = ZF-BG)	3e-ZF-BI	GCTTTGAACTACGCCCACGCGCGTT^{BI}TTCGCGCGTGGGCGTAGAATAAAG AAATTGCGTTGCACGT
	11g-ZF-BC	AACAGGTCCTTACGCCCACGGCGTT^{BC}TTCGCCGTGGGCGTAAGGAACCAG ACCGGAAGATTCGAGC
	19d-ZF-BG	GGGTAGCACTTACGCCCACGGCGTT^{BG}TTCGCCGTGGGCGTAAGATATTCG GTCGCTGACCCACGCA

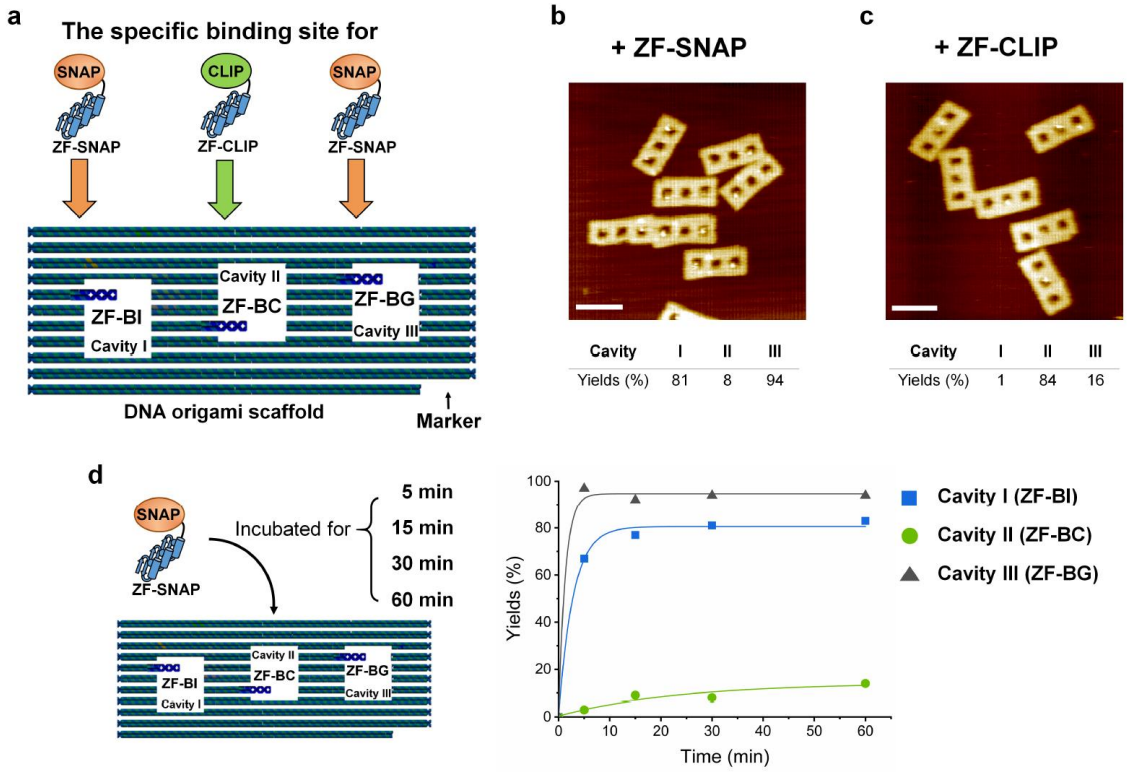


Figure 3.7 Investigation of chemoselectivity of modular adaptors on DNA scaffold. (a) An illustration of DNA scaffold used for the characterization of chemoselectivity. Three single binding sites at cavity I, II, and III of 3-well DNA scaffold were modified by ODN-ZF-BI, ODN-ZF-BC and ODN-ZF-BG, respectively. Binding sites at cavity I and III contain the substrates BI and BG for ZF-SNAP. Binding site at cavity II contains the substrate BC for ZF-CLIP. (b-c) Representative AFM images recorded after the reaction between 5 nM DNA scaffold with 250 nM ZF-SNAP (b) or ZF-CLIP (c). The reactions were carried out for 30 min at room temperature. (d) The origami design (left) and time course reaction (right) of ZF-SNAP (250 nM) with DNA scaffold (5 nM). An aliquot of the reaction mixture was purified by size-exclusion chromatography after 5, 15, 30 and 60 min incubations, then analyzed by AFM. The assembly yield at each cavity was estimated by counting the number of cavities occupied by ZF-SNAP. Scale bars: 100 nm.

3.2.4 Selective assembly of two series of modular adaptors on DNA scaffolds

With two series of sequence-selective MA in hand, chemoselectivity would be a powerful strategy to assemble over 3 POIs on DNA scaffold. As MA-SNAP and MA-CLIP showed very good orthogonality to each other, the selectivity of four types of MAs, including 2 MA-SNAPs and 2 MA-CLIPs, was verified further by both individual and simultaneous reactions. A DNA scaffold with five cavities was prepared to display a single binding site for ZF-SNAP, AZ-SNAP, ZF-CLIP, and AZ-CLIP at cavities I, II, III, and V, respectively (Figure 3.8a, Table 3.4).[12–14] When a single MA was reacted with the DNA scaffold, the designated specific binding site was selectively modified (Figures 3.8b-e). The individual ZF-SNAP, AZ-SNAP, and ZF-CLIP were assembled in over 82% or more, while AZ-CLIP resulted in 77% yield at the respective target sites with less than 7% modification at the unmatched positions. When both ZF-SNAP and ZF-CLIP were simultaneously reacted with the DNA scaffold (Figure 3.8f), modification of cavity I and III reached to similar yield with that of the individual reactions as shown in Figures 3.8b and 3.8d. The result again indicated that the selective binding of zif268 DBD to ODN-ZF was responsible for the selective cross-linking of two types of MAs with BI as the substrate. Further, simultaneous reactions of ZF-SNAP and AZ-SNAP on the DNA scaffold highlighted the contribution of chemoselectivity of SNAP-tag for BI over BC (Figure 3.8g). The observed selective modification at cavities I and II by these MAs proved that the combination of DBD and the pair of SNAP-tag and BI provides specific MAs with sufficiently high conjugation yield.

Table 3.4 Nucleotide sequences of staple strands containing zif268, AZP4 and GCN4 binding sites with BI, BG or BC modified T^R (R = BI : T^{BI}, R = BG : T^{BG}, R = BC : T^{BC}). The GCN4, zif268 and AZP4 binding sites were written in red, blue and green, respectively.

DNA scaffold	Oligo name	Sequence (from 5′ to 3′)
5 well DNA scaffold (cavity I = ZF-BI, cavity II = AZ-BI, cavity III = ZF-BC, cavity IV = empty, cavity V = AZ-BC)	8d-ZF-BI	CAAAATTATGGAAGGGTTAGAACCCTACGCCCACGCGTT^{BI}TTCGCGTGGGCGTAGATTATCAT
	16d-AZ-BI	CGGGAGGTCTATGCCACGTACGTT^{BI}TTCGTACGTGGCATAGTACCGCGCCCAATAGCTCATCGTA
	12h-ZF-BC	AAGGCCGGCTACGCCCACGCGTT^{BC}TTCGCGTGGGCGTAGAAAGACACCACGGAATCATATAAA
	16l-AZ-BC	ACACTGAGCTATGCCACGTACGTT^{BC}TTCGTACGTGGCATAGCAGAACCGCCACCCTCTTAGTACC

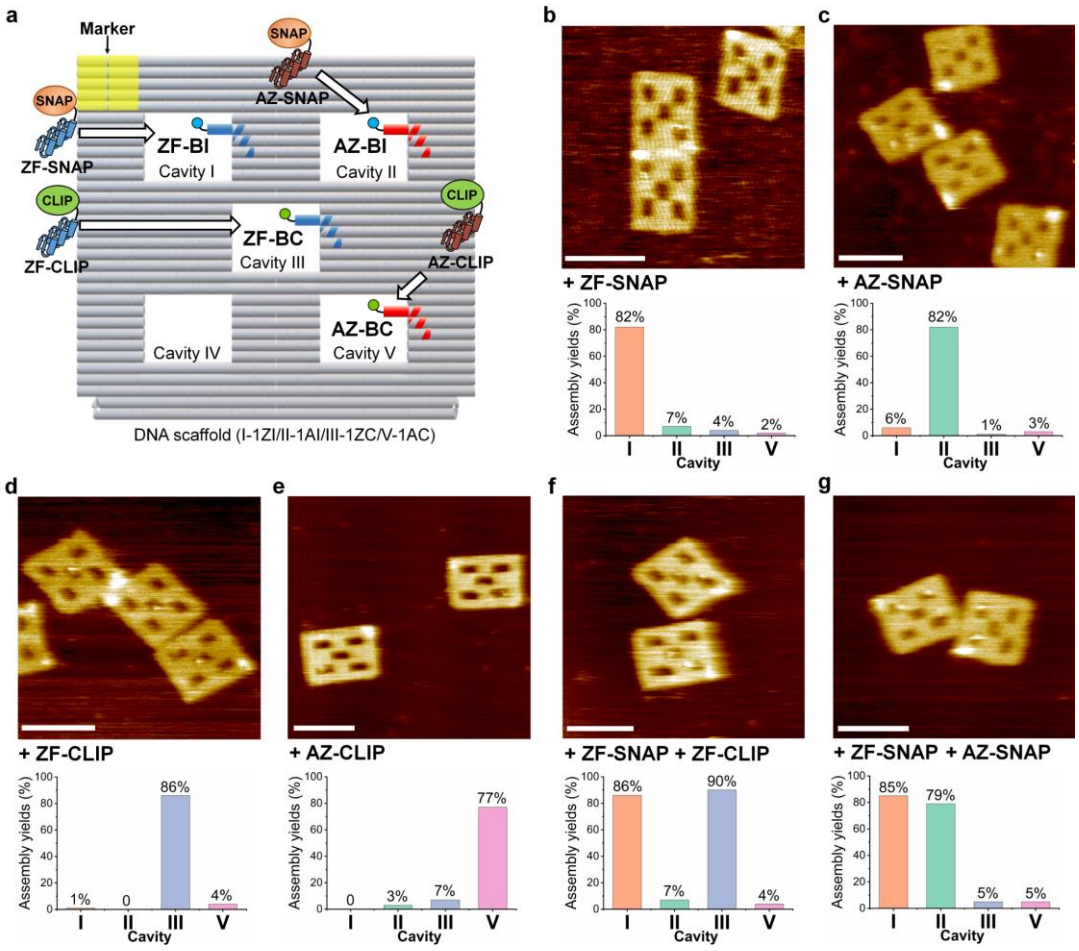


Figure 3.8 Selective assembly of four types of MAs on the DNA scaffold. (a) Illustration of DNA origami scaffold containing five cavities: cavity I contained target sequence for ZF-SNAP modified with BI, cavity II contained target sequence for AZ-SNAP modified with BI, cavity III contained target sequence for ZF-CLIP modified with BC, and cavity V contained target sequence for AZ-CLIP modified with BC. Representative AFM images of the scaffold modified

with individual MAs: (b) ZF-SNAP, (c) AZ-SNAP, (d) ZF-CLIP and (e) AZ-CLIP. Representative AFM images of the co-assembly of ZF-SNAP and ZF-CLIP (f), and ZF-SNAP and AZ-SNAP (g) on the DNA scaffold. The histogram under each AFM image implies the assembly yield of MA in each cavity. The solution of MA (each 250 nM) and DNA scaffold (5 nM) was incubated at room temperature for 30 min. The samples were purified by size exclusion chromatography before AFM imaging. Scale bars: 100 nm.

3.3 Conclusion

Considering that previously MA-CLIP provided us 3 sequence-selective modular adaptors, our newly designed 3 MA-SNAPs by virtue of BI exploited another group of 3 sequence-selective modular adaptors. Fortunately, high discrimination of BI against CLIP-tag robustly restricted the unspecific reaction between BI-modified ODN and MA-CLIP. Therefore, the applicable chemoselectivity between MA-SNAP and MA-CLIP enlightened us to comprehend the chance that using 3 MA-SNAPs and 3 MA-CLIPs together would facilely double the number of unique modular adaptor from 3 to 6. These results elucidated that the available number of specific modular adaptor can be increased by taking advantage of both sequence-selectivity governed by K_D and chemoselectivity performed by substrate-specificity of protein tag. By inheriting the idea, it is adventurous to assume that with n types of specific DNA binding protein in hand, m types of specific protein tag can expand their selectivity to create m×n unique modular adaptors by combining each protein tag with different DNA binding proteins.

3.4.1 Materials and methods

3.4.1 Materials

Custom synthesized staple strand oligonucleotides for the preparation of DNA origami scaffolds were purchased from Sigma-Aldrich (St. Louis, MO), Thermo Fisher

Scientific Inc. (Waltham, MA, USA), Japan Bio Services Co., LTD (Saitama, Japan) and Gene Design Inc. Sephacryl S-400 was purchased from GE Healthcare Japan Corporation (Tokyo, Japan). Ultrafree-MC-DV centrifugal filters were purchased from Merck Millipore (Billerica, MA, USA). A Cosmosil 5C18-MS II column (4.6 i.d. × 150 mm) was purchased from Nacalai Tesque Inc. (Kyoto, Japan). ULTRON VX-ODS packed columns were purchased from Shinwa Chemical Industries (Kyoto, Japan). Reagents for chemical synthesis, HPLC-grade trifluoroacetic acid, gel electrophoresis-grade acrylamide, bis(acrylamide), phenol, and all other chemicals and reagents were purchased from FUJIFILM Wako Pure Chemical Industries, Ltd. (Osaka, Japan), Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) or Nacalai Tesque, Inc. Phosphate buffer (PB) was prepared as 20 mM Na_2HPO_4 and 20 mM NaH_2PO_4. DNA origami buffer (pH 8.0) contained 40 mM Tris-HCl, 20 mM acetic acid, and 12.5 mM $MgCl_2$.

3.4.2 ^{32}P-labelling for characterization of cross-linking reaction

The 5′ end of ODN was ^{32}P-labelled as reported method. To study the covalent bond formation between the substrate and protein-tag in DNA-MA complex, ^{32}P-labelled ODN containing protein-tag substrate (0.5 nM) was incubated with MA in 1× origami buffer (pH 8.0) supplemented with 100 μM mercaptoethanol, 1 mM $ZnCl_2$, 0.02% Tween 20, 200 nM BSA, 100 nM calf thymus DNA and 200 mM NaCl at room temperature. The cross-linking reaction was quenched by adding SDS and formamide to the reaction solution, and the resultant mixture was heated at 80°C for 3 min and rapidly cooled down on ice. The aliquots were analyzed by denaturing PAGE containing 8 M urea, and the gel images were recorded by using Storm 860 Molecular Imager (Amersham).

3.4.3 Preparation of the DNA origami scaffold

A solution (50 ml) was prepared containing M13mp18 (20 nM), substrate-modified ODNs (10 equiv., 200 nM) and staple strands (5 equiv., 100 nM of each strand) in 1× origami buffer (pH 8.0). The solution was heated at 95°C for 1 min, rapidly cooled down to 53°C, incubated for 30 min, and finally rapidly cooled down to 4°C by using a thermal cycler. The sample was then purified by using Sephacryl S-400 (400 ml) Ultrafree-MC-DV centrifugal filters, and 1× origami buffer (pH 8.0, for washing) to get rid of the excess ODNs.

3.4.4 Specific attachment of modular adaptors on DNA scaffold

DNA scaffold (5 nM) was taken in 1× origami buffer (pH 8.0) supplemented with 200 mM NaCl, 100 μM mercaptoethanol, 1 μM $ZnCl_2$, and 0.005% Tween 20, and incubated with MAs (250 nM) for the specified reaction time at room temperature. The reaction was quenched by adding ODN competitor (2 μM) that has binding sequence for the MA while lacking the sequence to bind on DNA scaffold. This ODN competitor can competitively dissociate the non-covalent binding complex between the target ODN on DNA scaffold and MA. The reaction mixture was then purified by using Sephacryl S-400 in Ultrafree-MC-DV centrifugal filter (400 ml) and 1× origami buffer (pH 8.0, for washing).

3.4.5 AFM imaging and statistical analysis

The purified sample was incubated on a freshly cleaved mica surface (ϕ 1.5 mm) for 5 min at room temperature. The surface was then washed for three times to remove the unbound excess sample by using 1× origami buffer (pH 8.0). The sample was scanned under tapping mode using a Sample-Scanning-Nano Explorer AFM system

(SS-NEX, RIBM Co., Ltd., Tsukuba, Japan) with a silicon nitride cantilever (resonant frequency: 1.5 MHz, spring constant: 0.1 N/m, Olympus, Tokyo, Japan). Several images were acquired from different regions on the mica surface. The yield of specific and non-specific binding of MAs, and the coassembly yield were estimated by counting the MAs bound to the well-folded DNA scaffold. In each case, over 100 DNA scaffold were counted for calculating the assembly yield (Table 3.5).

Table 3.5 Total number of counted DNA origami and yields of the DNA scaffold assembled with the modular adaptors analysed by AFM.

DNA scaffold	Modular adaptor derivatives	Number of well-formed DNA scaffold	Yields of the modified cavities			Coassembly yield	AFM image
			Cavity I	Cavity II	Cavity III		
3 well DNA scaffold (cavity I = AP-2BI, cavity II = ZF-BI, cavity III = AZ-BI)	SNAP-GCN4	350	**89%**	**3%**	**5%**	-	Figure 3.5b
	ZF-SNAP	260	**2%**	**86%**	**8%**	-	Figure 3.5c
	AZ-SNAP	223	**1%**	**8%**	**89%**	-	Figure 3.5d
	SNAP-GCN4, ZF-SNAP and AZ-SNAP	213	**96%**	**87%**	**91%**	**85%**	Figure 3.5e
3 well DNA scaffold (cavity I = ZF-BI, cavity II = ZF-BC, cavity III = ZF-BG)	ZF-SNAP (Incubated for 5 min)	230	**67%**	**3%**	**97%**	-	Figure 3.7d
	ZF-SNAP (Incubated for 15 min)	175	**77%**	**9%**	**92%**	-	Figure 3.7d
	ZF-SNAP (Incubated for 30 min)	316	**81%**	**8%**	**94%**	-	Figure 3.7b and 3.7d
	ZF-SNAP (Incubated for 60 min)	107	**83%**	**14%**	**94%**	-	Figure 3.7d
	ZF-CLIP	368	**1%**	**84%**	**16%**	-	Figure 3.7c

DNA scaffold	Modular adaptor derivatives	Number of well-formed DNA scaffold	Yields of the modified cavities				AFM image
			Cavity I	Cavity II	Cavity III	Cavity V	
5 well DNA scaffold (cavity I = ZF-BI, cavity II = AZ-BI, cavity III = ZF-BC, cavity IV = empty, cavity V = AZ-BC)	ZF-SNAP	253	**82%**	**7%**	**4%**	**2%**	Figure 3.8b
	AZ-SNAP	250	**6%**	**82%**	**1%**	**3%**	Figure 3.8c
	ZF-CLIP	208	**1%**	**0**	**86%**	**4%**	Figure 3.8d
	AZ-CLIP	230	**0**	**3%**	**7%**	**77%**	Figure 3.8e
	ZF-SNAP ZF-CLIP	196	**86%**	**7%**	**90%**	**4%**	Figure 3.8f
	ZF-SNAP AZ-SNAP	148	**85%**	**79%**	**5%**	**5%**	Figure 3.8g

3.5 References

(1) Nakata, E.; Dinh, H.; Ngo, T. A.; Saimura, M.; Morii, T. A Modular Zinc Finger Adaptor Accelerates the Covalent Linkage of Proteins at Specific Locations on DNA Nanoscaffolds. *Chem. Commun.* **2015**, *51*, 1016–1019.

(2) Ngo, T. A.; Nakata, E.; Saimura, M.; Morii, T. Spatially Organized Enzymes Drive Cofactor-Coupled Cascade Reactions. *J. Am. Chem. Soc.* **2016**, *138*, 3012–3021.

(3) Nakata, E.; Dinh, H.; Nguyen, T. M.; Morii, T. DNA Binding Adaptors to Assemble Proteins of Interest on DNA Scaffold. *Methods Enzymol.*, **2019**, *617*, 287–322.

(4) Gautier, A.; Juillerat, A.; Heinis, C.; Corrêa, I. R.; Kindermann, M.; Beaufils, F.; Johnsson, K. An Engineered Protein Tag for Multiprotein Labeling in Living Cells. *Chem. Biol.* **2008**, *15*, 128–136.

(5) Nguyen, T. M.; Nakata, E.; Zhang, Z.; Saimura, M.; Dinh, H.; Morii, T. Rational Design of a DNA Sequence-Specific Modular Protein Tag by Tuning the Alkylation Kinetics. *Chem. Sci.* **2019**, *10*, 9315–9325.

(6) Nguyen, T. M.; Nakata, E.; Saimura, M.; Dinh, H.; Morii, T. Design of Modular Protein Tags for Orthogonal Covalent Bond Formation at Specific DNA Sequences. *J. Am. Chem. Soc.* **2017**, *139*, 8487–8496.

(7) Keppler, A.; Gendreizig, S.; Gronemeyer, T.; Pick, H.; Vogel, H.; Johnsson, K. A General Method for the Covalent Labeling of Fusion Proteins with Small Molecules in Vivo. *Nat. Biotechnol.* **2003**, *21*, 86–89.

(8) Pavletich, N. P.; Pabo, C. Zinc Finger-DNA Recognition : Crystal Structure of a Zif268-DNA Complex at 2 . 1. *Science* **1991**, *252*, 809–818.

(9) Sera, T.; Uranga, C. Rational Design of Artificial Zinc-Finger Proteins Using a Nondegenerate Recognition Code Table. *Biochemistry* **2002**, *41*, 7074–7081.

(10) Hope, I. A.; Mahadevan, S.; Struhl, K. Structural and Functional Characterization of the Short Acidic Transcriptional Activation Region of Yeast GCN4 Protein. *Nature* **1988**, *333*, 635–640.

(11) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; et al. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3*, 373–382.

(12) Dinh, H.; Nakata, E.; Mutsuda-Zapater, K.; Saimura, M.; Kinoshita, M.; Morii, T. Enhanced Enzymatic Activity Exerted by a Packed Assembly of a Single Type of Enzyme. *Chem. Sci.* **2020**, *11*, 9088–9100.

(13) Nakata, E.; Liew, F. F.; Uwatoko, C.; Kiyonaka, S.; Mori, Y.; Katsuda, Y.; Endo, M.; Sugiyama, H.; Morii, T. Zinc-Finger Proteins for Site-Specific Protein Positioning on DNA-Origami Structures. *Angew. Chemie - Int. Ed.* **2012**, *51*, 2421–2424.

(14) Ngo, T. A.; Nakata, E.; Saimura, M.; Kodaki, T.; Morii, T. A Protein Adaptor to Locate a Functional Protein Dimer on Molecular Switchboard. *Methods* **2014**, *67*, 142–150.

CHAPTER 4

Functional molecules with a DNA handle for application on DNA scaffold

4.1 Introduction

As introduced in Chapter 2 and Chapter 3, protein of interest (POI) can be localized on DNA scaffold through recognition-driven conjugation mediated by modular adaptor.[1–4] Beside protein, functional molecules, such as fluorophore,[5] peptide,[6] glycan [7] and lipids,[8] can also be site-specifically modified on DNA scaffold via a conjugated DNA handle. DNA handle for the complementary DNA base-pairing enables an excellent class of precise recognition and efficient hybridization, contributing to the remarkable specificity and affinity for unlimited molecular design.[9,10] Hence, DNA handle have been very essential to be introduced on the target molecules through conjugation reaction, which could impart advanced programmability and unprecedented addressability to the protein for extending its potential applicability.[11–13]

Cathepsin B (CtB), a major lysosomal cysteine protease, plays important role in protein degradation and maintenance of cellular homeostasis in lysosome.[14–16] The elevated activity of CtB has been uncovered in various cancers, particularly in aggressive cancers, to be regarded as a promising biomarker for tumor-specific prodrug design.[17,18] For detecting active CtB, various types of fluorescence-based probes have been designed, such as fluorogenic probe [19,20] or FRET probe.[21–24] However, in order to realize signal amplification and more diverse signal output for clinical usage, CtB sensor are desirable to be modified on biomacromolecular assemblies simultaneously with other sensors for pH, metabolites or other enzymes.[25–27] Given that DNA scaffold showed robust programmability for loading multi-types of functional molecules, I conjugated an DNA sequence as a handle on a CtB sensor (CtBsub-ODN1) (Figure 4.1). In the meanwhile, a sensor for cathepsin D (CtD) was also coupled with another DNA sequence (CtDsub-ODN2) to mimic the dual-signal outputs coupled with the

CtBsub-ODN1 (Figure 4.1). Because cathepsin D is an important aspartic protease inside the lysosome to show complex relationship with CtB (Figure 4.1b).[28,29] CtD is processed to be matured in cells by subsequent proteolytic cleavage involving CTSB and other cysteine proteases and is also related to the activation of CtB [26,29–32] The signal output from CtD sensor is assumed to indicate the potential existence of active cathepsin B. Besides, the design of both sensors are based on FRET which contains a donor fluorophore and a quencher. Two orthogonal FRET pairs were used on the substrates of CtB and CtD, respectively, for detecting the signal from both CtB and CtD independently (Figure 4.1). Furthermore, CtBsub-ODN1 was further assembly on DNA scaffold through DNA hybridization to detect the activity of CtB. Additionally, I modified two types of SNARF fluorophores with maleimido group, which could be used to conjugate with DNA. It was expected to be further assembled on DNA scaffold as pH sensor coupled with cathepsin sensors as a multiple signal-response sensor.[33–35]

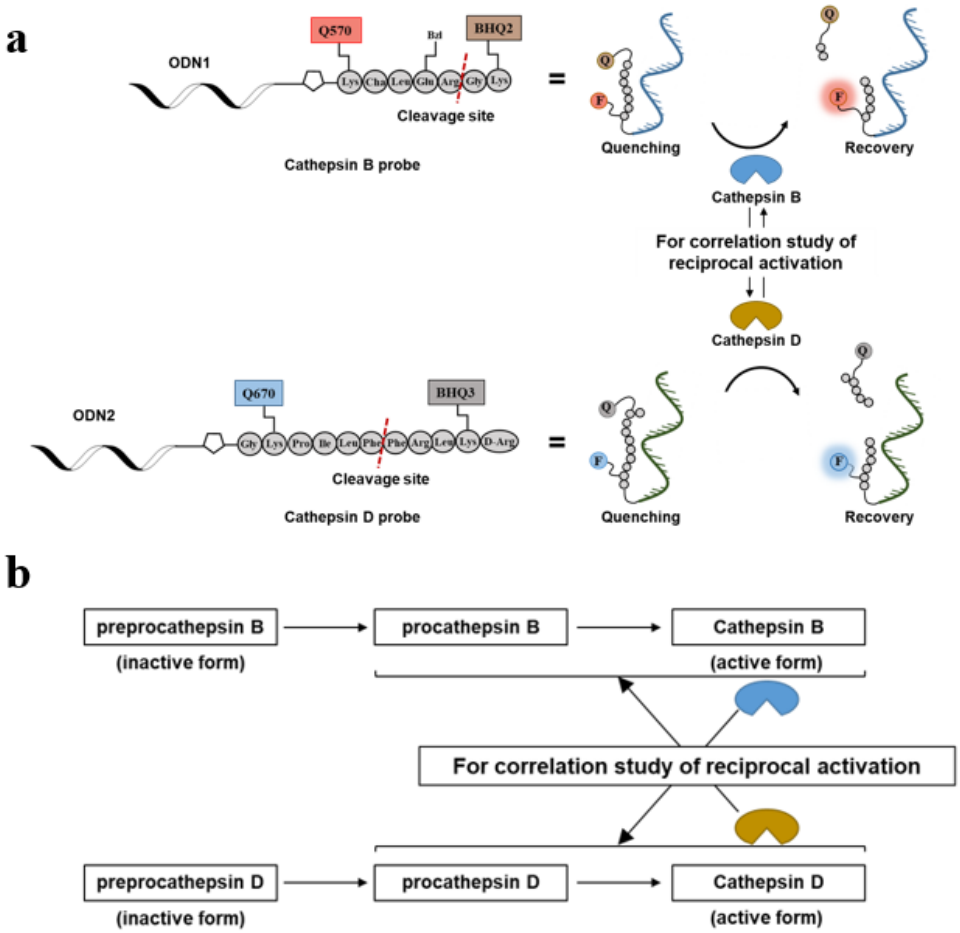
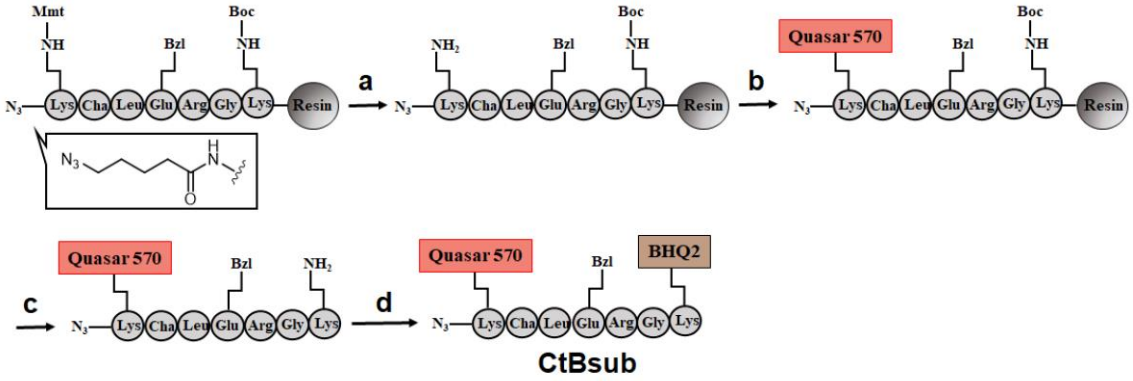


Figure 4.1 Scheme illustration for enzymatic reaction of cathepsin with DNA-conjugated sensor, (a) Probes for cathepsin B (CtBsub-ODN1) and Cathepsin D (CtDsub-ODN2) containing distinguishable FRET pairs, Q570-BHQ2 and Q670-BHQ3, respectively. The emission of fluorophore was efficiently quenched before the reaction, while it was recovered after the cleavage of peptide bond between fluorophore and quencher by the target cathepsin. (b) The detail of the proposed correlation of reciprocal activation of cathepsin B and cathepsin D.

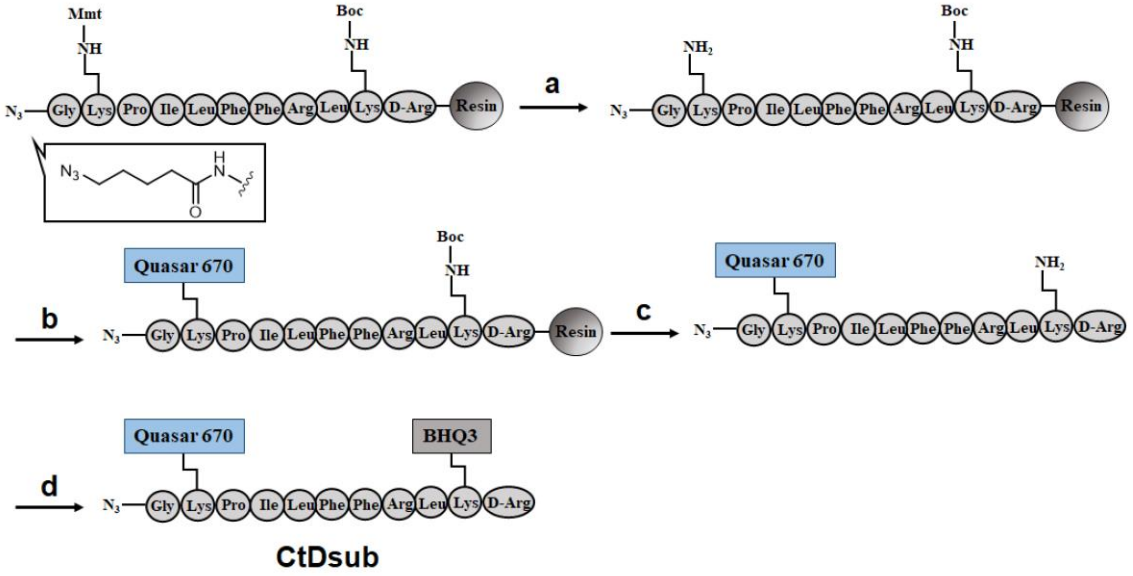
4.2 Results and discussion

4.2.1 Design and synthesis of FRET-based sensors for cathepsin B and D

FRET-based cathepsin probes for orthogonally detecting CtB and CtD were synthesized as shown in Scheme 4.1 and 4.2. Two cathepsin probes were designed to specifically monitor the activity of CtB and CtD *via* the change of the FRET efficiency of each distinguishable FRET type of fluorophore-quencher pair, that is, Quasar 570 (Q570)-BHQ2 and Quasar 670 (Q670)-BHQ3 (Figure 4.2). These combinations ensured the availability of dual-signal outputs and stable readouts. The substrates of CtB (CtBsub) and CtD (CtDsub) were designed by referring the reported peptide sequence.[36–38] Two lysines were added to both terminals of CtBsub and CtDsub for the modification of Q570 and BHQ2, and Q670 and BHQ3, respectively. The newly designed substrates were prepared via solid phase peptide synthesis (SPPS) and featured a terminal azido group in order to conjugate with the DNA handle. Stepwise modification of the fluorophore and quencher on the residue of two lysines were conducted by the selective deprotection of the amino residue of lysine. The fluorophore was coupled on the N-terminal of the substrate, which performed on the solid phase after the selective deprotection of a lysine protected by monomethoxytrityl (Mmt), to detect the fluorescent signal from the side of DNA handle. Quencher was then coupled on the lysine close to the C-terminal after the peptide excision from the resin. After the purification of these peptides by reverse phase HPLC, the cathepsin substrates containing azido group (CtBsub and CtDsub) were prepared. Both substrates were characterized by HPLC, UV and MS (Figure 4.3 and Figure 4.4).



Scheme 4.1 Synthetic scheme of CtBsub. Reagents and conditions: (a) 2% TFA in DCM (b) Q570-COOH (1.5 eq), HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF (c) 2.5% H_2O in TFA (d) BHQ2-NHS (1.2 eq), and TEA (2 eq.) in DMF.



Scheme 4.2 Synthetic scheme of CtDsub. Reagents and conditions: (a) 2% TFA in DCM (b) Q670-COOH (1.5 eq), HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF (c) 2.5% H_2O in TFA (d) BHQ3-NHS (1.2 eq), and TEA (2 eq.) in DMF.

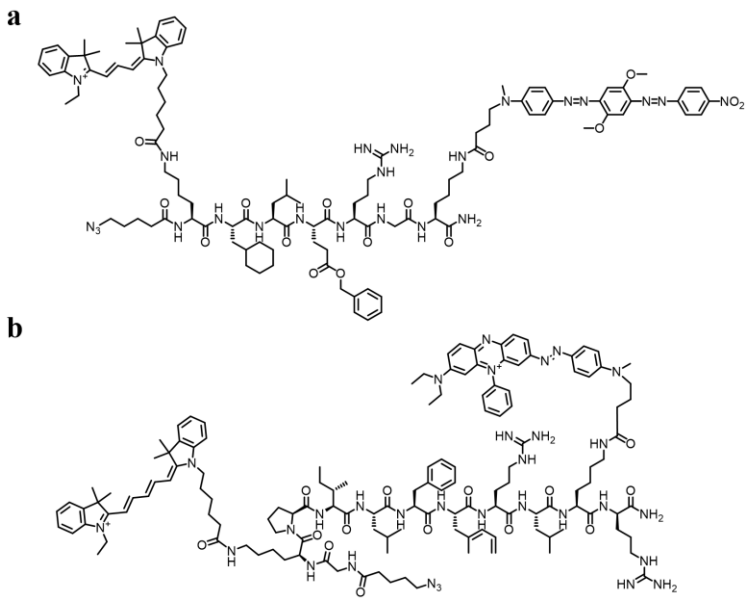


Figure 4.2 Chemical structures of (a) CtBsub and (b) CtDsub

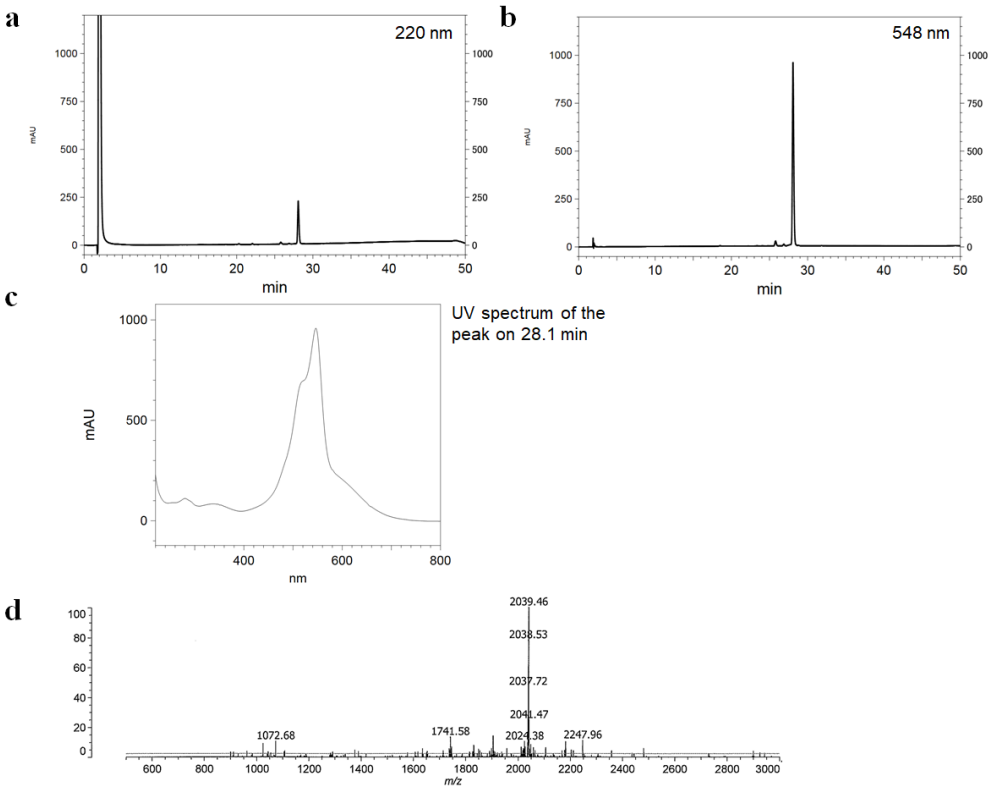


Figure 4.3 HPLC analysis of substrate of cathepsin B (CtBsub) modified by Quasar 570 and BHQ2. (a-b) HPLC profiles of the compound detected by UV at (a) 220 nm and (b) 548 nm. (c) UV-vis spectrum of a peak eluted at 28.1 min. HPLC analysis were conducted on a ULTRON

VX-ODS column (4.6 × 150 mm, eluent A: 0.05% TFA containing H_2O, eluent B : 0.05% TFA in 80% acetonitrile). Sample was eluted with the gradient of eluent B increased from 35% to 95% in 30 min at a flow rate of 1.0 mL·min^{-1}. (d) MS spectra of CtBsub modified by Quasar 570 and BHQ2. *m*/*z* 2039.5, for $[M+H]^+$, MALDI-TOF-MS positive linear mode, CHCA matrix (calcd. 2038.2).

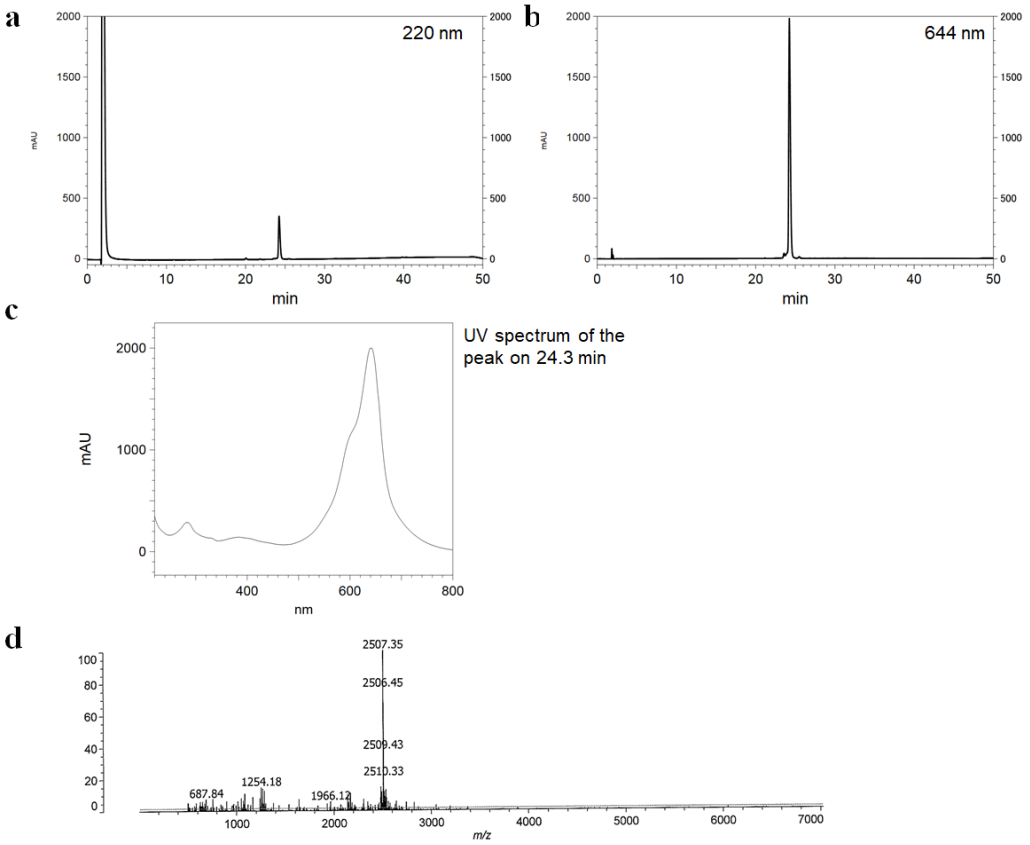
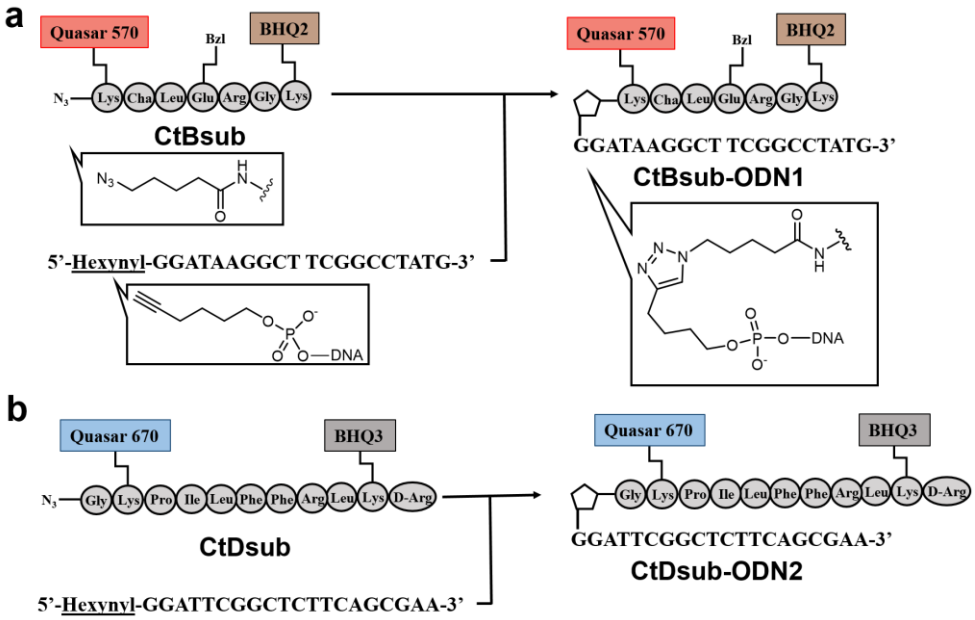


Figure 4.4 HPLC analysis of substrate of cathepsin D (CtDsub) modified by Quasar 670 and BHQ3. (a-b) HPLC profiles of the compound detected by UV at (a) 220 nm and (b) 644 nm. (c) UV-vis spectrum of a peak eluted at 24.3 min. HPLC analysis were conducted on a ULTRON VX-ODS column (4.6 × 150 mm, eluent A: 0.05% TFA containing H_2O, eluent B : 0.05% TFA in 80% acetonitrile). Sample was eluted with the gradient of eluent B increased from 35% to 95% in 30 min at a flow rate of 1.0 mL·min^{-1}. (d) MS spectra of CtDsub modified by Quasar 670 and BHQ3. *m*/*z* 2507.4, for $[M+H]^+$, CHCA matrix, MALDI-TOF-MS positive linear mode, CHCA matrix (calcd. 2506.5).

4.2.2 Conjugation of cathepsin sensor with a DNA handle

Conjugation reaction of azido-modified cathepsin substrates and 5'-hexynyl-modified DNA was carried out *via* a copper-catalyzed azide-alkyne cycloaddition to generate two specific cathepsin probes with DNA handle, CtBsub-ODN1 and CtDsub-ODN2 (Scheme 4.3). Because cathepsin substrate contained several hydrophobic amino acids and were modified by bulky hydrophobic fluorophore and quencher, organic solvents were used for dissolving peptide-based species. And DNA and other reagents for CuAAC were dissolved in the buffer and mixed with peptide solution for processing conjugation. The DNA handle here, ODN-1 or ODN-2, enabled the further assembly of the cathepsin probe onto the scaffolds through DNA hybridization. The isolated CtBsub-ODN1 and CtDsub-ODN2 were identified by MALDI-TOF-MS, HPLC and UV spectrum (Figure 4.5 and 4.6).



Scheme 4.3 Synthesis scheme of CuAAC conjugation for producing CtBsub-ODN1 (a) and CtDsub-ODN2 (b).

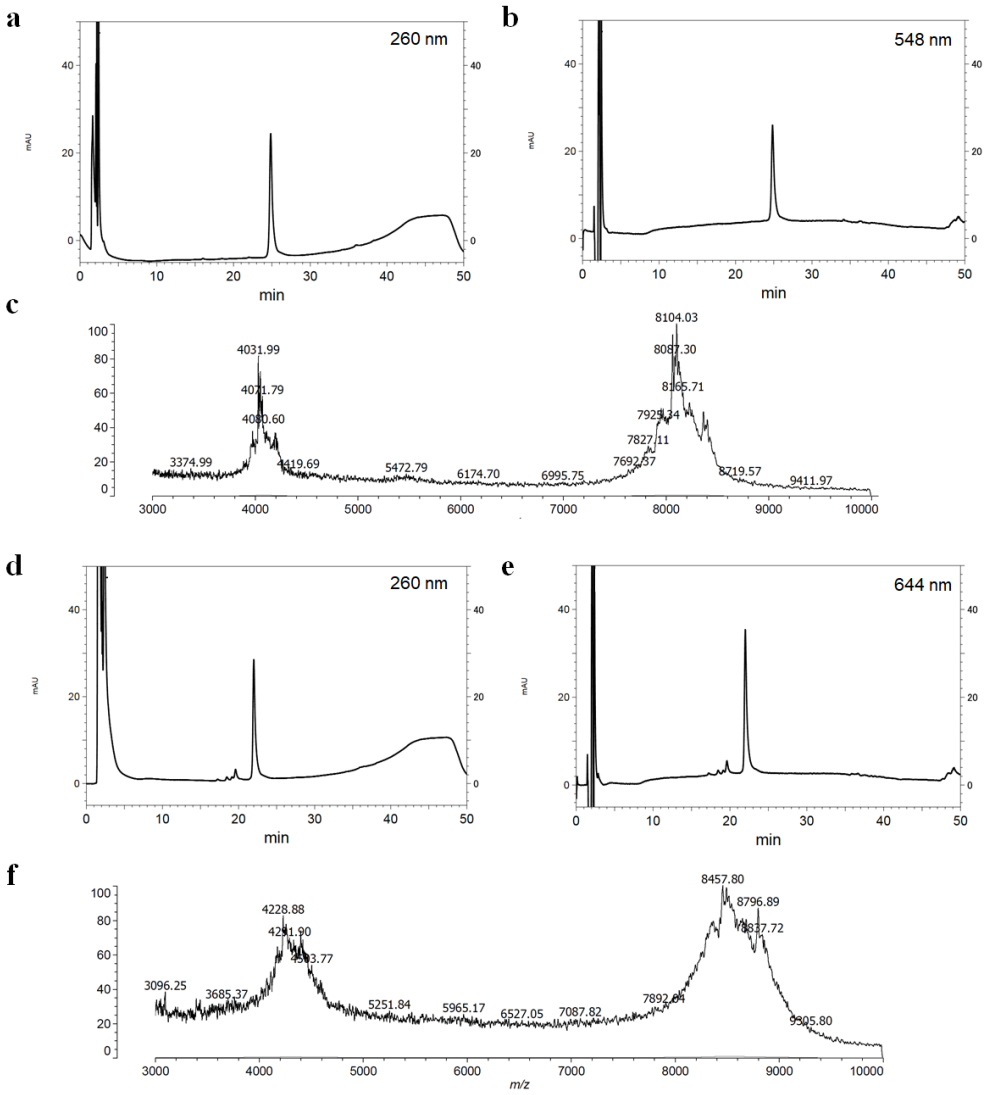


Figure 4.5 Characterization of CtBsub-ODN1 and CtDsub-ODN2. (a-b) HPLC profiles of CtBsub-ODN1 detected by UV at (a) 260 nm and (b) 548 nm. Sample was eluted with the gradient of eluent B increased from 30% to 90% in 30 min at a flow rate of 1.0 mL·min^{-1} on a COSMOSIL 5C_{18}-MS-II column (4.6 × 150 mm, eluent A: 0.1 M TEAA containing H_2O, eluent B : 0.1 M TEAA in 80% acetonitrile). (c) MS spectra of CtBsub-ODN1, *m*/*z* 8104.03 (positive linear mode), matrix THAP and DAHC (calcd. 8210.2). (d-e) HPLC profiles of CtDsub-ODN2 detected by UV at (d) 260 nm and (e) 644 nm. (f) MS spectra of CtBsub-ODN1. *m*/*z* 8457.80 for the conjugates, matrix THAP and DAHC, MALDI-TOF-MS positive linear mode (calcd. 8639.5).

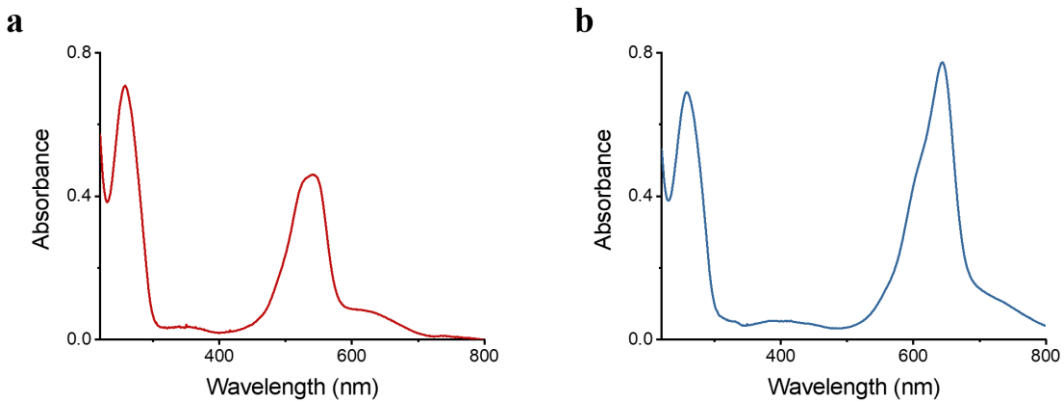


Figure 4.6 UV spectrum related to CtBsub-ODN1 and CtDsub-ODN2. (a-b) UV spectrum of (a) CtBsub-ODN1 and (b) CtDsub-ODN2. The concentration of each was 3 μM, measured in 0.1 M TEAA buffer by SHIMADZU UV-2550 spectrophotometer.

4.2.3 Detection of cathepsin *in vitro* by means of FRET-based cathepsin sensor with the DNA handle

I initially compared the fluorescence enhancement of the cathepsin probe with or without DNA handle by enzyme cleavage. In the presence of activated CtB, , a strong increase of CtB-ODN1 was able to be observed in donor fluorescence, indicating cleavage of the cathepsin sensors and loss of FRET (Figure 4.7c). On the other hand, CtBsub, without the DNA handle, showed less fluorescent signal than CtBsub-ODN1 at the quenched state, owing to its poor solubility (Figure 4.7a). And there is no signal increment from CtBsub to be observed after treated by CtB. CtDsub can cleavage by CtD, but show less fluorescence enhancement than CtDsub-ODN1 (Figure 4.7b and 4.7d). In the reaction of CtDsub-ODN2 with CtD, the increase in the emission of Q670 was observed with much faster kinetics than that of CtDsub (Figure 4.7f). These results indicated ODN tethered cathepsin substrate, especially CtBsub-ODN1, showed drastic improvement of the reactivity to the target cathepsin in the physiological aqueous media. The conversion yields were determined by HPLC as 37% for CtBsub-ODN1 with CtB

incubated for 12 hours at 37 ˚C, and 94% for CtDsub-ODN2 with CtD after the reaction for 3 hours at 37 ˚C (Figure 4.8).

The orthogonality in the cathepsin reaction of CtBsub-ODN1 and CtDsub-ODN2 was verified by treating them with the unmatched cathepsin. The emission intensities of probes treated with the matched cathepsin, that is, CtBsub-ODN1 with CtB (Figure 4.9a) and CtDsub-ODN2 with CtD (Figure 4.9c), showed 13 fold and 4 fold enhancement, respectively. Both the probes showed much lower signal increase by treatment with the unmatched cathepsin (Figure 4.9). Especially, CtDsub-ODN2 showed little or no signal increase in the presence of CtB (Figure 4.9c and 4.9d). CtBsub-ODN1 treated with CtD showed a slight and slow increase of the emission of Q570 (Figure 4.9a and 4.9b). These results indicate that CtBsub-ODN1 and CtDsub-ODN2 orthogonally reported the enzymatic activity of CtB and CtD.

To investigate the sensitivity of DNA-conjugated cathepsin probe, DNA-tethered cathepsin probe with different concentration of its target enzyme were performed. For both cases of CtBsub-ODN1 and CtDsub-ODN2, the increase of fluorescent signal could be observed along with the increasing concentration of target enzyme (Figure 4.10a and 4.10b). The fact showed the good response and sensitivity of DNA-conjugated cathepsin probe to the concentration of target enzyme. The parameter k_{cat}/K_m was measured by reacting different concentration of CtBsub-ODN1 and CtDsub-ODN2 with their target enzyme. After plotting the initial velocity in a dose-dependent mannar, k_{cat}/K_m of CtBsub-ODN1 with CtB was determined as $(2.6 \pm 0.1) \times 10^3$ $M^{-1}s^{-1}$, and CtDsub-ODN2 with CtD was $(1.7 \pm 0.1) \times 10^5$ $M^{-1}s^{-1}$ (Figure 4.11).

As CtB and CtD are both involved in the processing and regulation towards each other, the dual-signal system targeting both CtB and CtD is essential for validating the investigation to active cathepsin B in proteolytic network. For this purpose,

CtBsub-ODN1 and CtDsub-ODN2 with same concentration were mixed in order to simultaneously detect CtB and CtD. The mixture was treated with CtB alone, CtD alone and mixture of CtB and CtD, respectively. The enzymatic reaction was monitored under different wavelength at the same time (Figure 4.10c and 4.10d). In the case of observing Q570 on CtBsub-ODN1, the significant signal increase triggered by CtB can be found to be differentiable to the signal change of the group with CtD. The fluorescent curve of adding both CtB and CtD showed the similar tendency as the total sum of fluorescent curves from the groups with CtB alone and CtD alone. This result showed the potency of CtBsub-ODN1 for probing CtB. However, it also indicated the case that the substrate can also be cleavaged by other protease, even the CtB substrate was reported to be specific against other human cysteine cathepsins. Therefore, multiple signal detection, like CtDsub-ODN2 here, are necessary for contributing to the reliability of the signal outputs. Under the wavelength for Q670, only the signal increase on the groups involving addition of CtD can be observed, which verified the truth that CtD did exist in the system and contributed to the signal increase under Q570 observation. As a result, CtBsub-ODN1 and CtDsub-ODN2 were able to be used as a dual-response system for sensing CtB and CtD. And the facts elucidated the importance of multiple response for reliable detection of CtB.

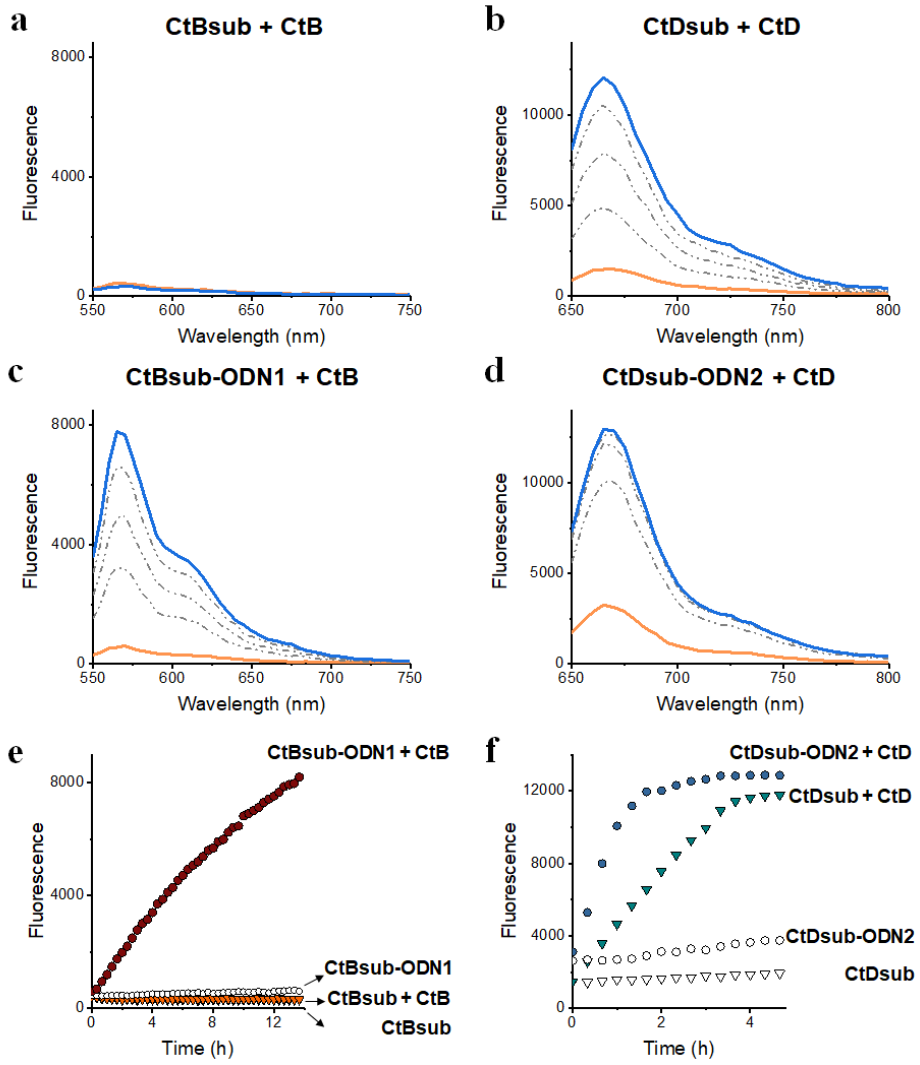


Figure 4.7 Time-dependent emission spectral changes of cathepsin probes by reactions with cathepsins. (a) CtBsub and (b) CtDsub treated with CtB for 13 hours and CtD for 4 hours, respectively. (c) CtBsub-ODN1 and (d) CtDsub-ODN2 treated with CtB for 13 hours and CtD for 4 hours, respectively. Emission spectra before and after the enzymatic reaction are shown in solid orange lines and solid blue line, respectively. After addition of CtB or CtD, the emission spectra was recorded for every 3 hours in (a) and (c) and for every 1 hour in (b) and (d) and illustrated in gray dash line. (e and f) Time-course changes of emission intensity (e) at 570 nm for CtBsub or CtB-ODN1 with or without CtB (excited at 520 nm) and (f) at 670 nm for CtDsub or CtD-ODN2 with or without CtD (excited at 620 nm). (e) CtBsub with CtB (filled triangle), CtBsub without cathepsin (open triangle), CtBsub-ODN1 with CtB (filled circles) and

CtBsub-ODN1 without cathepsin (open circles) (f) CtDsub with CtD (filled triangle), CtDsub without cathepsin (open triangle), CtDsub-ODN1 with CtD (filled circles) and CtDsub-ODN1 without cathepsin (open circles). The concentration of probes and cathepsins were 200 nM and 5 nM, respectively.

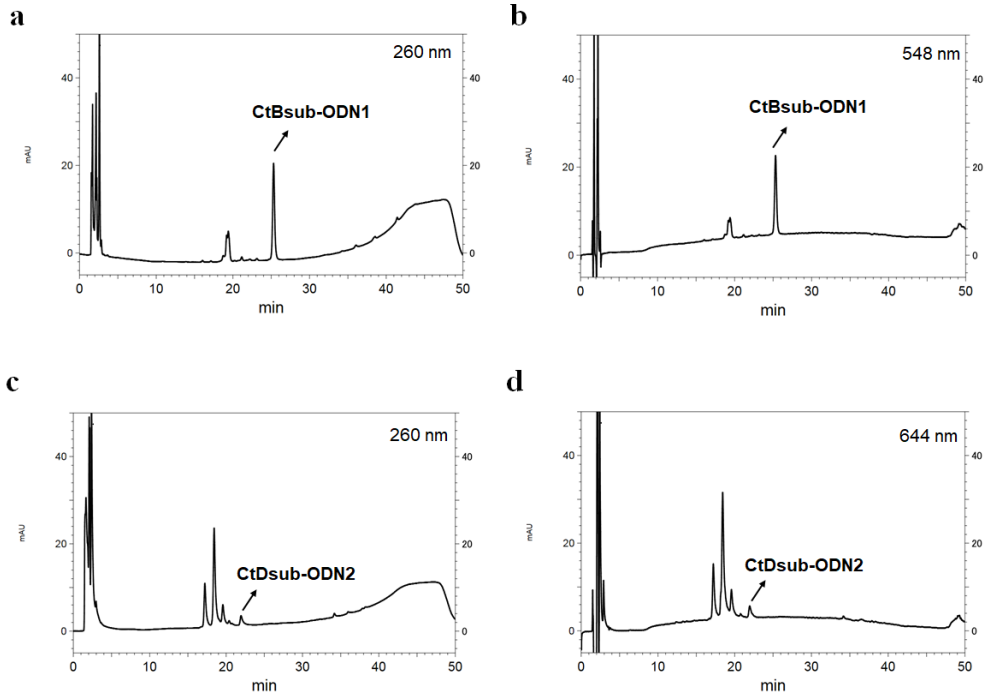


Figure 4.8 Cleavage yield of CtBsub-ODN1 and CtDsub-ODN2 treated by their target enzyme. (a-b) HPLC profiles of CtBsub-ODN1 (200 nM) treated by CtB (5 nM) was detected by UV at (a) 260 nm and (b) 548 nm. The reaction solution was incubated at 37°C for 12 hours and then checked by HPLC. (c-d) HPLC profiles of CtDsub-ODN2 (200 nM) treated by CtD (5 nM) was detected by UV at (a) 260 nm and (b) 644 nm. The reaction solution was incubated at 37°C for 3 hours. The conversion yield was calculated according to the integration of peaks at 260 nm. The unreacted CtBsub-ODN1 or CtDsub-ODN2 can be distinguished by referring to the retention time indicated from Figure 4.5. HPLC analysis were conducted on a COSMOSIL $5C_{18}$-MS-II column (4.6 × 150 mm, eluent A: 0.1 M TEAA containing H_2O, eluent B: 0.1 M TEAA in 80% acetonitrile). Sample was eluted with the gradient of eluent B increased from 30% to 90% in 30 min at a flow rate of 1.0 mL·min^{-1}.

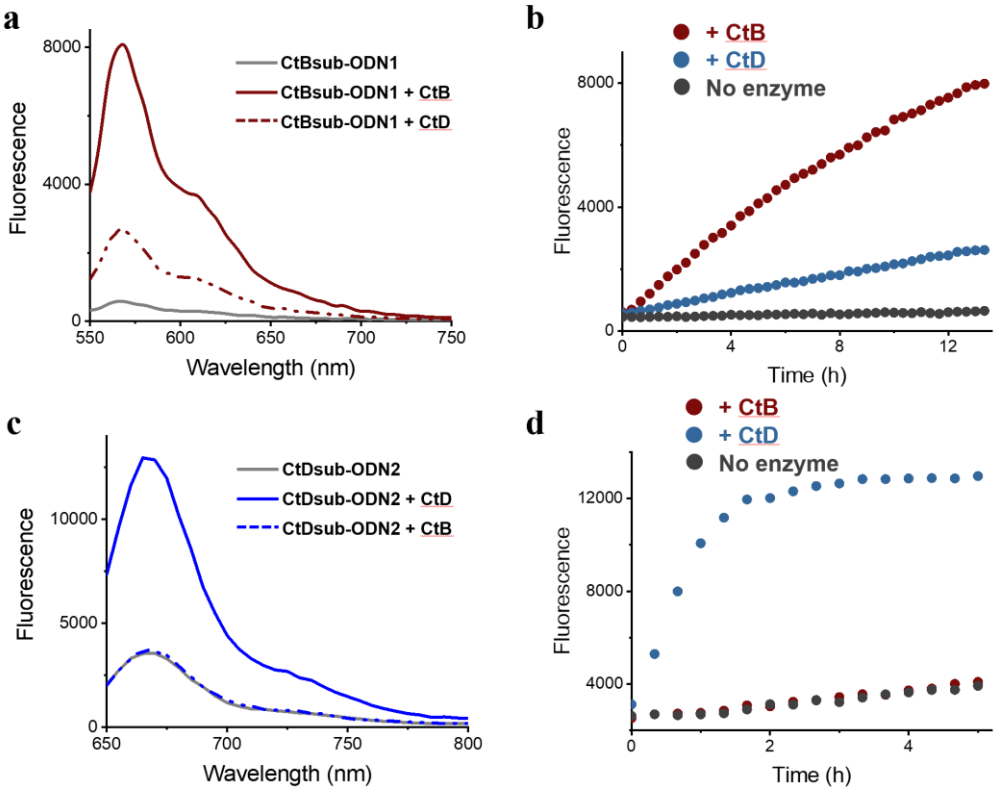


Figure 4.9 Emission spectra of cathepsin probes by reactions with or without cathepsins. (a) CtBsub-ODN1 (200 nM) treated with CtB (5 nM, red solid line), with CtD (5 nM, red dash line) or without cathepsin (gray solid line) after 13 hours incubation (excited at 520 nm). (b) The time-course change of fluorescence intensity in (a). (c) CtDsub-ODN2 (200 nM) treated with CtD (5 nM, blue solid line), with CtB (5 nM, blue dash line) or without cathepsin (gray solid line) after 4 hours incubation (excited at 620 nm). (d) The time-course change of fluorescence intensity in (d).

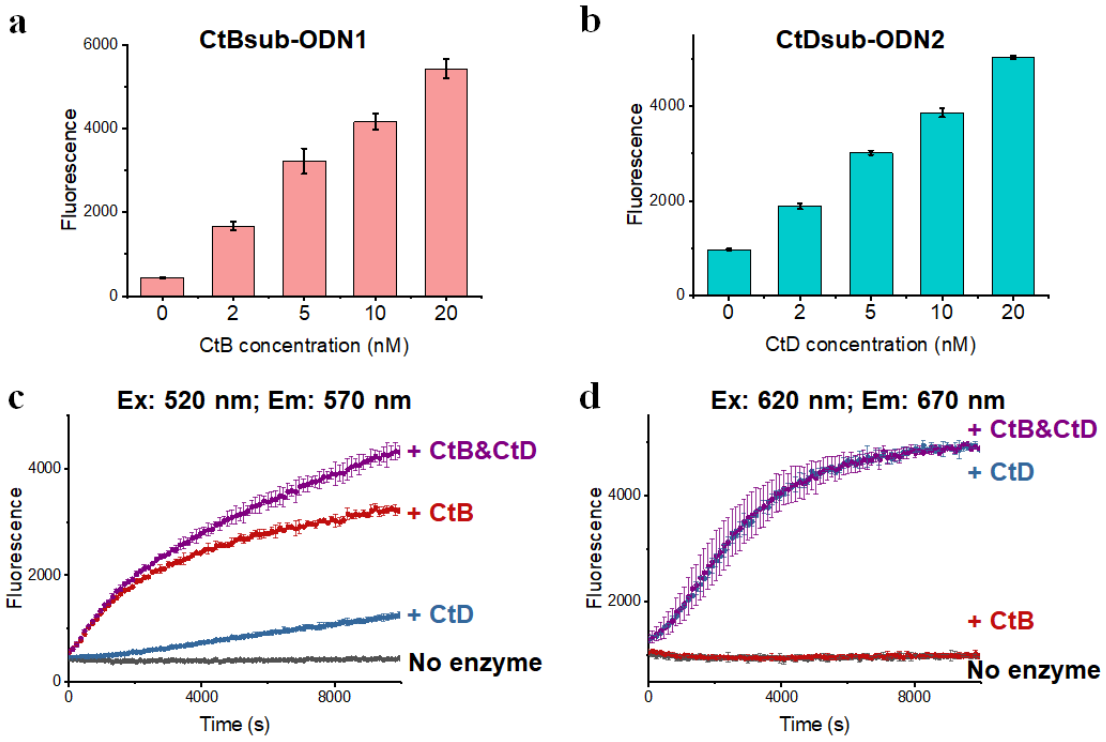


Figure 4.10 Enzymatic reaction with CtBsub-ODN1 and CtDsub-ODN2. (a) CtBsub-ODN1 (200 nM) was incubated with 0 nM, 2 nM, 5 nM, 10 nM and 20 nM of CtB, respectively. The fluorescent intensity was recorded after 3 hours incubation. Excitation is 520 nm and emission is 570 nm. (b) CtDsub-ODN2 was treated with different concentration of CtD (0 nM, 2 nM, 5 nM, 10 nM and 20 nM) and incubated for 1 hour. The fluorescent signal was read under excitation 620 nm and emission 670. Same concentration (200 nM) of CtB and CtD were mixed together to probe cathepsin reactivity simultaneously. The experimental groups were designed with the addition of CtB alone, CtD alone or mixture of CtB and CtD, respectively. The concentration of each cathepsin was kept as 5 nM. The progress of enzymatic reaction was recorded under (c) emission 570 nm with excitation 520 nm and (d) emission 670 nm with excitation 620 nm.

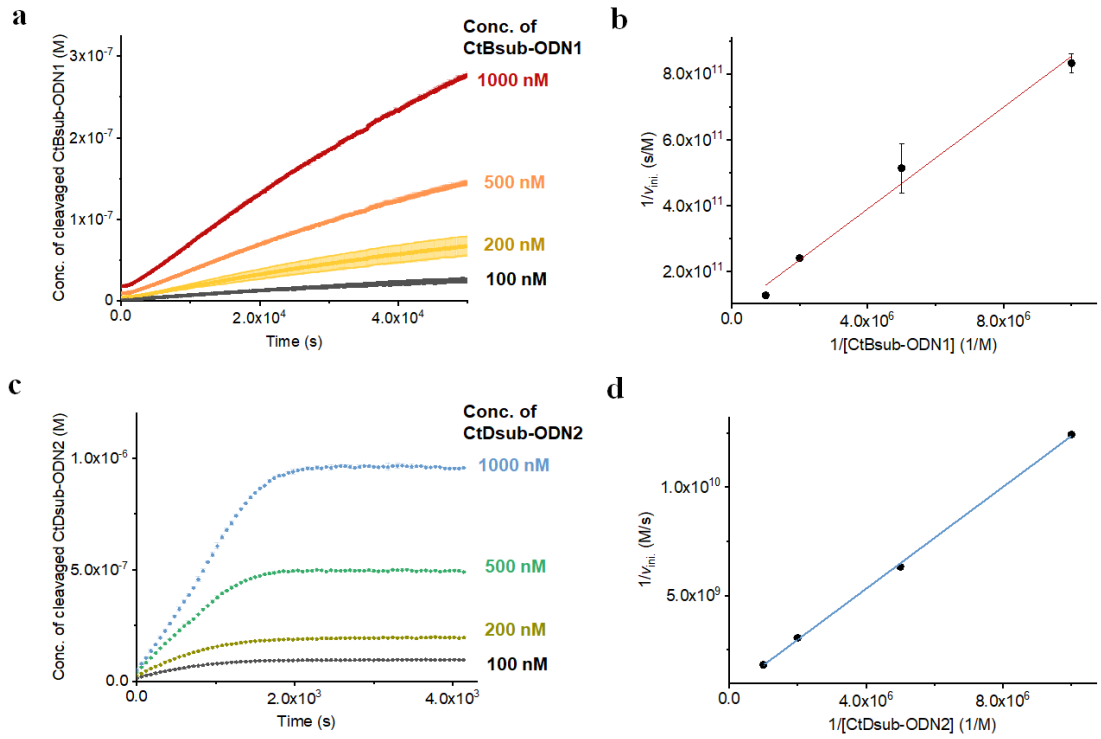


Figure 4.11 Characterization of kinetic parameters for the enzymatic reaction. (a) CtBsub-ODN1 in different concentration (100 nM, 200 nM, 500 nM and 1000 nM) were incubated with CtB (5 nM), respectively. And the fluorescent signal was then converted to the concentration of cleavaged CtBsub-ODN1 according to HPLC analysis. (b) Plotting the reciprocals of the initial velocity (the linear region from 0 s to 2.0×10^4 s) of every experimental groups in (a) following the reciprocals of their concentration. The k_{cat}/K_m ($(2.6 \pm 0.1) \times10^3$ $M^{-1}s^{-1}$) of CtBsub-ODN1 was calculated from the slope value of the linear fitting with CtB concentration (5 nM). (c) CtDsub-ODN2 in different concentration (100 nM, 200 nM, 500 nM and 1000 nM) were incubated at 37°C with CtD (5 nM), respectively. The concentration of cleavaged CtDsub-ODN2 were converted following the same way as (a). (d) The reciprocals of the initial velocity (the linear region from 0 s to 800 s) of every experimental groups in (c) were plotted along with the reciprocals of their concentration. The k_{cat}/K_m of CtDsub-ODN2 was calculated out to be $(1.7 \pm 0.1) \times10^5$ $M^{-1}s^{-1}$.

4.2.4 Assembly of cathepsin B sensor on DNA scaffold through its DNA handle

Besides the improvement of tethered DNA handle to solubility of FRET-based cathepsin sensor, the most significant merit is that DNA handle imparted the probe with precise addressability and enabled it to be assembly on DNA scaffold. Therefore, rectangular DNA origami was designed with 10 sites for complementary DNA sequence of ODN-1 (Figure 4.12a). In this case, 1 nmol of DNA origami is expected to load 10 nmol CtBsub-ODN1 through hybridization on the assigned positions. For loading cathepsin probes, DNA origami was prepared first *via* annealing process and then hybridization between CtBsub-ODN1 and DNA origami proceeded. 10 probes binding on DNA origami can be observed under AFM (Figure 4.12b). With confirmation of probe binding, DNA origami was treated with cathepsin B and the time-course fluorescence enhancement was recorded. The signal increment of DNA origami bearing probes showed similar fluorescence enhancement compared with the free CtBsub-ODN1 (Figure 4.12c and 4.12d). The concentration of free CtBsub-ODN1 was 10 times higher than that of DNA origami, on account of 10 probes assumed to be assembled on DNA origami. And DNA origami without binding site mixed with CtBsub-ODN1 demonstrated same tendency of signal change as the group only containing free CtBsub-ODN1, which indicated DNA origami itself had no influence on the enzymatic reaction (Figure 4.12d). The result verified that CtBsub-ODN1 assembled on DNA origami can be used to detect cathepsin B. Hence, the DNA handle conjugated with cathepsin substrate is the premise to functionalize the DNA scaffold and to realize multiple-signal detection on DNA scaffold.

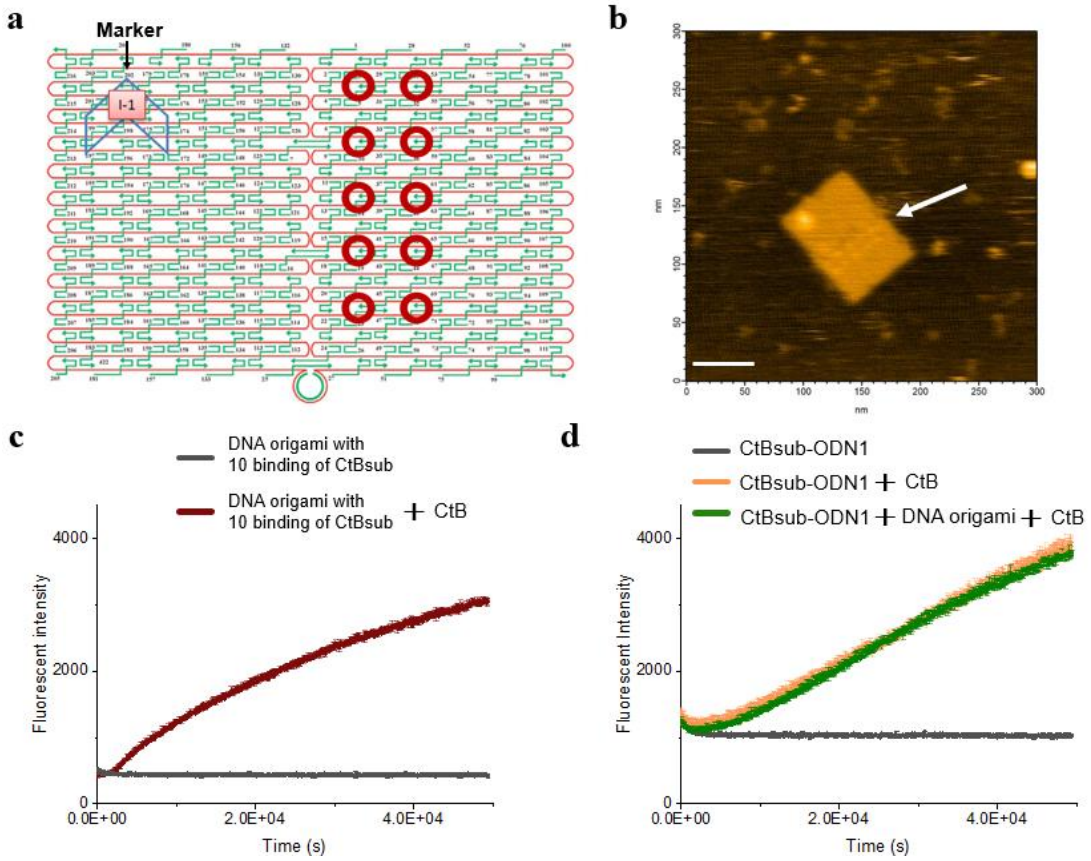
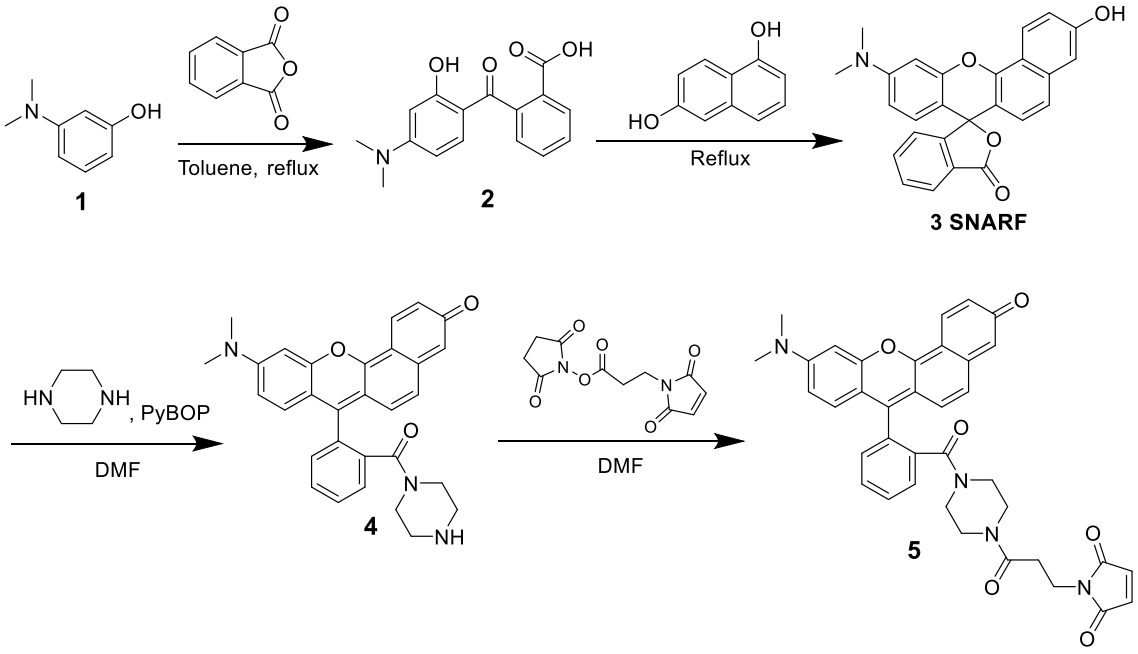


Figure 4.12 (a) Design of rectangular DNA origami. Red circles indicated the binding site for CtBsub-ODN1 (b) AFM images of rectangular DNA origami with 10 bindings of cathepsin B probes. The arrow indicated the bound cathepsin B probes through DNA hybridization. Scale bar: 50 nm. (c) DNA origami with 10 bindings of cathepsin B probes (3 nM) treated with or without CtB (5 nM). (d) CtBsub-ODN1 (30 nM) incubated with or without CtB (5 nM). One of groups with CtBsub-ODN1 was added with DNA origami (3 nM) without binding site for probes as a control group.

4.2.5 Modification of pH-sensitive fluorophore for its conjugation with DNA handle

The pH-sensitive fluorophores based on the SNARF structure have been designed and investigated to present different pK_a towards acidic and neutral condition.[33–35] In order to assembly SNARF on DNA scaffold, SNARF was designed and synthesis to be tethered with a maleimido group. Maleimido group was used to react with thiol group modified on the oligodeoxynucleotide (ODN). With SNARF-modified ODN, SNARF is able to be assembled on DNA scaffold as pH-response sensor for monitoring the intracellular pH changes. For the synthesis procedures, a piperazine was synthesized between SNARF and maleimido, to avoid the structural isomers if maleimido was directly modified on SNARF (Scheme 4.4).[39] The final products were characterized by ^{1}H NHR and ESI-TOF-MS (Figure 4.13).



Scheme 4.4 Synthesis scheme of maleimido-modified SNARF.

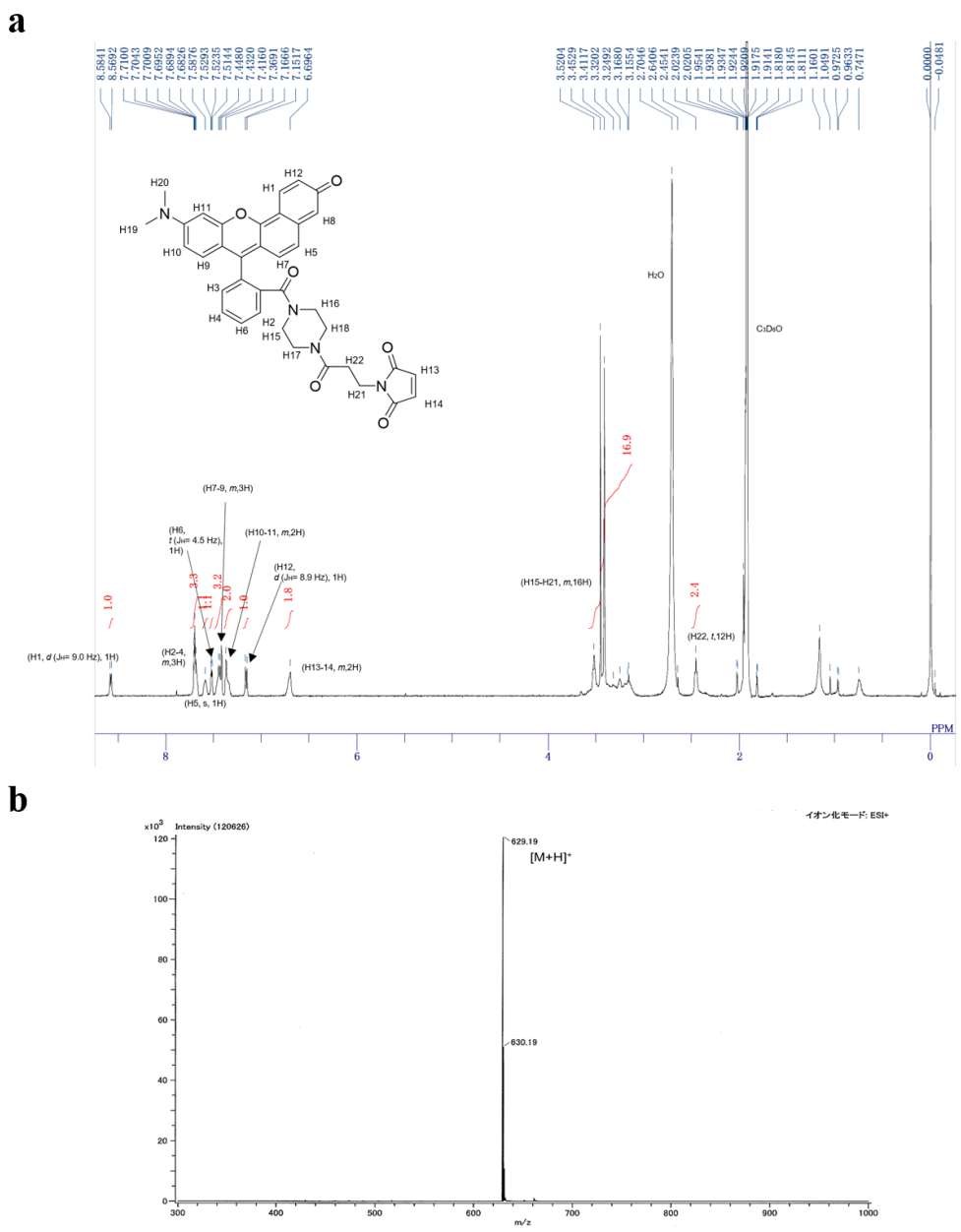


Figure 4.13 (a) ^{1}H NMR spectrum of maleimido-modified SNARF. Assignment of each hydrogen atom was marked on the NMR spectrum. (b) MS spectrum of maleimido-modified SNARF, *m*/*z* 629.2, for $[M+H]^{+}$, ESI-TOF-MS positive mode (calcd. 628.2).

4.3 Conclusion

I conjugated the DNA handle with two types of cathepsin probe to make CtBsub-ODN1 and CtDsub-ODN2. Two cathepsin probes were modified with orthogonal FRET pairs and showed sufficient reactivity and good specificity to detect target enzymes. The fact manifested that introducing DNA handles to FRET-based probes could not only expand its application to assemble on the scaffold, but also mitigated the barrier of hydrophobicity of FRET pairs for enabling multiple FRET pairs used on the design of the protease probes. As reported previously, both CtB and CtD apparently take part in the regulation to the other. At this point, CtBsub-ODN1 and CtDsub-ODN2 were able to be used simultaneously for detect reactivity of CtB and CtD. As a result, it showed that these probes were potent tools to monitor the activity of CtB and CtD in spontaneous manner. Significantly, CtBsub-ODN1 was successfully assembly on DNA scaffold and DNA scaffold with probe binding was able to detect activity of CtB even the concentration of DNA scaffold was very less. The results demonstrated remarkable availability of DNA handle tethered on functional molecules and powerful functionality of DNA scaffold. And Coupled with pH sensor, it is prospect to simultaneously use cathepsin sensor and pH sensor loaded on DNA scaffold to detect active cathepsin inside lysosomes.

4.4 Materials and methods

4.4.1 Materials

Fmoc-protected amino acids, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBt), Rink Amide resin, distilled N,N-dimethylformamide (DMF) and Trifluoroacetic acid (TFA) were obtained from Watanabe Chemical Industries (Hiroshima, Japan).

5-Azidovaleric Acid was obtained from Tokyo Chemical Industry (Tokyo, Japan). Quasar 570 carboxylic acid, Quasar 670 carboxylic acid, BHQ-2 succinimidyl ester and BHQ-3 succinimidyl ester were purchased from Biosearch Technologies (California, USA). Active human recombinant cathepsin B, cathepsin D from bovine spleen, N,N-diisopropylethylamine (DIEA), tris-hydroxypropyltriazolylmethylamine (THPTA) and 20,40,60-trihydroxyacetophenone monohydrate (THAP) were obtained from Sigma-Aldrich (St. Louis, MO, USA). 5'-Hexynyl-modified oligonucleotides were purchased from JBioS (Saitama, Japan). Custom synthesized staple strand oligonucleotides for the preparation of DNA origami scaffolds were purchased from Sigma-Aldrich (St. Louis, MO), Thermo Fisher Scientific Inc. (Waltham, MA, USA), Japan Bio Services Co., LTD (Saitama, Japan) and Gene Design Inc. Sephacryl S-400 was purchased from GE Healthcare Japan Corporation (Tokyo, Japan). Diammonium hydrogen citrate (DAHC) and a reversed-phase C18 column COSMOSIL $5C_{18}$-MS-II (For analysis: 4.6 × 150 mm, particle size 5 μm) for purifying DNA-protein conjugates were obtained from Nacalai Tesque (Kyoto, Japan). A reversed-phase C18 column ULTRON VX-ODS (For analysis: 4.6 × 150 mm; For purification, 20 × 250 mm, particle size 5 μm) for peptide purification was purchased from Shinwa Chemical Industries (Kyoto, Japan). Copper sulfate, sodium ascorbate and all other chemicals and reagents were purchased from FUJIFILM Wako Pure Chemical Industries, Ltd. (Osaka, Japan).

4.4.2 Synthesis of FRET pair-labeled peptide

The peptide synthesis of cathepsin substrate was performed according to the Fmoc chemistry protocol. Rink Amide resin was treated by DCM 30 min before using for peptide synthesis. Deprotection was proceeded by 30% pipyridine for 10 min under

37℃. Fmoc-protected amino acid (4 eq.) was pre-activated by HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF. And the solution was mixed with resin for 60 min under 37℃. 5-Azidovaleric acid was used to react on the N-terminal of the peptide by using the same coupling condition after the peptide sequence was completed. For modifying both fluorophore and quencher on the peptide, two lysines were designed on the N-terminal and C-terminal of the peptide. The lysine close to N-terminal and C-terminal was protected by Mmt and Boc group, respectively. After the synthesis of peptide with 5-azidovaleric acid, Mmt group on the lysine close to N-terminal was removed by 2% trifluoroacetic acid (TFA) in DCM. The deprotection reaction was carried out at room temperature for 4 min and subsequently washing it three times with DCM. This was repeated 5 times to assure total removal of the protecting group. A solution of Quasar 570 or 670 (1.5 eq), HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF was added to the resin with the deprotected lysine and incubated at 37℃ for 15 hours. For the cleavage of the peptide, a mixture of TFA and water (97.5:2.5) was added to the resin and shaken for 4 hours at room temperature. Peptides were then treated by diethyl ether precipitation and dissolved in acetonitrile/water (1:1) with TFA (0.1 % v/v) for reversed-phase HPLC purification. Purified peptides were acquired as white powder by lyophilization.

For the quencher modification, cathepsin substrate (1 eq.) modified by fluorophore was dissolved in DMF and mixed with BHQ2-NHS or BHQ3-NHS (1.2 eq) in DMF and trimethylamine (2 eq.) for 120 min under 30℃. The reaction solution was then diluted by acetonitrile/water (1:2) for 5 times and purified by HPLC. The synthesized cathepsin sensors were characterized by MALDI-TOF mass spectrometry (AXIMA-LNR, Shimadzu, CHCA matrix)

4.4.3 Conjugation of the DNA handle with hydrophobic peptide

The DNA handle was conjugated with cathepsin sensor through CuAAC click reaction. The 5'-hexynyl-modified ODN, ODN-1 or ODN-2 (6 nmol, 1 eq.), was used after the solvent was removed. Thereafter, the reagent mix was prepared as follows: 20 mM THPTA, 32 mM sodium ascorbate and 4 mM $CuSO_4$ were dissolved in sodium phosphate buffer (30 μL of 0.1 M, pH=7.0) and the buffer was used to re-dissolve the oligonucleotides. After degassing the ODN solution, it was mixed with cathepsin sensor (180 nmol, 30 eq.) in DMF (30 μL) and degassed three times. And next the reaction mixture was incubated under 37℃ for 5 min. The target product was monitored and purified by means of HPLC (0.1 M TEAA and 50% acetonitrile in 0.1 M TEAA).

The final products were characterized by MALDI-TOF mass spectrometry (AXIMA-LNR, Shimadzu). A matrix solution was prepared as a mixture (9:1) of 10 mg/mL of THAP in acetonitrile/H_2O (1:1) and 50 mg/mL of DAHC in pure water.[40] The sample solution was mixed with the same volume of matrix solution (1 μL each). All measurements were performed in linear positive mode.

4.4.4 Fluorescence measurements and FRET analyses

All *in vitro* experiments were performed with an Infinite 200 PRO microplate reader (TECAN Austria GmbH) at 37°C. Enzymatic assays were carried out in 100 mM sodium acetate, 100 mM sodium chloride, 1 mM EDTA, 0.02% Tween20 (pH 5.5) at 37°C. The cathepsin enzyme was activated in a buffer containing 100 mM sodium acetate, 100 mM sodium chloride, 1 mM EDTA, 1 mM DTT (pH 5.5) for 15 min under 37°C. CtBsub-ODN1 were observed at 520 nm (excitation) and 570 nm (emission) wavelengths (gain 130). And CtDsub-ODN2 were monitored at 620 nm (excitation) and 670 nm (emission) wavelengths (gain 130).

4.4.5 Preparation of DNA origami scaffold assembled with cathepsin sensors

A solution (450 μL) was prepared containing M13mp18 (20 nM) and staple strands (5 equiv., 100 nM of each strand) in a buffer containing 33 mM sodium acetate, 33 mM sodium chloride, 0.33 mM EDTA, 12.5 $MgCl_2$ (pH 5.5). The solution was heated at 95°C for 1 min, rapidly cooled down to 53°C, incubated for 30 min, and finally rapidly cooled down to 4°C by using a thermal cycler. The sample was then purified by using Sephacryl S-400 (400 ml) Ultrafree-MC-DV centrifugal filters, and in a buffer containing 33 mM sodium acetate, 33 mM sodium chloride, 0.33 mM EDTA, 12.5 $MgCl_2$ (pH 5.5, for washing) to get rid of the excess ODNs. The purified samples were lyophilized to remove the water and re-dissolved by water (150 μL) again to make a condensed solution (100 μL). A solution containing cathepsin probe-conjugated ODN (50 μL, 600 nM) was mix with DNA origami solution to make a solution (150 μL) containing DNA origami (3 times condensed), cathepsin probe-conjugated ODN (200 nM), 100 mM sodium acetate, 100 mM sodium chloride, 1 mM EDTA, 37.5 $MgCl_2$ (pH 5.5). The solution was incubated at 41 °C for 5 min, temp. 0.1 °C reduced every 3s to 4 °C using a thermal cycler. The sample was then purified by using Sephacryl S-400 (400 ml) Ultrafree-MC-DV centrifugal filters, and in a buffer containing 100 mM sodium acetate, 100 mM sodium chloride, 1 mM EDTA, 12.5 mM $MgCl_2$ (pH 5.5, for washing) to get rid of the excess peptide-conjugated ODNs and to exchange the concentration of $MgCl_2$ (37.5 mM to 12.5 mM). The acquired solution was the origami with the 10 bindings of cathepsin probes.

4.4.6 AFM imaging

The purified sample was incubated on a freshly cleaved mica surface (ϕ 1.5 mm) for 5 min at room temperature. The surface was then washed for three times to

remove the unbound excess sample by using a buffer containing 100 mM sodium acetate, 100 mM sodium chloride, 1 mM EDTA (pH 5.5). The sample was scanned under tapping mode using a Sample-Scanning-Nano Explorer AFM system (SS-NEX, RIBM Co., Ltd., Tsukuba, Japan) with a silicon nitride cantilever (resonant frequency: 1.5 MHz, spring constant: 0.1 N/m, Olympus, Tokyo, Japan). Several images were acquired from different regions on the mica surface.

4.4.7 Modification of maleimido on SNARF derivatives

The synthesis of SNARF was conducted by following previous published protocol. The maleimide-modified SNARF derivative (SNARF-mal) was prepared following the reference. The compound numbers follow Scheme 4.4.

*Synthesis of Compound **4*** [39]

Piperazine (88 mg, 1.28 mmol) was dissolved in CH_2Cl_2 (5 mL). And 1.4 M trimethylaluminium in hexane (731 μL, 1.02 mmol) was added dropwise into piperazine solution and white precipitations were produced during the addition. And the solution was stirred for 1 hour. SNARF (104 mg, 0.25 mmol) was totally dissolved in CH_2Cl_2 (20 mL) and the solution was added dropwise into the solution which contained piperazine. After 2 h reaction under reflux, the reaction solution was put on ice bath for cooling down. And 0.1 M aqueous solution of HCl was added dropwise until gas evolution ceased. The solution was filtered and the retained solids were washed by CH_2Cl_2 (25 mL) for twice and CH_2Cl_2/MeOH (4:1, 25 mL) for once. The combined filtrate was concentrated and the residue was dissolved in CH_2Cl_2, filtered to remove insoluble salts, and concentrated again. The resulting solid was then re-dissolved in chloroform and washed by water. After concentration, the crude products were purified by amino-coated silica column under the gradient MeOH 0% to 8% (in $CHCl_3$). Finally,

the products (58 mg, 48%) were collected. ^{1}H NMR (600 MHz; DMSO): δ 8.72 (d, 1H, J_H = 9.0 Hz, *5H*-naphthalene *H7*), 7.75-7.64 (m, 4H, *4H*-benzene *H2, H4, H5*; *5H*-naphthalene *H3*), 7.52 (d, 1H, J_H = 9.1 Hz , *4H*-benzene *H3*), 7.37-7.28 (m, 5H, *5H*-naphthalene *H4, H2*; *3H*-benzene *H3, H4, H6*), 7.09 (d, 1H, J_H = 9.0 Hz , *5H*-naphthalene *H8*), 3.63 (s, 2H, -CON-2(CH_2)-), 3.32-3.27 (m, 8H, -CON-2(CH_2)-; -Ar-N-2(CH_3)), 2.99 (s, 2H, -2(CH_2)-NH-), 2.84 (s, 2H, -2(CH_2)-NH). ESI-TOF-MS *m/z* = 478.17 $[M+H]^+$ (calcd. for $C_{30}H_{28}N_3O_3$, 478.21)

*Synthesis of Compound **5*** (Maleimido-modified SNARF)

*Compound **4*** (30 mg, 0.06 mmol) was dissolved in DCM (5 mL). N-Succinimidyl 3-Maleimidopropionate (33 mg, 0.12 mmol) and trimethylamine (18 μL, 0.12 mmol) were added into the solution. Reacting for 2 h, DCM was removed and the crude products was purified by rp-column under the gradient MeOH 10% to 90% (in H_2O). The products containing *compound **5*** was collected from rp-column purification and further purified by HPLC under the gradient phase B 33% to 45% in 24 min, which the flowing phase A and B were 0.05% TFA and 80% acetonitrile with TFA (0.05%), respectively. The solution containing *compound **5*** was evaporated to remove the acetonitrile and lyophilized to get rid of water. The final products (6.5 mg, 17%) were acquired. ^{1}H NMR (600 MHz; Acetone-d6): δ 8.57 (d, 2H, J_H = 9.0 Hz, *5H*-naphthalene *H7*), 7.71-7.68 (m, 3H, *4H*-benzene *H2, H4, H5*), 7.59 (s, 1H, *5H*-naphthalene *H3*), 7.52 (t, 1H, J_H = 4.5 Hz, *4H*-benzene *H3*), 7.45-7.37 (m, 5H, *5H*-naphthalene *H4, H2*; *3H*-benzene *H3, H4, H6*), 7.16 (d, 1H, J_H = 8.9 Hz , *5H*-naphthalene *H8*), 6.70 (s, 2H, maleimide), 3.52-3.16 (m, 16H, piperazine; -Ar-N-2(CH_3); -CH_2-N-2(CO)-), 2.45 (s, 2H, -CH_2-CO-N-). ESI-TOF-MS *m/z* = 629.19 $[M+H]^+$ (calcd. for $C_{37}H_{33}N_4O_6$, 629.24)

4.5 References

(1) Nguyen, T. M.; Nakata, E.; Saimura, M.; Dinh, H.; Morii, T. Design of Modular Protein Tags for Orthogonal Covalent Bond Formation at Specific DNA Sequences. *J. Am. Chem. Soc.* **2017**, *139*, 8487–8496.

(2) Nguyen, T. M.; Nakata, E.; Zhang, Z.; Saimura, M.; Dinh, H.; Morii, T. Rational Design of a DNA Sequence-Specific Modular Protein Tag by Tuning the Alkylation Kinetics. *Chem. Sci.* **2019**, *10*, 9315–9325.

(3) Dinh, H.; Nakata, E.; Mutsuda-Zapater, K.; Saimura, M.; Kinoshita, M.; Morii, T. Enhanced Enzymatic Activity Exerted by a Packed Assembly of a Single Type of Enzyme. *Chem. Sci.* **2020**, *11*, 9088–9100.

(4) Nakata, E.; Dinh, H.; Ngo, T. A.; Saimura, M.; Morii, T. A Modular Zinc Finger Adaptor Accelerates the Covalent Linkage of Proteins at Specific Locations on DNA Nanoscaffolds. *Chem. Commun.* **2015**, *51*, 1016–1019.

(5) Nakata, E.; Hirose, H.; Gerelbaatar, K.; Arafiles, J. V. V.; Zhang, Z.; Futaki, S.; Morii, T. A Facile Combinatorial Approach to Construct a Ratiometric Fluorescent Sensor: Application for the Real-Time Sensing of Cellular PH Changes. *Chem. Sci.* **2021**, *12*, 8231–8240.

(6) Jiang, T.; Meyer, T. A.; Modlin, C.; Zuo, X.; Conticello, V. P.; Ke, Y. Structurally Ordered Nanowire Formation from Co-Assembly of DNA Origami and Collagen-Mimetic Peptides. *J. Am. Chem. Soc.* **2017**, *139*, 14025–14028.

(7) Wijesekara, P.; Liu, Y.; Wang, W.; Johnston, E. K.; Sullivan, M. L. G.; Taylor, R. E.; Ren, X. Accessing and Assessing the Cell-Surface Glycocalyx Using DNA Origami. *Nano Lett.* **2021**, *21*, 4765–4773.

(8) Dong, Y.; Chen, S.; Zhang, S.; Sodroski, J.; Yang, Z.; Liu, D.; Mao, Y. Folding DNA into a Lipid-Conjugated Nanobarrel for Controlled Reconstitution of

Membrane Proteins. *Angew. Chemie - Int. Ed.* **2018**, *57*, 2072–2076.

(9) Trads, J. B.; Tørring, T.; Gothelf, K. V. Site-Selective Conjugation of Native Proteins with DNA. *Acc. Chem. Res.* **2017**, *50*, 1367–1374.

(10) Madsen, M.; Gothelf, K. V. Chemistries for DNA Nanotechnology. *Chem. Rev.* **2019**. *119*, 6384–6458

(11) Chandrasekaran, A. R.; Anderson, N.; Kizer, M.; Halvorsen, K.; Wang, X. Beyond the Fold: Emerging Biological Applications of DNA Origami. *ChemBioChem* **2016**, 1081–1089.

(12) Yang, Y. R.; Liu, Y.; Yan, H. DNA Nanostructures as Programmable Biomolecular Scaffolds. *Bioconjug. Chem.* **2015**, *26*, 1381–1395.

(13) McCluskey, J. B.; Clark, D. S.; Glover, D. J. Functional Applications of Nucleic Acid–Protein Hybrid Nanostructures.*Trends Biotechnol.* **2020**, *38*, 976–989

(14) Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. Cysteine Cathepsins: From Structure, Function and Regulation to New Frontiers. *Biochim. Biophys. Acta - Proteins Proteomics* **2012**, *1824*, 68–88.

(15) Mort, J. S.; Buttle, D. J. Cathepsin B. *Int. J. Biochem. Cell Biol.* **1997**, *29*, 715–720.

(16) Olson, O. C.; Joyce, J. A. Cysteine Cathepsin Proteases: Regulators of Cancer Progression and Therapeutic Response. *Nat. Rev. Cancer* **2015**, *15*, 712–729.

(17) Aggarwal, N.; Sloane, B. F. Cathepsin B: Multiple Roles in Cancer. *Proteomics - Clin. Appl.* **2014**, *8*, 427–437.

(18) Roshy, S.; Sloane, B. F.; Moin, K. Pericellular Cathepsin B and Malignant Progression. *Cancer Metastasis Rev.* **2003**, *22*, 271–286.

(19) Fujii, T.; Kamiya, M.; Urano, Y. In Vivo Imaging of Intraperitoneally Disseminated Tumors in Model Mice by Using Activatable Fluorescent

Small-Molecular Probes for Activity of Cathepsins. *Bioconjug. Chem.* **2014**, *25*, 1838–1846.

(20) Wang, Y.; Li, J.; Feng, L.; Yu, J.; Zhang, Y.; Ye, D.; Chen, H. Y. Lysosome-Targeting Fluorogenic Probe for Cathepsin B Imaging in Living Cells. *Anal. Chem.* **2016**, *88*, 12403–12410.

(21) Ryu, J. H.; Kim, S. A.; Koo, H.; Yhee, J. Y.; Lee, A.; Na, J. H.; Youn, I.; Choi, K.; Kwon, I. C.; Kim, B. S.; et al. Cathepsin B-Sensitive Nanoprobe for in Vivo Tumor Diagnosis. *J. Mater. Chem.* **2011**, *21*, 17631–17634.

(22) Segal, E.; Prestwood, T. R.; Van Der Linden, W. A.; Carmi, Y.; Bhattacharya, N.; Withana, N.; Verdoes, M.; Habtezion, A.; Engleman, E. G.; Bogyo, M. Detection of Intestinal Cancer by Local, Topical Application of a Quenched Fluorescence Probe for Cysteine Cathepsins. *Chem. Biol.* **2015**, *22*, 148–158.

(23) Verdoes, M.; Oresic Bender, K.; Segal, E.; Van Der Linden, W. A.; Syed, S.; Withana, N. P.; Sanman, L. E.; Bogyo, M. Improved Quenched Fluorescent Probe for Imaging of Cysteine Cathepsin Activity. *J. Am. Chem. Soc.* **2013**, *135*, 14726–14730.

(24) Schleyer, K. A.; Cui, L. Molecular Probes for Selective Detection of Cysteine Cathepsins. *Org. Biomol. Chem.* **2021**, *19*, 6182–6205.

(25) Hoogendoorn, S.; Habets, K. L.; Passemard, S.; Kuiper, J.; Van Der Marel, G. A.; Florea, B. I.; Overkleeft, H. S. Targeted PH-Dependent Fluorescent Activity-Based Cathepsin Probes. *Chem. Commun.* **2011**, *47*, 9363–9365.

(26) Liu, F.; Li, X.; Lu, C.; Bai, A.; Bielawski, J.; Bielawska, A.; Marshall, B.; Schoenlein, P. V.; Lebedyeva, I. O.; Liu, K. Ceramide Activates Lysosomal Cathepsin B and Cathepsin D to Attenuate Autophagy and Induces ER Stress to Suppress Myeloidderived Suppressor Cells. *Oncotarget* **2016**, *7*, 83907–83925.

(27) Song, Y.; Wright, J. G.; Anderson, M. J.; Rajendran, S.; Ren, Z.; Hua, D. H.; Koehne, J. E.; Meyyappan, M.; Li, J. Quantitative Detection of Cathepsin B Activity in Neutral PH Buffers Using Gold Microelectrode Arrays: Toward Direct Multiplex Analyses of Extracellular Proteases in Human Serum. *ACS Sensors* **2021**, *6*, 3621–3631.

(28) Liaudet-Coopman, E.; Beaujouin, M.; Derocq, D.; Garcia, M.; Glondu-Lassis, M.; Laurent-Matha, V.; Prébois, C.; Rochefort, H.; Vignon, F. Cathepsin D: Newly Discovered Functions of a Long-Standing Aspartic Protease in Cancer and Apoptosis. *Cancer Lett.* **2006**, *237*, 167–179.

(29) Aghdassi, A. A.; John, D. S.; Sendler, M.; Ulrich Weiss, F.; Reinheckel, T.; Mayerle, J.; Lerch, M. M. Cathepsin d Regulates Cathepsin b Activation and Disease Severity Predominantly in Inflammatory Cells during Experimental Pancreatitis. *J. Biol. Chem.* **2018**, *293*, 1018–1029.

(30) Laurent-Matha, V.; Derocq, D.; Prébois, C.; Katunuma, N.; Liaudet-Coopman, E. Processing of Human Cathepsin D Is Independent of Its Catalytic Function and Auto-Activation: Involvement of Cathepsins L and B. *J. Biochem.* **2006**, *139*, 363–371.

(31) Wille, A.; Gerber, A.; Heimburg, A.; Reisenauer, A.; Peters, C.; Saftig, P.; Reinheckel, T.; Welte, T.; Bühling, F. Cathepsin L Is Involved in Cathepsin D Processing and Regulation of Apoptosis in A549 Human Lung Epithelial Cells. *Biol. Chem.* **2004**, *385*, 665–670.

(32) Williams, A. C.; Maciewicz, R. A. Activation of cathepsin B, secreted by a colorectal cancer cell line requires low pH and is mediated by cathepsin D. *Int. J. Cancer* **1996**, *554*, 547–554.

(33) Nakata, E.; Yukimachi, Y.; Nazumi, Y.; Uto, Y.; Maezawa, H.; Hashimoto, T.;

Okamoto, Y.; Hori, H. A Newly Designed Cell-Permeable SNARF Derivative as an Effective Intracellular PH Indicator. *Chem. Commun.* **2010**, *46*, 3526–3528.

(34) Nakata, E.; Yukimachi, Y.; Nazumi, Y.; Uwate, M.; Maseda, H.; Uto, Y.; Hashimoto, T.; Okamoto, Y.; Hori, H.; Morii, T. A Novel Strategy to Design Latent Ratiometric Fluorescent PH Probes Based on Self-Assembled SNARF Derivatives. *RSC Adv.* **2014**, *4*, 348–357.

(35) Nakata, E.; Yukimachi, Y.; Uto, Y.; Hori, H.; Morii, T. Latent PH-Responsive Ratiometric Fluorescent Cluster Based on Self-Assembled Photoactivated SNARF Derivatives. *Sci. Technol. Adv. Mater.* **2016**, *17*, 431–436.

(36) Poreba, M.; Groborz, K.; Vizovisek, M.; Maruggi, M.; Turk, D.; Turk, B.; Powis, G.; Drag, M.; Salvesen, G. S. Fluorescent Probes towards Selective Cathepsin B Detection and Visualization in Cancer Cells and Patient Samples. *Chem. Sci.* **2019**, *10*, 8461–8477.

(37) Tung, C. H.; Bredow, S.; Mahmood, U.; Weissleder, R. Preparation of a Cathepsin D Sensitive Near-Infrared Fluorescence Probe for Imaging. *Bioconjug. Chem.* **1999**, *10*, 892–896.

(38) Yasuda, Y.; Kageyama, T.; Akamine, A.; Shibata, M.; Kominami, E.; Uchiyama, Y.; Yamamoto, K. Correction: Characterization of New Fluorogenic Substrates for the Rapid and Sensitive Assay of Cathepsin E and Cathepsin D. *J. Biochem.* **1999**, *126*, 260.

(39) Nguyen, T.; Francis, M. B. Practical Synthetic Route to Functionalized Rhodamine Dyes. *Org. Lett.* **2003**, *5*, 3245–3248.

(40) Formation, C. B.; Nakano, S.; Seko, T.; Zhang, Z.; Morii, T. RNA-Peptide Conjugation through an Efficient Covalent Bond Formation *Appl. Sci.* **2020**, *10*, 8920.

CHAPTER 5

Multi-signal reporters on dextran for the real-time sensing of lysosomal proteases

5.1 Introduction

Polysaccharides are an important class of biomacromolecular assembilies to acquire significant research interest for a variety of drug delivery and tissue regeneration applications thanks to their assured biocompatibility and bioactivity.[1–3] Among different types of polysaccharides, dextran is one of the most commonly used macropinosome markers, and a range of dextran sizes have been used to identify macropinosomes.[4] Typically, 70 kDa dextran has been reported to be more selective for labelling of macropinosomes compared with smaller size of dextran.[5] And Macropinocytosis can subsequently traffic dextran to lysosomes for degradation along the endosomal–lysosomal pathway. Therefore, 70 kDa dextran is an ideal vehicle for our studies on the detection of lysosomal proteases.[6]

Cysteine proteases, one of the essential protease families with 11 members, are responsible for the non-specific, bulk proteolysis in the acidic environment of the endosomal or lysosomal compartment for degrading intracellular and extracellular proteins.[7] Increased expression and proteolytic activity of lysosomal cysteine proteases have often been correlated with poor prognosis for patients with a variety of malignancies.[8] Cathepsin B (CtB) is known to be overexpressed in various cancers, particularly in aggressive cancers, making it an attractive target for tumor-specific prodrug design.[9–12] In the acidic environment of lysosomes where the pH is kept at 4.5-5, procathepsin B can undergo autocatalytic activation, leading to formation of active CtB.[12–15] Alternatively, CtB can be activated by cathepsin D (CtD), which is an aspartic endo-protease universally found in lysosomes.[16–19] Therefore, a reporter for acidic environment or CtD would be the robust evidence for characterizing the activated CtB inside the lysosome. The simultaneous detections of cellular pH would be a potent sensor for lysosomal environment and CtB or CtD detection.

In this chapter, I designed a dual-reporter approach to reporting activated CtB and pH in lysosome. 70 kDa dextran here was used as a lysosomal-target molecular vehicle for loading and trafficking the dual-reporter system into lysosome through macropinosome. The abundant nucleophilic groups on dextran enabled the chemical modifications of chemical handles for further conjugation with designed reporters. The peptidic substrates for CtB and CtD were synthesized and modified with FRET pairs for directly sensing the activity of CtB and CtD. Separately from that, a ratiometric fluorescent pH probe was prepared by assembling two types of fluorophores on dextran for reporting the pH around the microenvironment of detected protease (Figure 5.1).[20] This pH sensor is expected to reported acidic environment for indication the lysosomal pH and activation condition of CtB. Upon characterizing each sensor, the cathepsin sensor and the pH probe would be modified on dextran as orthogonal reporters.

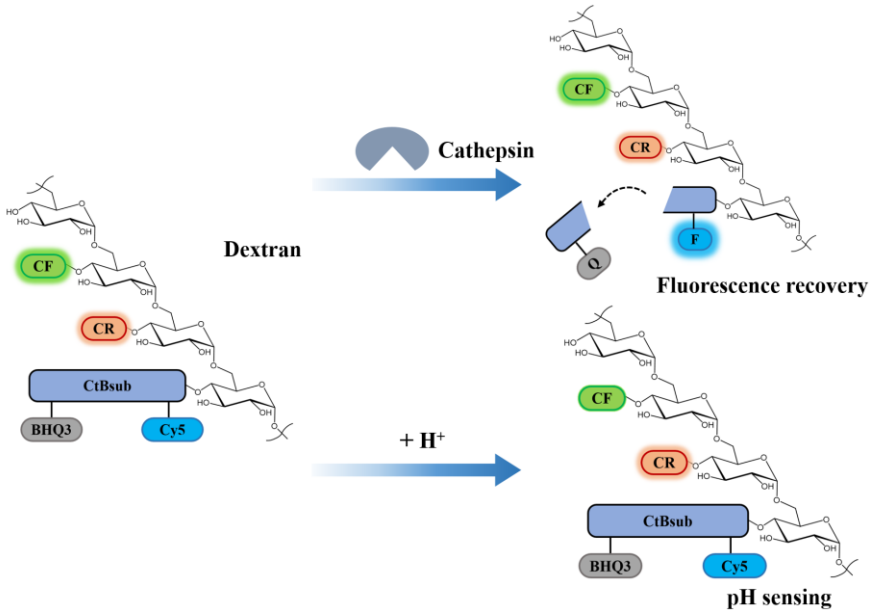
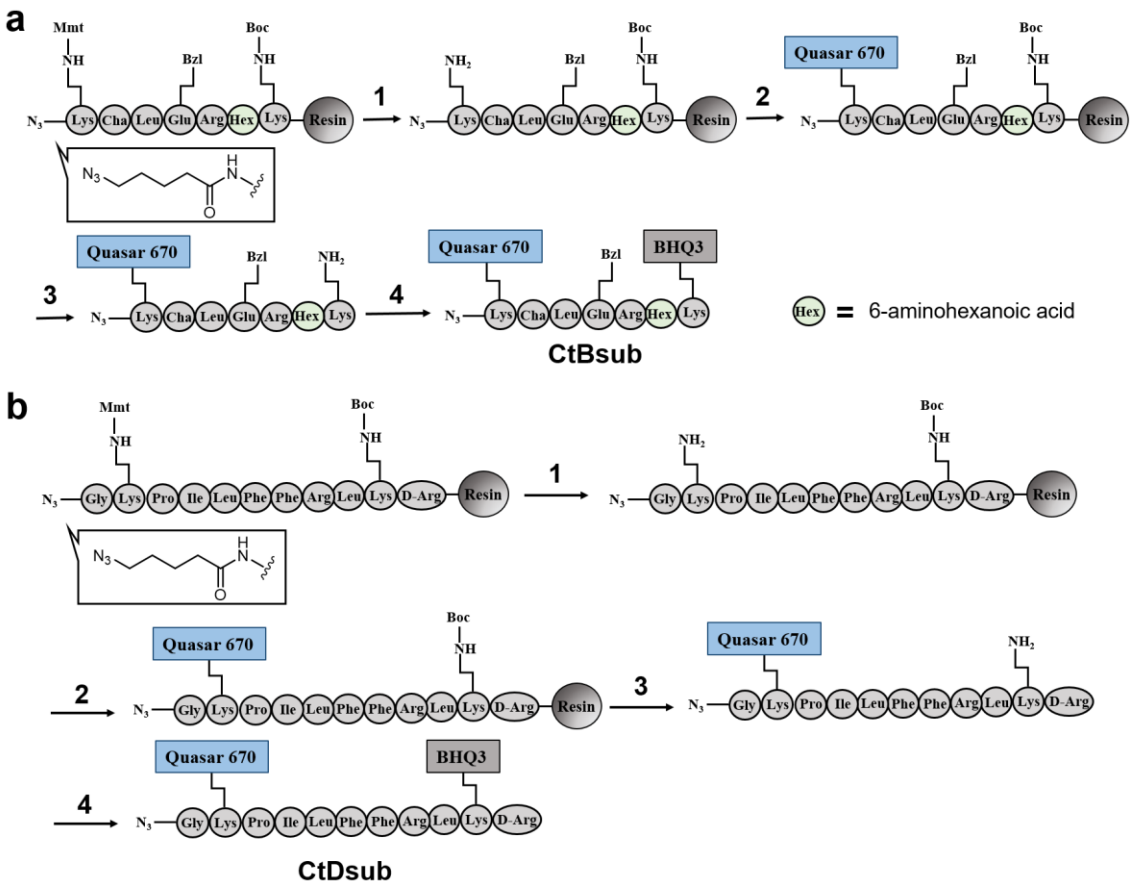


Figure 5.1 Scheme illustration of the dual-responsive reporters on dextran for sensing pH and cathepsin.

5.2 Results and discussion

5.2.1 Design and synthesis of FRET-based sensors for cathepsin B and D

The cathepsin reporters were designed based on FRET dye-quencher pairs, which would show significant fluorescence enhancement after the quencher was cleavaged by cathepsin enzyme. The distinguishable FRET type of fluorophore-quencher pair used for this study was Quasar 670 (Q670)-BHQ3. The excitation and emission wavelength of Q670 was 620 nm and 670 nm, respectively, to be orthogonal with pH sensor. The probes for CtB (CtBsub) and CtD (CtDsub) were designed by referring the reported peptide sequence.[21–23] For CtBsub, a 6-aminohexanoic acid was synthesized on the C-terminal of reported substrate sequence for improving the cleavage efficiency.[24] Two lysines were added to both terminals of CtBsub and CtDsub for Q670 and BHQ3 (Figure 5.2). The newly designed substrates were prepared via solid phase peptide synthesis (SPPS) and a terminal azide group could mediate the conjugation of cathepsin probe on dextran via Cu(I)-free [2+3] cycloaddition. For couple fluorophore and quencher on designed peptide sequence, two lysines located at N- and C-terminals of peptide bore monomethoxytrityl (Mmt) and tert-Butoxycarbonyl (Boc), respectively. Mmt can be orthogonal removed during SPPS for the modification of Q670. And Boc would be removed after peptide cleavage from the resin and enable the modification on C-terminal amino group with BHQ3 in the solution phase. The synthesis procedure was shown in Scheme 5.1. After the purification of these peptides by reverse phase HPLC, the cathepsin substrates containing azido group (CtBsub and CtDsub) were prepared. Both substrates were characterized by HPLC, UV and MS (Figure 5.3 and Figure 5.4).



Scheme 5.1 Synthetic scheme of (a) CtBsub and (b) CtDsub. Reagents and conditions: (1) 2% TFA in DCM (2) Q670-COOH (1.5 eq), HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF (3) 2.5% H_2O in TFA (4) BHQ3-NHS (1.2 eq), and TEA (2 eq.) in DMF.

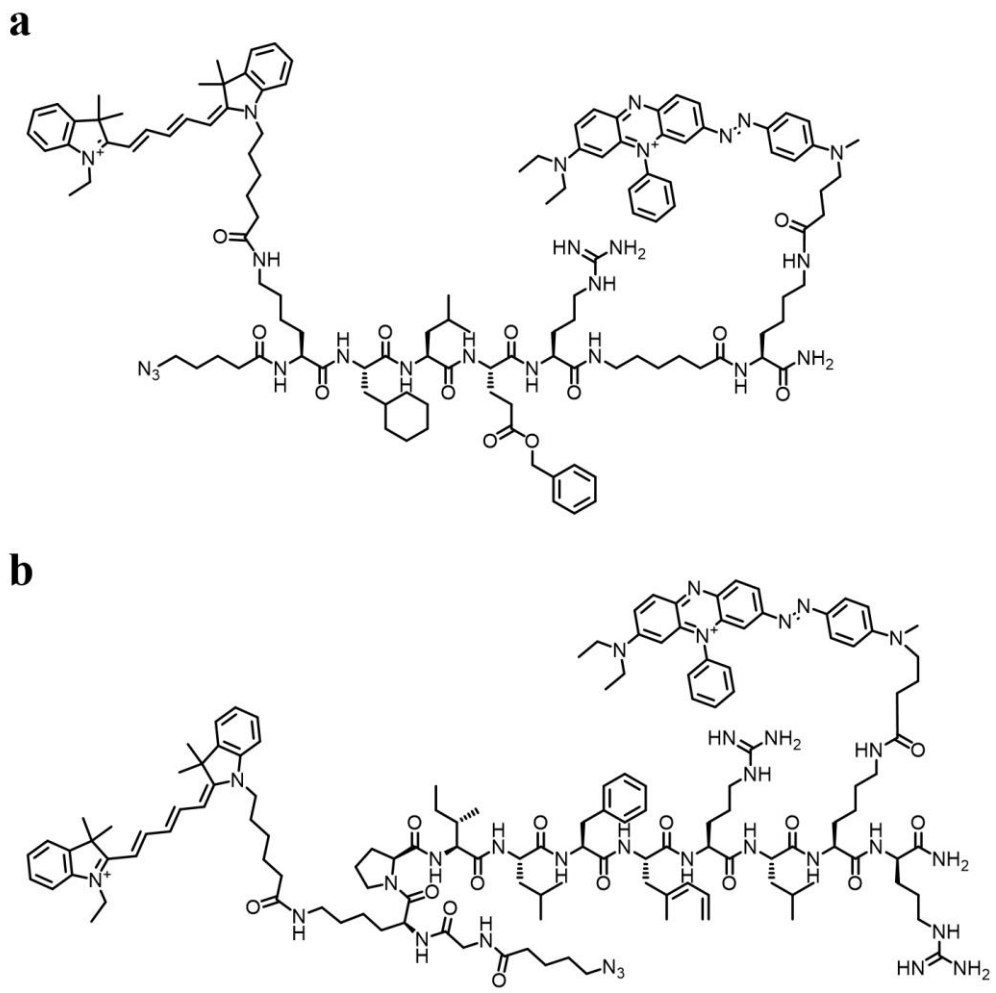


Figure 5.2 Chemical structures of (a) CtBsub and (b) CtDsub

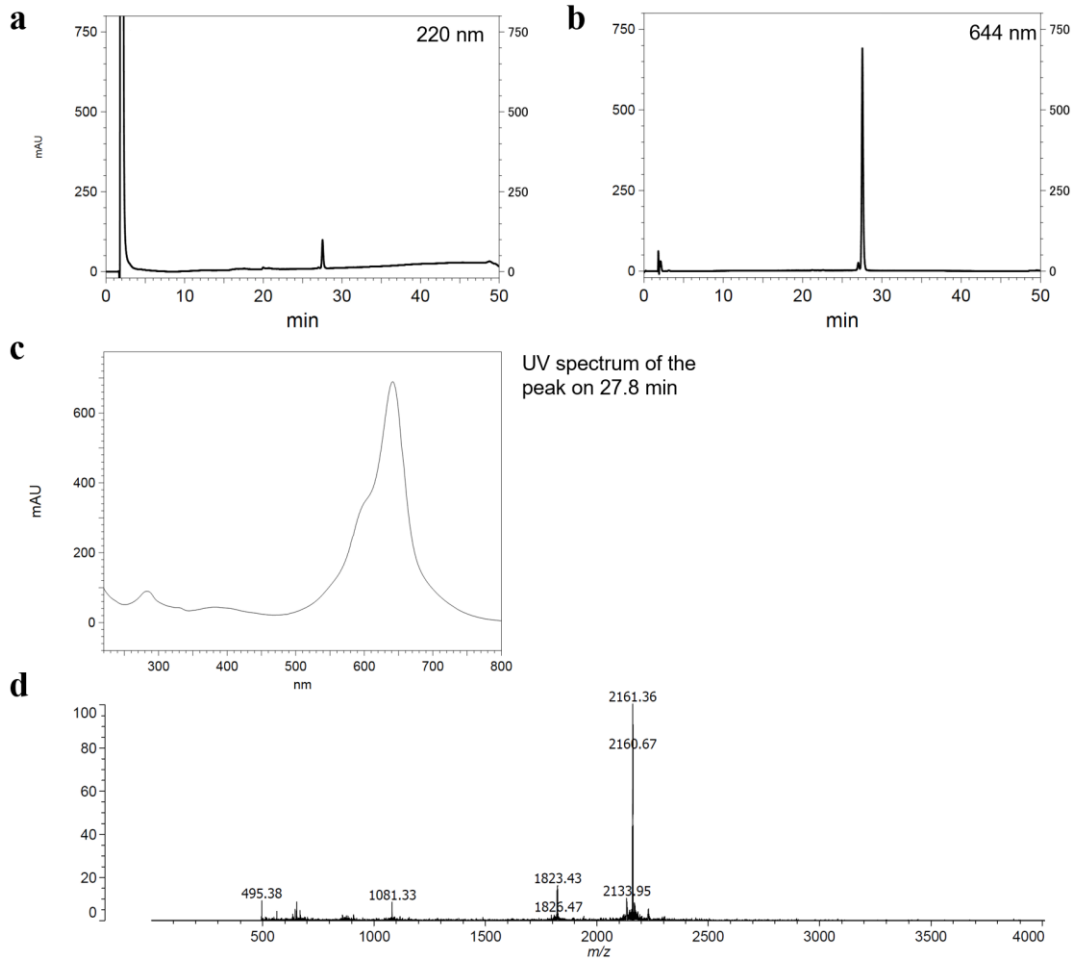


Figure 5.3 HPLC analysis of cathepsin B (CtBsub) modified by Quasar 670 and BHQ3. (a-b) HPLC profiles of the compound detected by UV at (a) 220 nm and (b) 548 nm. (c) UV-vis spectrum of a peak eluted at 27.8 min. HPLC analysis were conducted on a ULTRON VX-ODS column (4.6 × 150 mm, eluent A: 0.05% TFA containing H_2O, eluent B : 0.05% TFA in 80% acetonitrile). Sample was eluted with the gradient of eluent B increased from 35% to 95% in 30 min at a flow rate of 1.0 mL·min^{-1}. (d) MS spectra of CtBsub. *m*/*z* 2161.4, for $[M+H]^+$, MALDI-TOF-MS positive linear mode, CHCA matrix (calcd. 2161.3).

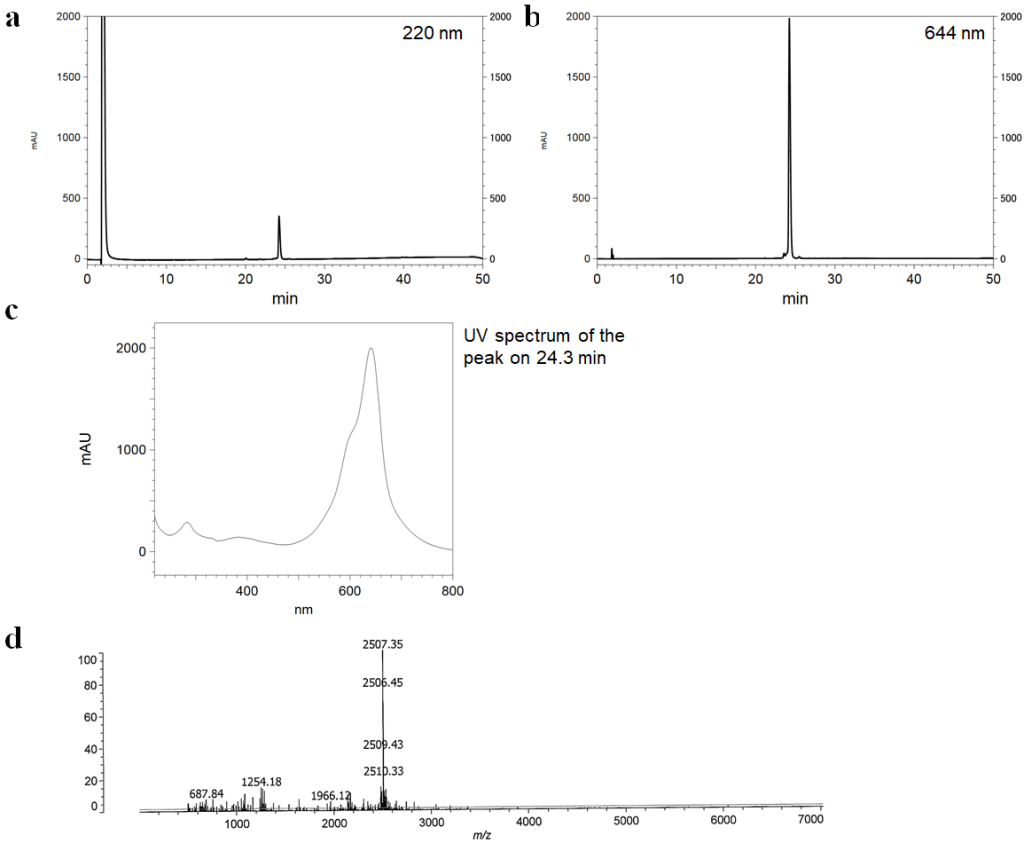
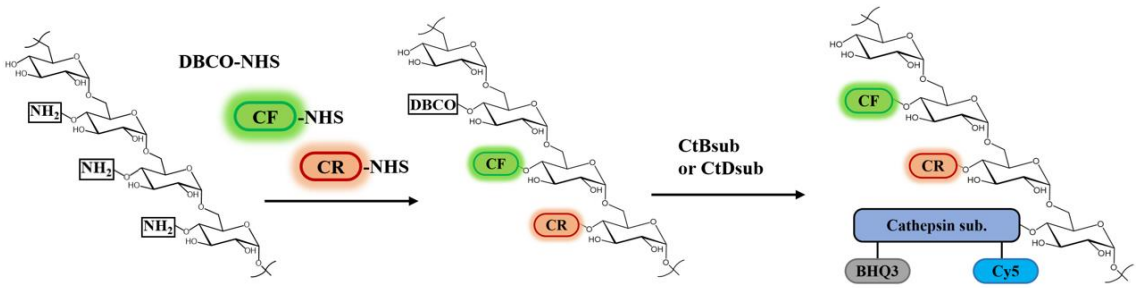


Figure 5.4 HPLC analysis of cathepsin D (CtDsub) modified by Quasar 670 and BHQ3. (a-b) HPLC profiles of the compound detected by UV at (a) 220 nm and (b) 644 nm. (c) UV-vis spectrum of a peak eluted at 24.3 min. HPLC analysis were conducted on a ULTRON VX-ODS column (4.6 × 150 mm, eluent A: 0.05% TFA containing H_2O, eluent B : 0.05% TFA in 80% acetonitrile). Sample was eluted with the gradient of eluent B increased from 35% to 95% in 30 min at a flow rate of 1.0 mL·min^{-1}. (d) MS spectra of CtDsub. *m*/*z* 2507.4, for $[M+H]^+$, CHCA matrix, MALDI-TOF-MS positive linear mode, CHCA matrix (calcd. 2506.5).

5.2.2 Modification of dextran with both pH sensor and cathepsin sensor

For the dual detection of cathepsin and pH inside the lysosome, both cathepsin and pH sensor were modified on dextran. For this purpose, ratiometric fluorescent pH probe, consisting of 6-carboxyfluorescein (CF) with pK_a of 6.5 and a pH-insensitive 6-carboxytetramethylrhodamine (CR) under the physiological conditions,[25] were chosen to be coupled with cathepsin sensor. 70 kDa dextran was purchased with modification of amino group and the loading amount of amino is 29 eq. for dextran (1 eq.). For the modification of cathepsin sensor, a chemical handle dibenzocyclooctyne (DBCO) was supposed to be modified on the amino groups on dextran for processing Cu(I)-free [2+3] cycloaddition with the azido on cathepsin probe. So, at the beginning, CR, CF and DNCO with N-hydroxysuccinimide (NHS) were used to be modified on dextran. And after purification, modified-dextran was mixed with CtBsub or CtDsub to carry out the Cu(I)-free cycloaddition to produce CtBsub-dextran or CtDsub-dextran (Scheme 5.2 and Figure 5.5). The modification process could be monitored on UV spectrum by the UV absorption of CR, CF and Q670-BHQ3 pair. And by using different amount of cathepsin sensor (1 eq., 3 eq. and 6 eq.) for the conjugation, the loading amount of cathepsin sensor on dextran (1 eq.) was able to be regulated (Figure 5.5d). The dextran sample which was modified by CtBsub or CtDsub (6 eq.) was used for the further characterization of the pH and cathepsin sensor loaded on dextran. From the UV spectrum, it could be worked out that dextran (1 eq.) loaded CF (0.4 eq.), CR (0.17 eq.) and CtBsub (0.91 eq.) or CtDsub (2 eq.).



Scheme 5.2 Synthetic scheme of modifying dextran with pH sensor, chemical handle and cathepsin sensor.

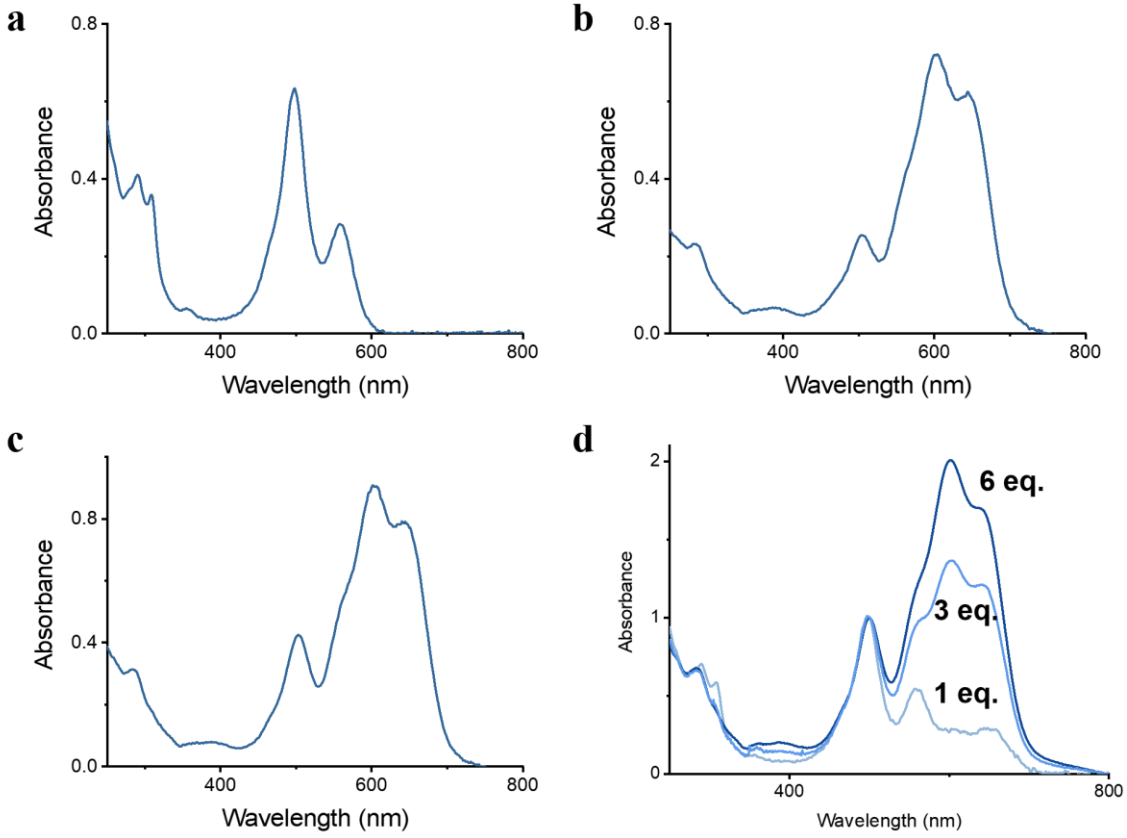


Figure 5.5 UV spectrum of modified dextran. (a) UV spectrum of dextran (1 eq.) modified with CR-NHS (5 eq.), CF-NHS (5 eq.) and DBCO-NHS (10 eq.). (b) UV spectrum of CtBsub-dextran. Dextran (1 eq.) was modified with CtBsub (6 eq.). (c) UV spectrum of CtDsub-dextran. Dextran (1 eq.) was modified with CtDsub (6 eq.). (d) Comparison of UV spectrum of CtBsub-dextran. Dextran (1 eq.) with CF, CR and DBCO were reacted with different equivalents of CtBsub (1 eq., 3 eq. and 6 eq.).

5.2.3 Characterization of the detection capability of the cathepsin sensors on dextran

To test the sensing capability for cathepsin, CtBsub-dextran or CtDsub-dextran were treated with CtB and CtD, respectively. Before mixed with a solution containing dextran-loaded sensor, cathepsin enzyme was activated in an acidic buffer and then added into dextran solution. One type of cathepsin probe on dextran was treated with both CtB and CtD, respectively, for checking reactivity and specificity. And the fluorescent emission spectrum was scanned in a time-dependent manner. As a result, CtBsub-dextran or CtDsub-dextran showed significant fluorescence enhancement while incubated with their specific enzyme, namely CtBsub-dextran to CtB and CtDsub-dextran to CtD (Figure 5.6). Weak signal response also appeared in the experimental group with CtBsub-dextran and CtD, which cannot be observed in the group of CtDsub-dextran and CtB. However, the signal increment of CtBsub-dextran with CtB showed much higher fluorescence enhancement and faster response to treatment of CtB (Figure 5.6e). As a result, cathepsin sensor bore on dextran showed potent sensing ability to their specific enzyme and good orthogonality against unspecific species.

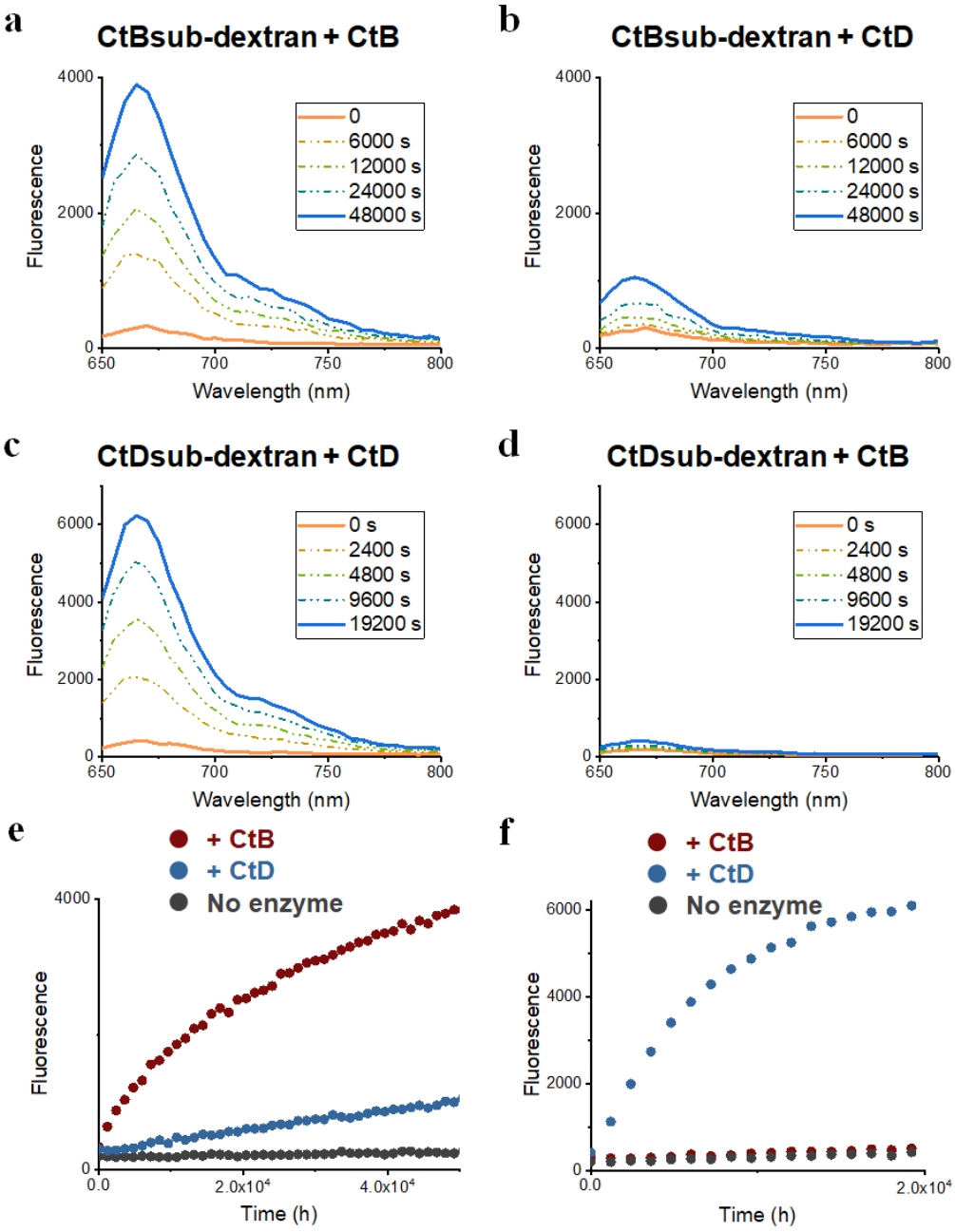


Figure 5.6 The time-dependent fluorescence emission spectral changes of CtBsub-dextran or CtDsub-dextran treated with CtB or CtD. (a) CtBsub-dextran (330 nM) bearing CtBsub (300 nM) treated with CtB (5 nM). (b)CtBsub-dextran (330 nM) bearing CtBsub (300 nM) treated with CtD (5 nM). Similarly, CtDsub-dextran (250 nM) bearing CtDsub (300 nM) treated with (c) CtD (5 nM) or (d) CtB (5 nM). CtBsub-dextran or CtDsub-dextran were excited at 620 nm, respectively. (e) Time-course changes of fluorescent intensity at 670 nm in (a) and (b). (f) Time-course changes of fluorescent intensity at 670 nm in (c) and (d).

5.2.4 Characterization of pH sensor on dextran

The fluorescence emission property of CF modified on dextran (CtBsub-dextran) was next investigated at various pH. The pH insensitive CR was co-assembled on dextran as an internal standard for the fluorescence intensity of CF, thus enabling to evaluate the ratio of fluorescence intensities of CF and CR within the confined space. Averagely, CF (0.4 eq.) and CR (0.2 eq.) were conjugated on dextran. Fluorescence emission intensities of CF and CR on dextran were measured at various pH ranging from pH 3.5 to 10. The fluorescence emission of CF showed a gradual increase in intensity at the emission maxima as the pH increased from 3.5 to 8 (Figure 5.7a). On the other hand, that of CR was not changed in the pH range of 3.5 to 10 (Figure 5.7b). A plot for the ratio of emission intensity of CF/CR at 518 nm excited at 470 nm, corresponding to CF, to that at 582 nm excited at 530 nm, corresponding to CR, at each pH was shown in Figure 3d. The plot allowed an estimation of pK_a for CF on dextran to be 6.09 ± 0.03.[25] These results indicated that CF and CR on dextran can sensitively respond to pH changes in the ratiometric manner from pH 3.5 to 8.

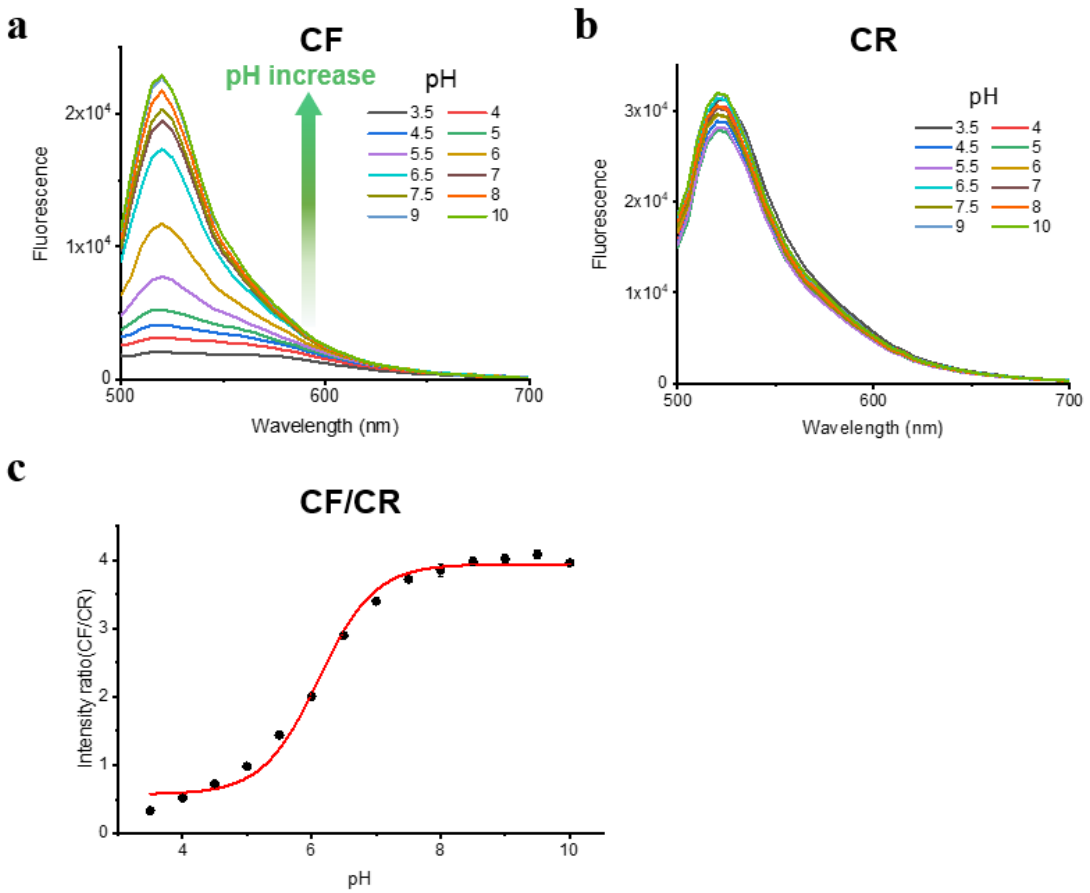


Figure 5.7 Fluorescent pH titration of CF and CR on dextran (330 nM). (a and b) Fluorescence emission spectral changes of (a) CF (excited at 470 nm) and (b) CR (excited at 530 nm) at various pH ranging from 3 to 10. (d) A plot for the ratio of emission intensity at 518 nm excited at 470 nm, corresponding to CF, to that at 582 nm excited at 530 nm, corresponding to CR, for CF5/CR5 at various pH ranging from 3 to 10. The emission intensity ratios (518/582) was calculated by dividing the fluorescence intensity at 518 nm with excitation at 470 nm by the fluorescence intensity at 582 nm with excitation at 530 nm. The pK_a value was estimated to be 6.09 ± 0.03 at ambient temperature.

5.2.5 Estimation of optimal pH of cathepsin reaction *in vitro*

CtB and CtD were known to show enzymatic reactivity for substrate cleavage under acidic condition inside the lysosome.[12,26] Since the pH sensor on dextran was capable of responding to pH changes in the ratiometric manner under acidic condition, CtBsub-dextran or CtDsub-dextran were used to investigate the enzymatic reactivity of CtB and CtD under different pH. CtBsub-dextran or CtDsub-dextran was treated with its target enzyme and time-course fluorescent intensities of Q670 on dextran were measured at various pH ranging from pH 4 to 7. It was found that CtBsub-dextran demonstrated quick response and significant fluorescence enhancement to CtB in the pH range of 4 to 5.5 and CtB showed highest cleavage capability at pH 5.5.[27–29] However, the enzymatic reactivity of CtB dramatically diminished while pH rose over 6. Similar phenomenon was also observed on CtDsub-dextran treated by CtD. CtD could efficiently cleavage CtDsub on dextran in the pH range of 4 to 5.5 and its optimal pH for the highest catalytic ability might be between 4.5 and 5.0.[30] Under the condition of pH from 6 to 7, very less fluorescence enhancement can be observed. Also, the ratio of emission intensity CF/CR were also detected with or without addition of cathepsin enzyme. Judging from the comparison, the ratio changes of CF/CR along with the increase of pH was not influence by the existence of both CtB and CtD. It implied that pH sensor and cathepsin sensor can work independently and orthogonally. Those results verified that pH sensor and cathepsin sensor on dextran could realize orthogonal detection of both pH and cathepsin simultaneously and it was able to be used inside the cell for sensing lysosomal protease.

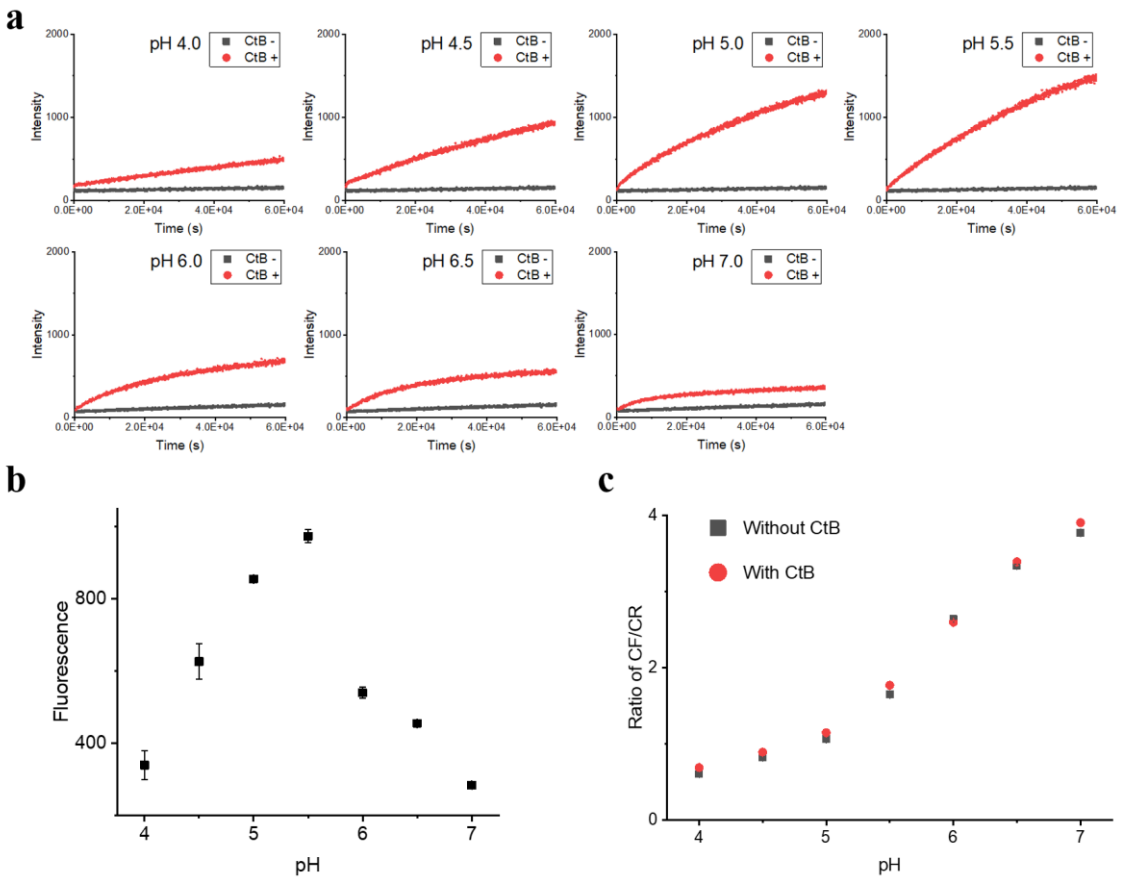


Figure 5.8 Measurement of pH and CtB sensor on CtBsub-dextran (a) Time-course fluorescence measurement of CtBsub-dextran (330 nM) treated with CtB (5 nM) at various pH ranging from pH 4 to 7. (b) Fluorescence in (a) with the incubation of 30000s. (c) A plot for the ratio of emission intensity of CF/CR at various pH ranging from 4 to 7. The comparison was made between the group treated with and without CtB after 60000s incubation.

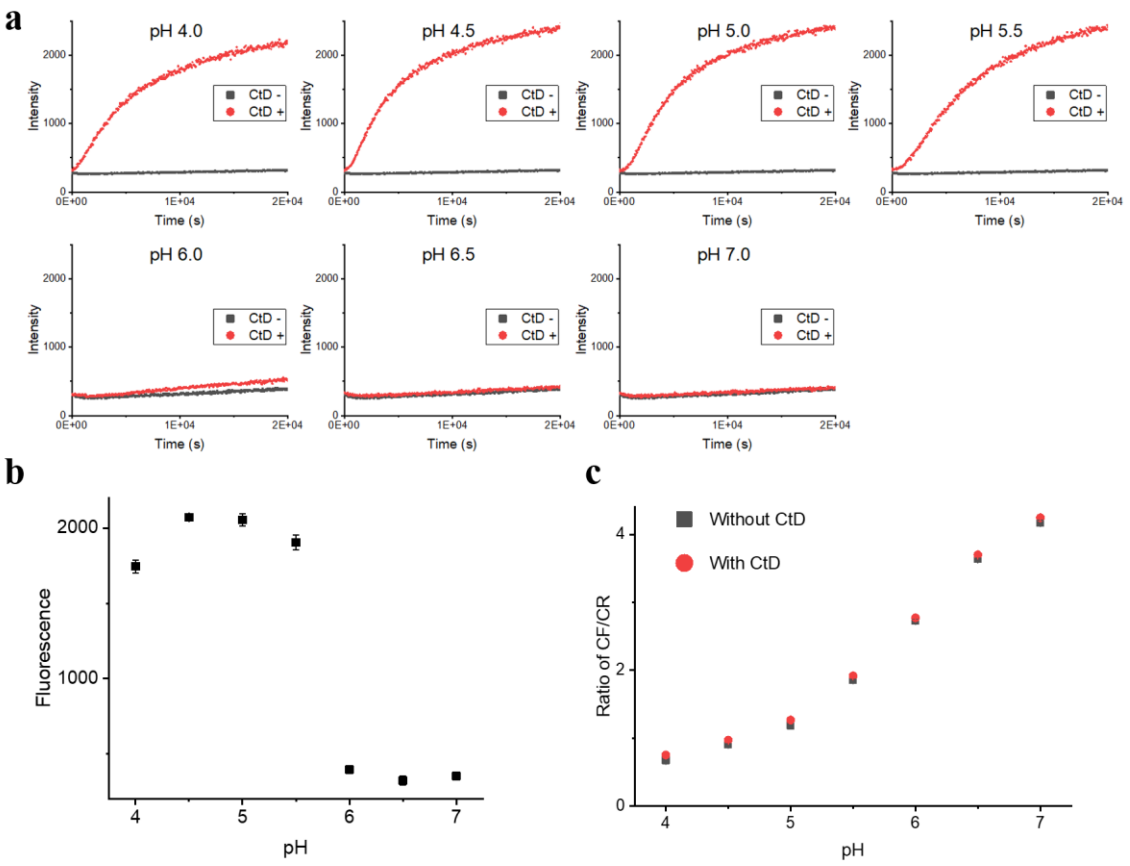


Figure 5.9 Measurement of pH and CtD sensor on CtDsub-dextran (a) Time-course fluorescence measurement of CtDsub-dextran (250 nM) treated with CtD (5 nM) at various pH ranging from pH 4 to 7. (b) Fluorescence in (a) with the incubation of 10000s. (c) A plot for the ratio of emission intensity of CF/CR at various pH ranging from 4 to 7. The comparison was made between the group treated with and without CtD after 20000s incubation.

5.3 Conclusion

CtBsub-dextran or CtDsub-dextran was designed to load both cathepsin sensor and pH sensor on dextran, based on that CtB and CtD were both in activated status under acidic condition of lysosome. Given 70 kDa dextran as an ideal vehicle, it bears abundant nucleophilic groups for chemical modification, amino-coated 70 kDa dextran were modified by CF, CR and DBCO through condensation reaction between NHS and amino groups. With DBCO modified on dextran, FRET-based cathepsin sensor can be conjugated with dextran through Cu(I)-free [2+3] cycloaddition. Q670-BHQ3 FRET pair allows orthogonal fluorescence detection coupled with CF and CR. Through a series of investigations, cathepsin sensor and pH sensor on dextran were verified to probe cathepsin and pH efficiently and independently. Therefore, enzymatic reactivity of CtB or CtD were measured under different pH conditions to show that the highest activity of CtB was around 5.5 which was higher than optimal pH for CtD (between 4.5 and 5). By virtue of dextran for trafficking into lysosome through macropinosome, CtBsub-dextran or CtDsub-dextran are promising for the detection of lysosomal active CtB coupled with the information of pH and existence of active CtD.

5.4 Materials and methods

5.4.1 Materials

Fmoc-protected amino acids, 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 1-hydroxybenzotriazole (HOBt), Rink Amide resin, distilled N,N-dimethylformamide (DMF) and trifluoroacetic acid (TFA) were obtained from Watanabe Chemical Industries (Hiroshima, Japan). 5-Azidovaleric Acid was obtained from Tokyo Chemical Industry (Tokyo, Japan). Quasar 670 carboxylic acid and BHQ-3 succinimidyl ester were purchased from

Biosearch Technologies (California, USA). Active human recombinant cathepsin B, cathepsin D from bovine spleen were obtained from Sigma-Aldrich (St. Louis, MO, USA). Amino-modified 70 kDa dextran, 6-carboxyfluorescein-NHS and 6-carboxytetramethylrhodamine-NHS were purchased from Invitrogen (Waltham, USA). PD-10 Columns were obtained from Cytiva (Marlborough, USA). A reversed-phase C18 column ULTRON VX-ODS (For analysis: 4.6 × 150 mm; For purification, 20 × 250 mm, particle size 5 μm) for peptide purification was purchased from Shinwa Chemical Industries (Kyoto, Japan). Copper sulfate, sodium ascorbate and all other chemicals and reagents were purchased from Wako Chemicals (Tokyo, Japan).

5.4.2 Synthesis of FRET pair-labeled peptide

The peptide synthesis of cathepsin substrate was performed according to the Fmoc chemistry protocol. Rink Amide resin was treated by DCM 30 min before using for peptide synthesis. Deprotection was proceeded by 30% pipyridine for 10 min under 37℃. Fmoc-protected amino acid (4 eq.) including Fmoc-protected 6-aminohexanoic acid was pre-activated by HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF. And the solution was mixed with resin for 60 min under 37℃. 5-Azidovaleric acid was used to react on the N-terminal of the peptide by using the same coupling condition after the peptide sequence was completed. For modifying both fluorophore and quencher on the peptide, two lysines were design on the N-terminal and C-terminal of the peptide. The lysine close to N-terminal and C-terminal was protected by Mmt and Boc group, respectively. After the synthesis of peptide with 5-azidovaleric acid, Mmt group on the lysine close to N-terminal was removed by 2% trifluoroacetic acid (TFA) in DCM. The deprotection reaction was carried out at room temperature for 4 min and subsequently washing it three times with DCM. This was repeated 5 times to assure total removal of

the protecting group. A solution of Quasar 670 (1.5 eq), HBTU (4 eq.), HOBt (4 eq.) and DIEA (8 eq.) in DMF was added to the resin with the deprotected lysine and incubated at 37℃ for 15 hours. For the cleavage of the peptide, a mixture of TFA and water (97.5:2.5) was added to the resin and shaken for 4 hours at room temperature. Peptides were then treated by diethyl ether precipitation and dissolved in acetonitrile/water (1:1) with TFA (0.1 % v/v) for reversed-phase HPLC purification. Purified peptides were acquired as white powder by lyophilization.

For the quencher modification, cathepsin substrate (1 eq.) modified by fluorophore was dissolved in DMF and mixed with BHQ3-NHS (1.2 eq) in DMF and trimethylamine (2 eq.) for 120 min under 30 ℃. The reaction solution was then diluted by acetonitrile/water (1:2) for 5 times and purified by HPLC. The synthesized cathepsin sensors were characterized by MALDI-TOF mass spectrometry (AXIMA-LNR, Shimadzu, CHCA matrix)

5.4.3 Modification of dextran

CF, CR and DBCO were modified on dextran through the reaction between NHS group on CF, CR or DBCO and amino groups on dextran. Amino-coated dextran (100 nmol, 1 eq.) was dissolved by $NaHCO_3$ (100 mM, 250 μL). CF (500 nmol, 5 eq.), CR (500 nmol, 5 eq.) and DBCO (1 μmol, 10 eq.) were dissolved in DMF (250 μL) and this solution was mixed with the solution containing dextran for 12 hours incubation at 30 ℃. The reaction mixture was then added into a PD-10 column which was washed by 50 mM phosphate buffer (pH 7.0). The solution eluted from PD-10 column by 50 mM phosphate buffer (pH 7.0) was monitored by UV spectrum. The solution (2.5 mL) containing CF-, CR- and DBCO-modified dextran was then added into a PD-10 column which was washed by H_2O. Target product was eluted by H_2O and collected for

lyophilization. The recovery yield of dextran from PD-10 was considered as 100% to be used for next step reaction.

Cathepsin probes were conjugated on dextran through copper-free click reaction. CtBsub or CtDsub (30 nmol, 6 eq.) were dissolve in DMF (30 μL) and mixed with 50 mM Phosphate buffer (pH 7.0, 40 μL) containing DBCO-modified dextran (5 nmol, 1 eq.). The reaction solution was incubation at 37 ℃ for 4 hours. And 50 mM phosphate buffer (pH 7.0, 400 μL) was used to dilute the reaction solution and added into a PD-10 colume. And the CtBsub- or CtDsub-conjugated dextran was processed for twice purification via PD-10 eluted by 50 mM phosphate buffer (pH 7.0) and H_2O as described above. Finally, CtBsub- or CtDsub-conjugated dextran was collected and lyophilized.

5.4.4 Fluorescence measurements and FRET analyses

All *in vitro* experiments were performed with an Infinite 200 PRO microplate reader (TECAN Austria GmbH) at 37°C. Enzymatic assays were carried out in 100 mM sodium acetate (pH 4.0-5.5), 1 mM EDTA, 0.02% Tween20 and 40 mM Phosphate buffer (pH 6.0-7.0) 1 mM EDTA, 0.02% Tween20 at 37°C. The cathepsin enzyme was activated in a buffer containing 100 mM sodium acetate, 1 mM EDTA, 1 mM DTT (pH 5.5) for 15 min under 37°C. p*K*a measurement was proceeded in different buffers depended on the required pH: 40 mM citrate (pH 3.0 to 5.5), 40 mM MES (pH 6.0 to 6.5), 40 mM HEPES (pH 7.0 to 8.0) or 40 mM CHES (pH 8.5 to 10.0). CtBsub-dextran and CtDsub-dextran were monitored at 620 nm (excitation) and 670 nm (emission) wavelengths (gain 130).

5.4 References

(1) Sood, A.; Gupta, A.; Agrawal, G. Recent Advances in Polysaccharides Based Biomaterials for Drug Delivery and Tissue Engineering Applications. *Carbohydr. Polym. Technol. Appl.* **2021**, *2*, 100067.

(2) Shelke, N. B.; James, R.; Laurencin, C. T.; Kumbar, S. G. Polysaccharide Biomaterials for Drug Delivery and Regenerative Engineering. *Polym. Adv. Technol.* **2014**, *25*, 448–460.

(3) Basu, A.; Kunduru, K. R.; Abtew, E.; Domb, A. J. Polysaccharide-Based Conjugates for Biomedical Applications. *Bioconjug. Chem.* **2015**, *26*, 1396–1412.

(4) Lin, X. P.; Mintern, J. D.; Gleeson, P. A. Macropinocytosis in Different Cell Types: Similarities and Differences. *Membranes (Basel).* **2020**, *10*, 1–21.

(5) Li, L.; Wan, T.; Wan, M.; Liu, B.; Cheng, R.; Zhang, R. The Effect of the Size of Fluorescent Dextran on Its Endocytic Pathway. *Cell Biol. Int.* **2015**, *39*, 531–539.

(6) Varshosaz, J. Dextran Conjugates in Drug Delivery. *Expert Opin. Drug Deliv.* **2012**, *9*, 509–523.

(7) Turk, V.; Stoka, V.; Vasiljeva, O.; Renko, M.; Sun, T.; Turk, B.; Turk, D. Cysteine Cathepsins: From Structure, Function and Regulation to New Frontiers. *Biochim. Biophys. Acta - Proteins Proteomics* **2012**, *1824*, 68–88.

(8) Olson, O. C.; Joyce, J. A. Cysteine Cathepsin Proteases: Regulators of Cancer Progression and Therapeutic Response. *Nat. Rev. Cancer* **2015**, *15*, 712–729.

(9) Roshy, S.; Sloane, B. F.; Moin, K. Pericellular Cathepsin B and Malignant Progression. *Cancer Metastasis Rev.* **2003**, *22*, 271–286.

(10) Verdoes, M.; Oresic Bender, K.; Segal, E.; Van Der Linden, W. A.; Syed, S.; Withana, N. P.; Sanman, L. E.; Bogyo, M. Improved Quenched Fluorescent Probe

for Imaging of Cysteine Cathepsin Activity. *J. Am. Chem. Soc.* **2013**, *135*, 14726–14730.

(11) Mikhaylov, G.; Klimpel, D.; Schaschke, N.; Mikac, U.; Vizovisek, M.; Fonovic, M.; Turk, V.; Turk, B.; Vasiljeva, O. Selective Targeting of Tumor and Stromal Cells by a Nanocarrier System Displaying Lipidated Cathepsin b Inhibitor. *Angew. Chem.* **2014**, *126*, 10241 –10245

(12) Aggarwal, N.; Sloane, B. F. Cathepsin B: Multiple Roles in Cancer. *Proteomics - Clin. Appl.* **2014**, *8*, 427–437.

(13) Mort, J. S.; Buttle, D. J. Cathepsin B. *Int. J. Biochem. Cell Biol.* **1997**, *29*, 715–720.

(14) Yoon, M. C.; Solania, A.; Jiang, Z.; Christy, M. P.; Podvin, S.; Mosier, C.; Lietz, C. B.; Ito, G.; Gerwick, W. H.; Wolan, D. W.; et al. Selective Neutral PH Inhibitor of Cathepsin B Designed Based on Cleavage Preferences at Cytosolic and Lysosomal PH Conditions. *ACS Chem. Biol.* **2021**, *16*, 1628–1643.

(15) Giusti, I.; D'Ascenzo, S.; Millimaggi, D.; Taraboletti, G.; Carta, G.; Franceschini, N.; Pavan, A.; Dolo, V. Cathepsin B Mediates the pH-Dependent Proinvasive Activity of Tumor-Shed Microvesicles. *Neoplasia* **2008**, *10*, 481–488.

(16) Premzl, A.; Kos, J. Cysteine and Aspartic Proteases Cathepsins B and D Determine the Invasiveness of MCF10A NeoT Cells. *Radiol. Oncol.* **2003**, *37*, 241–248.

(17) Aghdassi, A. A.; John, D. S.; Sendler, M.; Ulrich Weiss, F.; Reinheckel, T.; Mayerle, J.; Lerch, M. M. Cathepsin d Regulates Cathepsin b Activation and Disease Severity Predominantly in Inflammatory Cells during Experimental Pancreatitis. *J. Biol. Chem.* **2018**, *293*, 1018–1029.

(18) Wille, A.; Gerber, A.; Heimburg, A.; Reisenauer, A.; Peters, C.; Saftig, P.;

Reinheckel, T.; Welte, T.; Bühling, F. Cathepsin L Is Involved in Cathepsin D Processing and Regulation of Apoptosis in A549 Human Lung Epithelial Cells. *Biol. Chem.* **2004**, *385*, 665–670.

(19) Laurent-Matha, V.; Derocq, D.; Prébois, C.; Katunuma, N.; Liaudet-Coopman, E. Processing of Human Cathepsin D Is Independent of Its Catalytic Function and Auto-Activation: Involvement of Cathepsins L and B. *J. Biochem.* **2006**, *139*, 363–371.

(20) Nakata, E.; Hirose, H.; Gerelbaatar, K.; Arafiles, J. V. V.; Zhang, Z.; Futaki, S.; Morii, T. A Facile Combinatorial Approach to Construct a Ratiometric Fluorescent Sensor: Application for the Real-Time Sensing of Cellular PH Changes. *Chem. Sci.* **2021**, *12*, 8231–8240.

(21) Poreba, M.; Groborz, K.; Vizovisek, M.; Maruggi, M.; Turk, D.; Turk, B.; Powis, G.; Drag, M.; Salvesen, G. S. Fluorescent Probes towards Selective Cathepsin B Detection and Visualization in Cancer Cells and Patient Samples. *Chem. Sci.* **2019**, *10*, 8461–8477.

(22) Tung, C. H.; Bredow, S.; Mahmood, U.; Weissleder, R. Preparation of a Cathepsin D Sensitive Near-Infrared Fluorescence Probe for Imaging. *Bioconjug. Chem.* **1999**, *10*, 892–896.

(23) Yasuda, Y.; Kageyama, T.; Akamine, A.; Shibata, M.; Kominami, E.; Uchiyama, Y.; Yamamoto, K. Correction: Characterization of New Fluorogenic Substrates for the Rapid and Sensitive Assay of Cathepsin E and Cathepsin D *J. Biochem.* **1999**, *126*, 260.

(24) Ofori, L. O.; Withana, N. P.; Prestwood, T. R.; Verdoes, M.; Brady, J. J.; Winslow, M. M.; Sorger, J.; Bogyo, M. Design of Protease Activated Optical Contrast Agents That Exploit a Latent Lysosomotropic Effect for Use in

Fluorescence-Guided Surgery. *ACS Chem. Biol.* **2015**, *10*, 1977–1988.

(25) Han, J.; Burgess, K. Fluorescent Indicators for Intracellular PH. *Chem. Rev.* **2010**, *110*, 2709–2728.

(26) Liaudet-Coopman, E.; Beaujouin, M.; Derocq, D.; Garcia, M.; Glondu-Lassis, M.; Laurent-Matha, V.; Prébois, C.; Rochefort, H.; Vignon, F. Cathepsin D: Newly Discovered Functions of a Long-Standing Aspartic Protease in Cancer and Apoptosis. *Cancer Lett.* **2006**, *237*, 167–179.

(27) Yoshida, A.; Ohta, M.; Kuwahara, K.; Cao, M. J.; Hara, K.; Osatomi, K. Purification and Characterization of Cathepsin B from the Muscle of Horse Mackerel Trachurus Japonicus. *Mar. Drugs* **2015**, *13*, 6550–6565.

(28) Yoon, M. C.; Christy, M. P.; Phan, V. V.; Gerwick, W. H.; Hook, G.; O'Donoghue, A. J.; Hook, V. Molecular Features of CA-074 PH-Dependent Inhibition of Cathepsin B. *Biochemistry* **2022**, *61*, 228–238.

(29) Hoogendoorn, S.; Habets, K. L.; Passemard, S.; Kuiper, J.; Van Der Marel, G. A.; Florea, B. I.; Overkleeft, H. S. Targeted PH-Dependent Fluorescent Activity-Based Cathepsin Probes. *Chem. Commun.* **2011**, *47*, 9363–9365.

(30) Conus, S.; Pop, C.; Snipas, S. J.; Salvesen, G. S.; Simon, H. U. Cathepsin D Primes Caspase-8 Activation by Multiple Intra-Chain Proteolysis. *J. Biol. Chem.* **2012**, *287*, 21142–21151.

CHAPTER 6

General conclusion

In this thesis, different conjugation strategies for orthogonally and simultaneously assembling functional units on biomacromolecular assemblies were used for evaluating cellular functions. Specially, I took advantage of chemoselective reaction and recognition-driven DNA-protein conjugation to (1) spatially arrange more than four types of proteins on a DNA scaffold, (2) assemble FRET-based cathepsin sensors on a DNA scaffold, and (3) functionalize polysaccharides with pH sensors and cathepsin sensors. And these studies were based on two prominent merits of biomacromolecular assemblies: bearing abundant reactive moieties and recognition modules for chemical modification and specific interaction.

In Chapter 2, I developed a set of modular adaptors (MA) and their substrates based on SNAP-tag to realize recognition-driven DNA-protein conjugation. Owing to the extremely high reaction rate between SNAP-tag and its original substrate benzylguanine (BG), two BG derivatives, i.e., benzylinosine (BI) and 7-deaza-benzylguanine (deBG), were designed by altering the interaction sites of BG with SNAP-tag. Both substrates exhibited the reduced rate constants for the reaction with SNAP-tag. For further characterization of the sequence-selective reaction of MA, BI was incorporated into DNA to elucidate the kinetic details of sequence-selective modification by MA consisting of SNAP-tag (MA-SNAP). As a result, I demonstrated sequence-selectivity of MA-SNAP with BI-modified DNA from three aspects: sequence-selective conjugation among multiple DNA sequences; estimation of all kinetic parameters to conclude $k_{cov} \ll k_{off}$; proportional relationship between k_{app} and K_D. Those results strongly proved that k_{cov} of MA reaction is a vital key for achieving the recognition-driven conjugation between MA and DNA.

Chapter 3 focused on the strategy that can spatially arrange over four unique MAs on a DNA scaffold was verified. Considering that high discrimination of BI against

CLIP-tag prevented the BI-modified ODN from nonspecific reaction with MA-CLIP. By three types of DNA binding sequence modified with BI or benzylcytosine (BC), the chemoselectivity between SNAP- and CLIP-tags allowed us to simultaneously use six unique MAs, each three MA-SNAPs and MA-CLIPs, which doubled the number of orthogonal modular adaptors. With the efficient and specific DNA-protein conjugation mediated by MA, both sequence-selective conjugation and chemoselective conjugation of MA could be used to locate proteins of interest on the same DNA scaffold. A breakthrough result was accomplished experimentally when two orthogonal series of MA were applied, where four different MAs were orthogonally located at the respective target sites on a DNA scaffold. This result provided the technical basis for constructing enzyme cascades consisting of at least four different enzymes on a DNA scaffold.

In Chapter 4, orthogonal FRET-based probes to detect cathepsin B (CtB) and D (CtD) were conjugated with a DNA handle for their modification on a biomacromolecular scaffold, such as the DNA scaffold. Coupling the CtB sensor with the CtD sensor enables clarification of the correlation between the activity of CtB and CtD. Peptidic substrates for CtB and CtD were designed with orthogonal FRET pairs and showed sufficient reactivity and specificity to detect the respective target enzymes. The DNA handle conjugated with FRET-based cathepsin probes showed drastic improvement of the reactivity toward cathepsins over their parent probes. Significantly, CtBsub-ODN1 was successfully assembly on DNA scaffold and DNA scaffold with probe binding was able to detect activity of CtB even the concentration of DNA scaffold was very less. To use the cathepsin sensor with a pH sensor simultaneously, I modified SNARF fluorophores with maleimido, which enabled the SNARF fluorophore to be modified to a DNA handle. The resulting SNARF fluorophore modified with a DNA handle would be further assembled on a DNA scaffold as a ratiometric fluorescent pH probe.

Chapter 5 emphasized the modification of multiple sensors on dextran. I presented a dual-reporter approach to the cellular imaging of activated CtB in the lysosome. Because CtB undergoes autocatalytic activation in the acidic environment of lysosomes and can also be activated through the proteolytic role of active CtD. Therefore, cell-permeable amino-coated 70 kDa dextran were modified by A ratiometric fluorescent pH probe and FRET-based cathepsin probe. FRET pair allows orthogonal fluorescence detection coupled with pH probe. Through a series of investigation, cathepsin sensor and pH sensor on dextran were verified to probe cathepsin and pH efficiently and independently. Therefore, enzymatic reactivity of CtB or CtD were measured under different pH conditions to show that the highest activity of CtB was around 5.5 which was higher than optimal pH for CtD (between 4.5 and 5).

The studies involved in this thesis demonstrated the crucial roles of diverse conjugation reactions for precisely functionalizing biomacromolecular assemblies with various functional units, such as fluorophores, peptides, and proteins. Especially, the spatial control of over four types of proteins on a DNA scaffold provides a new strategy for chemically investigating *in vitro* complex enzyme cascades, which provide the basis of artificial metabolic processes in the carbon-neutral society. At the same time, biomacromolecular assemblies feature good biocompatibility and multiple functions are essential platforms for disease diagnosis and drug delivery. The rational application of conjugation reactions for functionalizing biomacromolecular scaffolds would provide biocompatible and highly functionalized tools to promote further development of pharmaceutical and clinical research.

List of publications

Publications directly related to the contents of this thesis

1. RNA-peptide conjugation through an efficient covalent bond formation

Shun Nakano, Taiki Seko, Zhengxiao Zhang and Takashi Morii

Appl. Sci., **2020**, *10*, 8920-8929.

(Chapter 1)

2. Tuning the reactivity of substrate for SNAP-tag expands its application for recognition-driven DNA-protein conjugation.

Zhengxiao Zhang, Eiji Nakata, Huyen Dinh, Masayuki Saimura, Arivazhagan Rajendran, Kazunari Matsuda and Takashi Morii

Chem. - Eur. J., **2021***,27,* 18118-18128.

(Chapter 2, 3)

3. FRET-Based cathepsin probes for simultaneous detection of cathepsin B and D Activities

Zhengxiao Zhang, Eiji Nakata, Yuya Shibano and Takashi Morii

ChemBioChem., **2022**, *In press*. https://doi.org/10.1002/cbic.202200319

(Chapter 4)

4. Orthogonal FRET reporters for the real-time sensing of lysosomal proteases

Zhengxiao Zhang, Eiji Nakata, Yuya Shibano, Hisaaki Hirose, Shiroh Futaki, Takashi Morii

Manuscript in preparation

(Chapter 5)

Publications not directly related to the contents of this thesis

1. Rational design of DNA sequence-specific modular protein tag by tuning the

alkylation kinetics

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

Chem. Sci., **2019***, 10,* 9315-9325.

2. Evaluation of the role of the DNA surface for enhancing the activity of scaffolded enzymes.

Peng Lin, Huyen Dinh, Yuuki Morita, Zhengxiao Zhang, Eiji Nakata, Masahiro Kinoshita and Takashi Morii

Chem. Commun., **2021***, 57,* 3925-3928.

3. A facile combinatorial approach to construct a ratiometric fluorescent sensor: application for the real-time sensing of cellular pH changes

Eiji Nakata, Hisaaki Hirose, Khongorzul Gerelbaatar, Jan Vincent V. Arafiles, Zhengxiao Zhang, Shiroh Futaki and Takashi Morii

Chem. Sci., **2021***, 57,* 8231-8240.

List of presentations

International Conferences

1. Design of sequence specific modular adaptors by tuning the reactivity of protein-tag substrate

Zhengxiao Zhang, Eiji Nakata, Thang Minh Nguyen, Takashi Morii, the 45th International Symposium on Nucleic Acids Chemistry, November 7–9, **2018**, Kyoto, Japan

2. Design of sequence specific modular adaptors by tuning the reactivity of protein-tag substrate

Zhengxiao Zhang, Eiji Nakata, Thang Minh Nguyen, Takashi Morii, the 46th International Symposium on Nucleic Acids Chemistry, October 29–31, **2019**, Tokyo, Japan

3. **Design of DNA sequence-specific modular adaptors by tuning the reactivity of protein-tag substrate**

Zhengxiao Zhang, Eiji Nakata, Dinh Huyen, Masayuki Saimura, Kazunari Matsuda, Takashi Morii, the 48th International Symposium on Nucleic Acids Chemistry, November 10–12, **2021**, Online, Japan

Domestic Conferences

1. Design of sequence specific modular adaptors by tuning the reactivity of protein-tag substrate

Zhengxiao Zhang, Eiji Nakata, Thang Minh Nguyen, Takashi Morii, the 12th Symposium on Biorelevant Chemistry, September 9–11, **2018**, Osaka, Japan

2. Tuning the reactivity of crosslink formation for sequence-selective DNA-protein conjugation

Zhengxiao Zhang, Eiji Nakata, Thang Minh Nguyen, Takashi Morii, the 99th CSJ

Annual Meeting, March 16–19, **2019**, Kobe, Japan

3. Orthogonal FRET reporters for the real-time sensing of lysosomal proteases

Zhengxiao Zhang, Eiji Nakata, Hisaaki Hirose, Shiroh Futaki, Takashi Morii, the 102nd CSJ Annual Meeting, March 23–26, **2022**, Online, Japan