Development of New Chemical Methods toward Lectin Engineering

Yoichiro Koshi

2008

Preface

The study presented in this dissertation has been carried out under the direction of Professor Itaru Hamachi at Department of Chemistry and Biochemistry of Kyushu University form April 2003 to March 2005 and at Department of Synthetic Chemistry and Biological Chemistry of Kyoto University form April 2005 to March 2008. The study is focused on development of new chemical methods toward lectin engineering, based on organic chemistry and protein science.

The author would like to express his sincere gratitude to Porf. Itaru Hamachi for his kind guidance, valuable suggestions, and encouragement throughout this work. The author is deeply grateful to Dr. Eiji Nakata and Dr. Shinya Tsukiji for their helpful advice, discussions and encouragement. The author is also indebted to Lecturer Akio Ojida and Assistant Prof. Masato Ikeda for their helpful suggestions.

The author wishes to Prof. Yasuo Mori, Assistant Prof. Shigeki Kiyonaka, Dr. Takashi Yoshida for the guidance of molecular biology. The author acknowledges to Associate Prof. Yuu Okumoto (Kyoto Univ.) for usage of gel imager in chapter 3. The author is grateful to Dr. Kenji Usui and Prof. Hisakazu Mihara (Tokyo Institute of Technology) for their advice about the statistical analytes in chapter 3. The author thanks Associate Prof. Tomohisa Ogawa (Tohoku Univ.) for providing Congerin in chapter 2. The author also thanks Shimadzu Corporation and Dr. Maki Yamada for mass-mass analysis in chapter 2.

The author thanks the graduates of Hamachi lab, Dr. Noriyuki Kasagi, Mr. Hiroki Yamane, Mr. Hiroki Takemoto, Dr. Kei Honda and Dr. Masaki Inoue, for their advice. The author also thanks Dr. Hiroshi Tsutusmi, Dr. Shunichi Tamaru and Dr. Satoshi Yamaguchi. The author is also grateful to the member of Protein Engineering Team and other member of Hamachi lab for their helpful suggestions and hearty encouragement.

The author thanks Japan Society for the Promotion of Science for financial support.

The author expresses his deep appreciation to his parents, Mr. Tadao Koshi and Mrs. Harumi Koshi, and his grand parents for their constant assistance and encouragement.

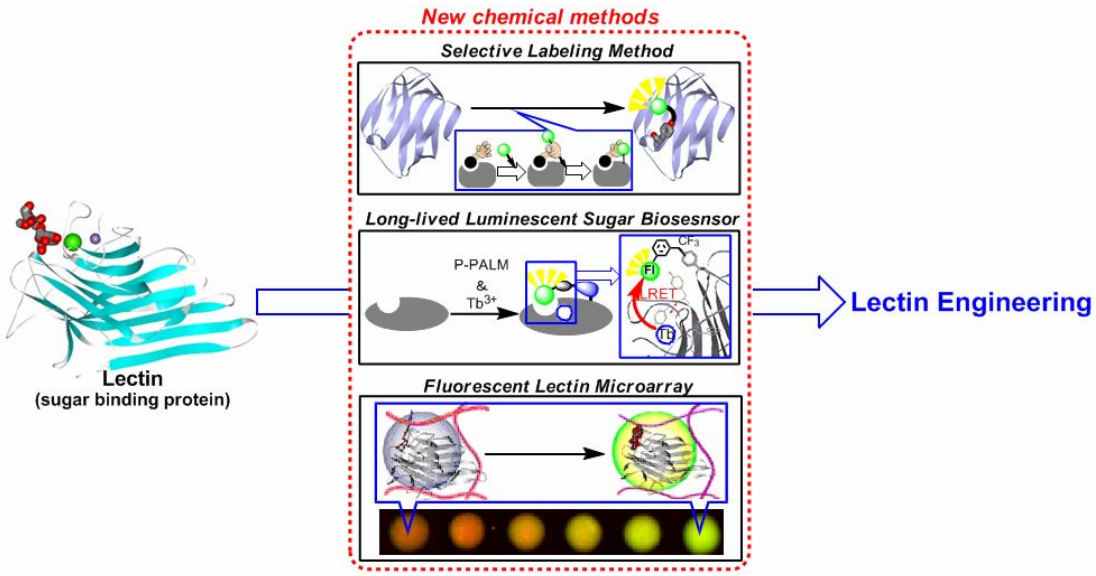
Yoichiro Koshi

March, 2008

Contents

General introduction

Advances in protein engineering enabled the functionalization of many proteins which is important in the field of molecular biology, medicine, and industry world. Recently, the importance of sugar which is regarded as a third biopolymer and sugar-binding proteins, lectins, are recognized.[1,2] By applying protein engineering method to lectins, more fruitful results are expected not only in glyco-science but also in other fields.



Protein engineering

Protein plays a vital role of biological activities, and they have important functionalities such catalysis and molecular recognition. Protein engineering is a methodology to modify and use natural or functionalized proteins based on engineering concept. Since this is an interdisciplinary research field, the development of many technologies was essential in protein engineering. For example, the improvement of X-ray crystallographic analysis and NMR technology made it possible to solve the detail structure of proteins. Gene manipulation enabled us to obtain mutants and recombinants of proteins easily. In addition, organic chemistry can be used for chemical modification of proteins in aqueous solution.

Antibody engineering

"Antibody engineering", a research to functionalize antibodies, is fundamental to "antibody therapy".[3] In antibody therapy, it is expected that cancer cells will be targeted by using the high affinity antibody-antigen reaction (Fig.1a).[4] Usually, antibodies are obtained from experimental animals such as mouse. However the antibody of the animal origin can not be used to human being because of immune rejection. To solve this problem, the immunogenicity could be reduced by the humanization of mouse monoclonal antibody or by using the human antibody producing mouse technology (Fig.1b).

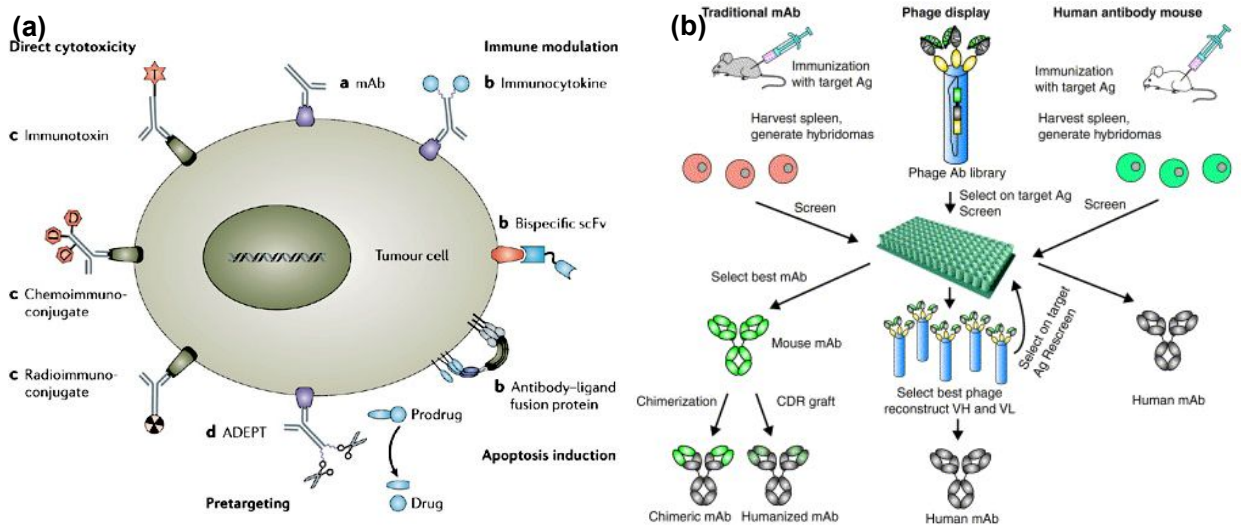


Figure 1. (a) Concepts of antibody therapy. (b)Generation of humanized antibody and human antibody.[4a]

Some action mechanisms are known to the antibody therapy (Fig.1a), and the anti-tumor activities can be increased by the fusion of anti-cancer agent. As shown in Fig.2, the modification of the chelating agent enabled radio isotope labeling (Fig2a).[5] Since the radio-isotope labeled antibody acts specifically in cancer cells, on contrast to conventional radiotherapy, the better anti-cancer results are expected by using this method. Attaching small cancer drug to antibody is another strategy in antibody therapy (Fig.2b).[6] For example, Trail and coworkers modified Doxorubicin, which is one of the small molecular anti-cancer drugs, to the antibody by hydrazon

bond.[7] After binding to cancer cell, antibody–drug conjugates were endocytosed, and the hydrazon bond would be cleaved under weakly acidic condition in lysosome. As a result, the anti-cancer drug is released inside the cell, and exerts its cytotoxic effects. Other antibody–drug conjugates using the similar strategy were reported. One example is to attach a drug molecule to antibody through disulfide bond. By doing that, disulfide bond would be cleaved under intracellular reducing conditions, releasing the drug in target cells.

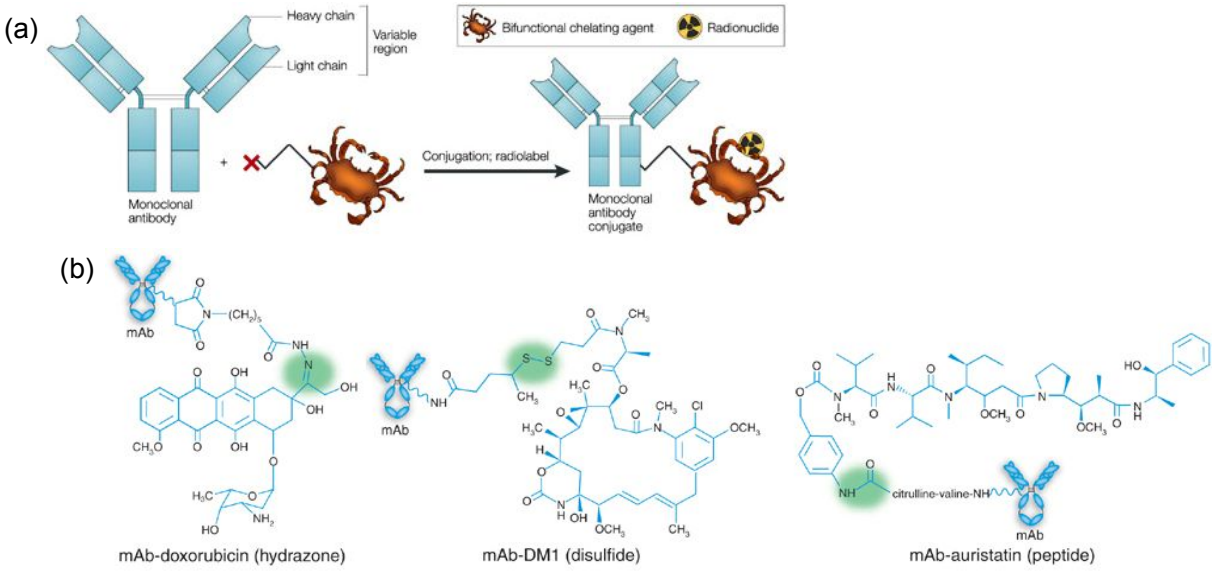


Figure 2. (a) A general view of the conjugation of chelating reagent to a monoclonal antibody.[5] (b) Chemical structure of some advanced mAb-drug conjugate. [6]

Modification of proteins including antibodies with polyethylene glycol (PEGylation) is commonly used in protein-based drug development.[8] PEGylation has several advantages of improving solubility, reducing antigenicity and protecting from proteolytic degradation (Fig.3). As a result, protein modified through PEGylation are expected to have longer drug effects in physiological systems.

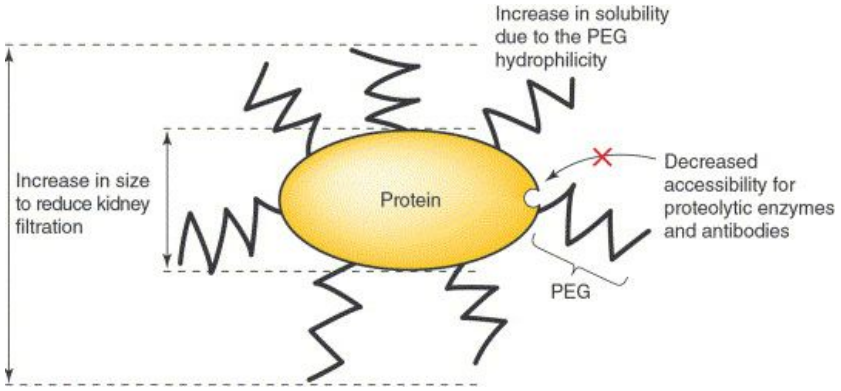


Figure 3. Main advantage of PEGylated proteins.[8a]

The engineering of fluorescent proteins.

Fluorescent proteins are known to play an important role in bio-imaging.[9a] In green fluorescent protein (GFP) which is the most well-known fluorescent protein, its fluorophore is synthesized in an auto-catalytic process in the protein interior. No co-factors or enzymatic components are required (Fig.4a). In the case of red Fluorescent Protein (RFP), a fluorophore is generated after one more oxidization step of GFP-mimic intermediate (Fig.4b). By genetically fusing fluorescent protein with a protein of interest, it is possible to monitor localization and movement of target proteins in a real time. Up to date, a large variety of fluorescent proteins with different excitation and emission wavelengths were developed (Fig.5a).[9b] In addition to the above fruitful results, a new photoconvertible fluorescent protein, named "Kaede", was developed by Miyawaki et.al.[10] Upon UV irradiation, the fluorophore of Kaede undergoes β-elimination, and the fluorescence changed from green to red (Fig.4c, 5b).

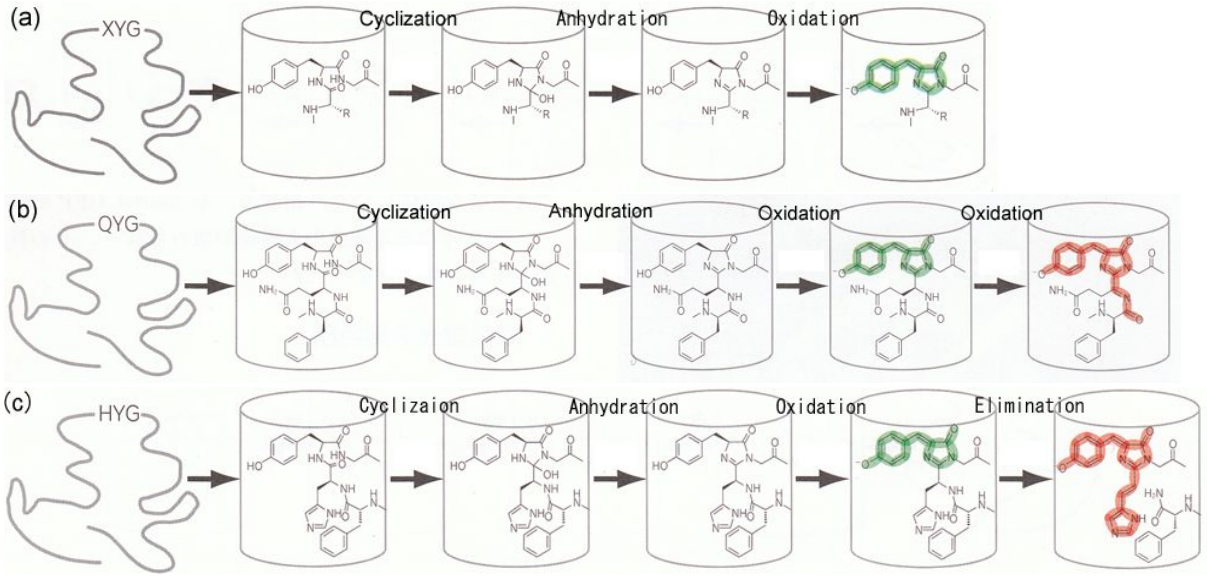


Figure 4. Chemical structure of fluorophore of fluorescent proteins (a) GFP, (b) DsRed, (c)Kaede.[9a]

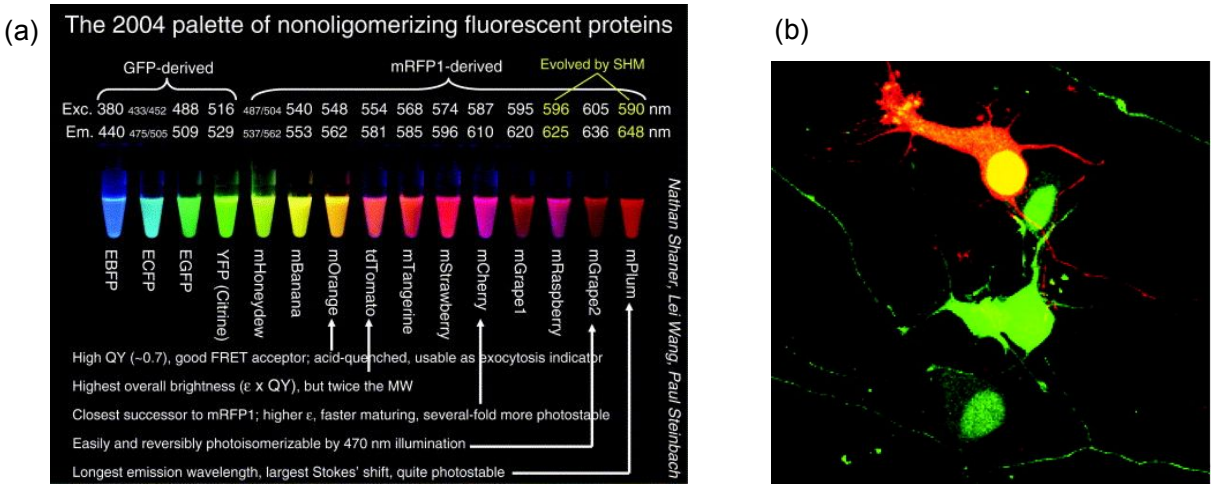


Figure 5. (a) Engineered fluorescent proteins cover the full visible spectrum of emissions.[9b]
(b) Neuron cells expressing Kaede (upper cell was irradiated by UV-pulsed).[10]

In addition to these fluorescent proteins, luciferase is another commonly used protein in bio-imaging.[11] Luciferase, a luminescent enzyme, produces bioluminescence by the oxidation of luciferin (Fig.7). The background noise during imaging can be reduced by using luciferase since no excitation light is required, unlike fluorescent proteins, to generate luminescence. The emission wavelength of bioluminescence depends on kinds of luciferase. Kato and coworkers have successfully analyzed the crystal structure of the complex of a luciferase and luciferin-analogue. Based on the information obtained from this crystal structure, they rationally tuned the emission wavelength of luminescence through luciferase mutation (Fig.8).[12] As this technology developing, selecting luminescence wavelength arbitrarily, as what we are able to do now with fluorescent proteins, might be possible in near feature.

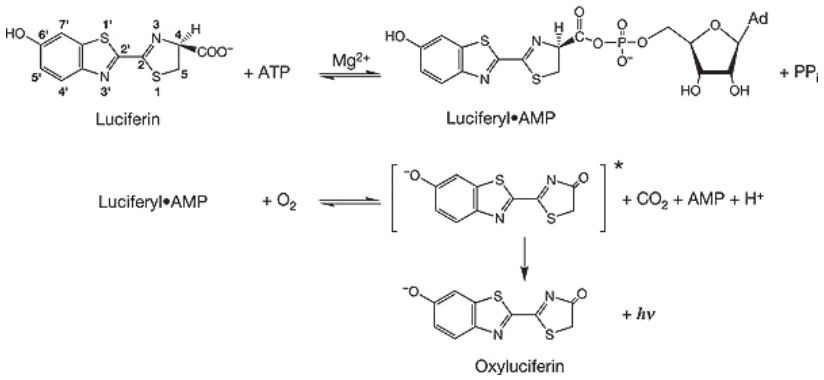


Figure 6. Luminescent mechanism of *Firefly* luciferase. [11]

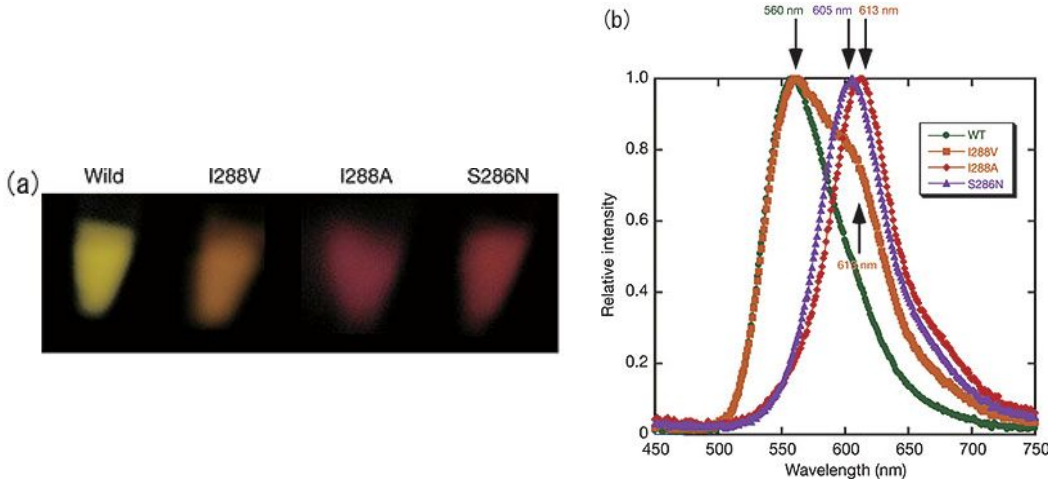


Figure 7. Bioluminescence color of wild-type and three mutant forms of Lcr luciferases. (a) Photographs of bioluminescence by a luciferase-catalyzed reaction. (b) Emission spectra of the four types of Luciola cruciata luciferases. [12]

Enzyme engineering.

Development of industrial enzymes is one of the main targets in protein engineering. Up to date, various industrial enzymes were generated.[13] One example of them is α-amylase, which is a glycosidase cleaving α1,4-glycoside bond. Termamyl™, a thermostable α-amylase, is commonly used in starch liquefaction. Disadvantage of using this enzyme is that Ca^{2+} addition and pH

adjustment are carefully required in this industrial process (Fig.8a). In order to improve the functionality of Termamyl™, mutation of this α-amylase was made and a new enzyme, Termamyl LC™, was then developed.[14] With the use of Termamyl LC™ in starch liquefaction, no Ca^{2+} is required and the enzyme works well under steady acidic conditions. As a result, efficiency of industrial starch liquefaction is greatly improved by using this engineered enzyme.

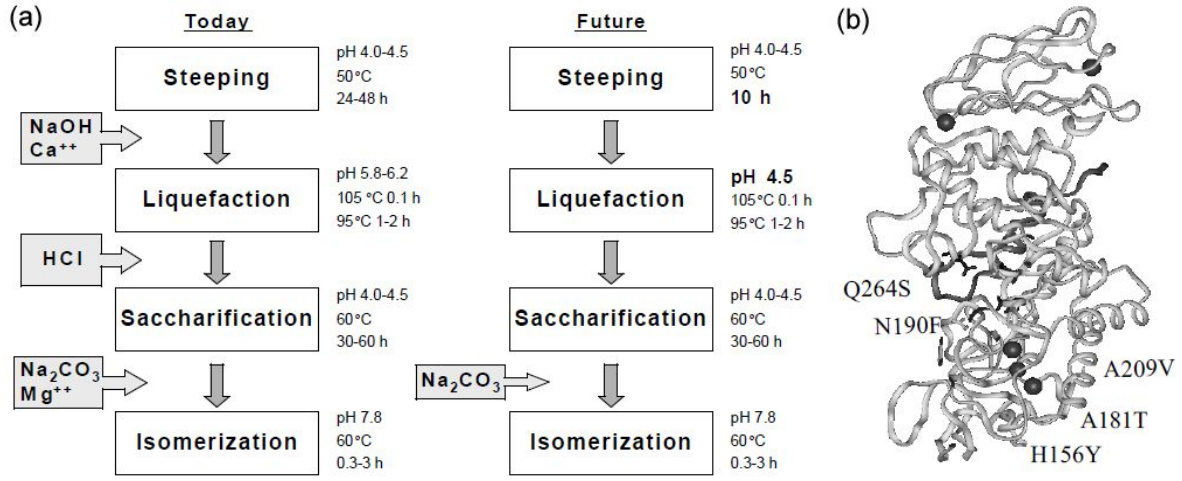


Figure 8 (a) Conversion of starch to isomerized sugar process today tomorrow. (b) A model structure of calcium independent α-amylase TermamylLC™.[14]

Other examples

Other than the example mentioned above, many applications of protein functions are known. One of these applications is the use of antifreeze protein (AFP) as an antifreeze reagent. It is known that AFP can inhibit the growth of ice crystals in the way of covering the waster-accessible surface of ice (Fig.9a).[15] AFPs were isolated from certain types of fishes. As an antifreeze reagent, AFP is known to be commercially used in preservations of cells and organs as well as food. Recently, Nishimura and coworkers have successfully synthesized an artificial AFP which is structurally similar to the functional part of natural AFP (Fig.9b).[16] In addition to these examples, the protein-based biosensors that are discussed in chapter 1 and the protein microarrays described in chapter 3 are also representative results in protein engineering.

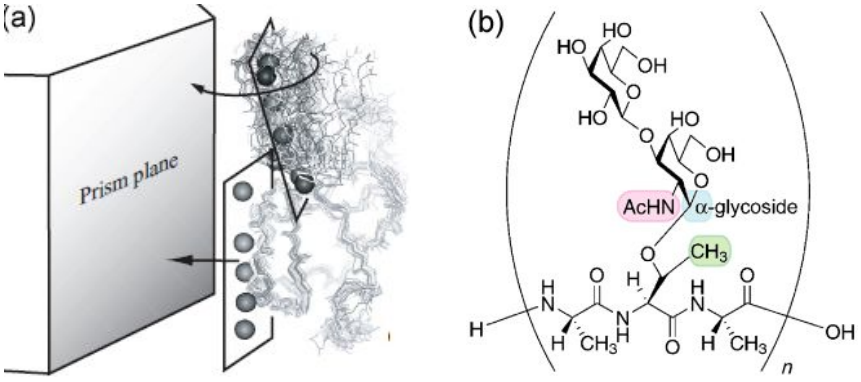


Figure 9 (a) Ice binding model of antifreeze protein. (b) Chemical structure of artificial antifreeze glycoprotein.

Lectins[2,17]

Lectins are sugar binding proteins. They are found in most of the organism such as fungi, plant, fishes, and the mammals. Interactions between lectin and sugar-chain play important physiological roles such as cell-cell communication and protein quality control. Moreover, lectins also play a very important role in glycoscience as a tool for analyzing sugar.

The function of lectins in the living body

Roles of lectins playing in physiological system vary in biological species, and greatly depend on kinds of lectins. As an example, legume lectins are crucial in maintaining the symbiotic relationship between leguminous plants and rhizobiums which can fix nitrogen (Fig.10a).[18] Rhizobiums bind to the root hair of plants through the mediation of lectins. Some other plant lectins play important roles in their defense system, by displaying toxicity to insects and herbivorous animals (Fig.10b).[19]

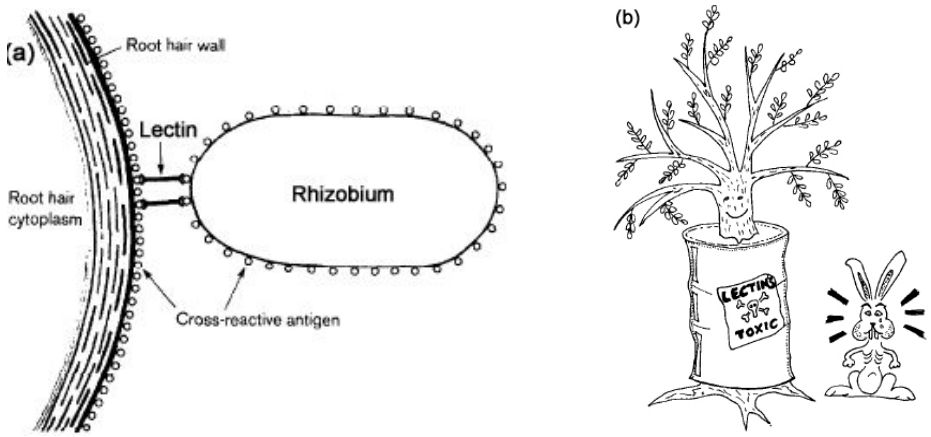


Figure 10 (a) Lectin cross binding model for root hair. [18] (b) Illustration of the protective effect of toxic. [19]

Functions of animal lectins are more diverse and complicated than the ones in plants. For example, asialo-glycoprotein receptor (AsGPR) is a calcium-dependent lectin (C-type lectin) in liver.[20] AsGPR recognizes glycoproteins which have sugar branches terminated in galactose and then induces endocytosis. It is believed that this mechanism is used as quality control of the glycoprotein in a physiological system. On the other hand, Mannan Binding Lectin (MBL), which is another type of C-type lectins, has an oligomeric structure through collagen domains (Fig.11a).[21] This kind of lectin is called "Colectin", and they are often involved in immunological activities. In blood, colectins activate the early stage of immune response, so called "lectin pathway", by the recognition of pathogens (Fig.11b).

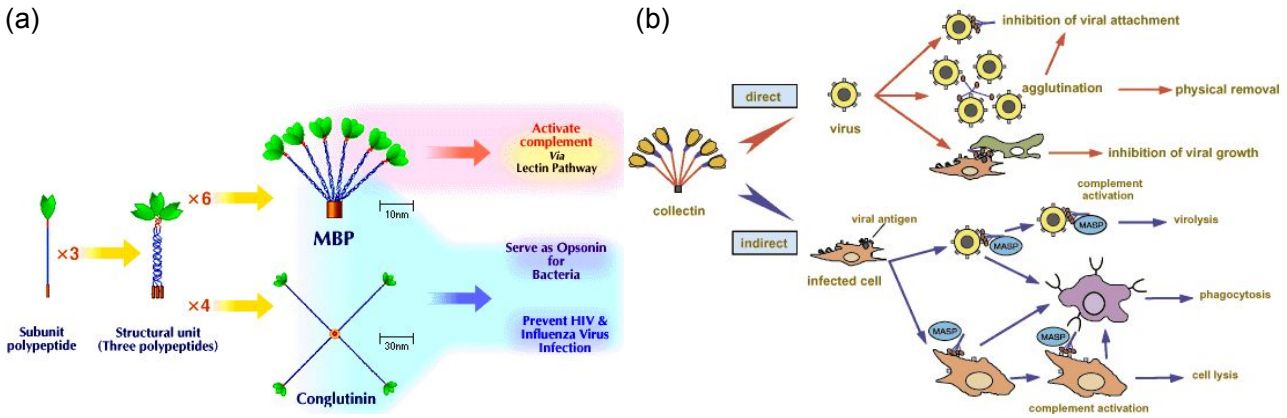


Figure 11. (a) Molecular structures and biological activities of MBL and conglutinin.[20d] (b) Anti-viral activities in collectins. [21d]

It is also known that lectins exit in virus. Hemagglutinins (HA) of influenza virus play an important role at the first step of the infection (Fig.12a).[22] Virus invasion starts from the binding between HA and sugar-chains on host cell membrane. Therefore, the specific combination of sugar chains on host cell and lectin presenting in virus is crucial in determining virus host cells (Fig.12b).

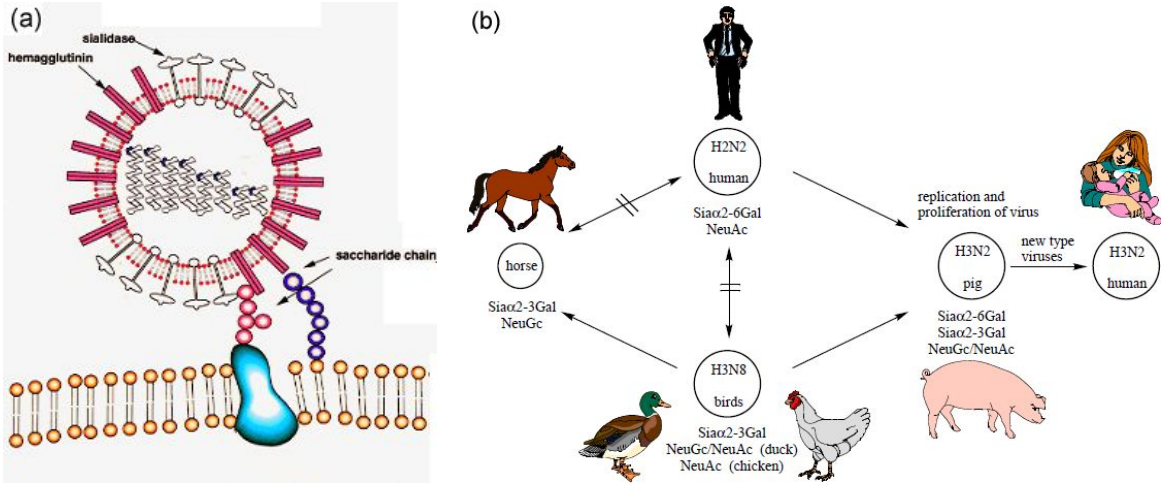


Figure 12. (a) Influenza virus binding to sugar chain on cell surface. (b) The host range of influenza a viruses is determined by combination of sialic acid species (NeuAc/NeuGc) and difference in sialyl-Gal linkages (α2-3/α2-6).

Lectin as the research tool for glycobiology

Not only that their biological activities are crucial, lectins are also important as sugar recognition tools. Since sugar-identification technologies are significant in glycoscience, some examples of chemically synthesized molecules working as sugar-recognition tools were reported. One example is the use of boronic acid as an artificial sugar receptor developed by Shinkai and coworkers (Fig.13a).[23a] Since boronic acid is capable of making a reversible ester bond with diol, it was successfully used as sugar (such as glucose) detector. Davis and coworkers recently developed

a kind of receptor that is able to recognize cellobiose through hydrophobic interaction.[23b] These examples are important in host-guest chemistry. However, it is generally hard to recognize sugars by using artificial sugar receptors with high accuracy due to difficulties involved in precisely recognizing the configurations of hydroxy groups in water.

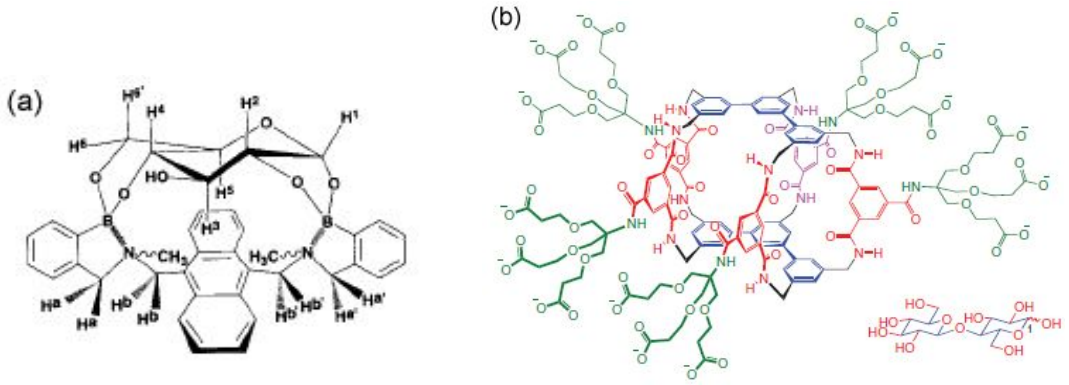


Figure 13. Molecular structure of artificial sugar receptor by Shinkai (a)[23a] or Davis (b)[23b].

On the other hand, natural lectins are capable of recognizing saccharides with high selectivity through precise hydrogen-bonding systems. A an example, Fig.14 illustrates how Concanavalin A (Con A) specifically binds to mannose by disposing precise hydrogen bonding to hydroxyl groups of mannose in binding pocket.[24] Through as technology developing, more analytical instruments such as mass spectroscopy are available today for analyzing saccharides, lectins still has some advantages of identifying complicated sugar structures.

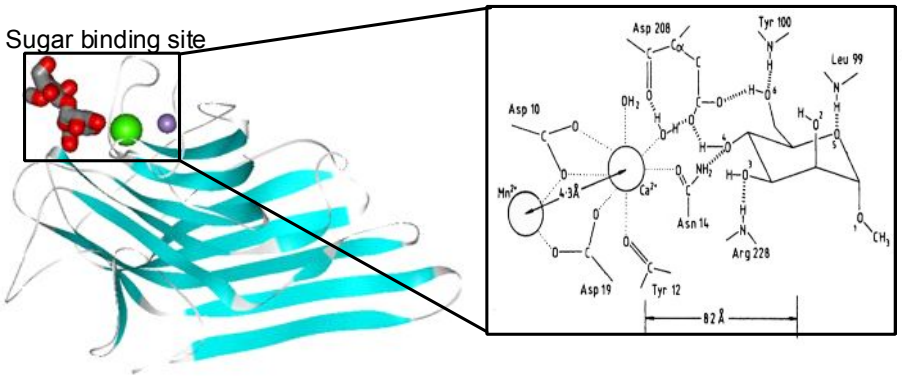


Figure 14. Structural model of sugar binding pocket of Con A.[24]

<u>Recognition of cell surface saccharide.</u>

Lectin-staining is a technique used for analyzing sugar-chain on cell and tissue surface. This technique was used in blood-type identification in the past. Recently, it is also used to diagnose cancer by recognizing abnormal sugar-chains on the cancer cell surface. For example, Helix pomatia agglutinin (HPA) which is a lectin derived from escargot is used as prognostic marker of breast cancer (Fig.15a).[25]

A method used to isolate certain types of cells by using lectin-sugar on cell surface interactions

is also studied (Fig.15b). In bone marrow transplantation, Soy bean agglutinin (SBA) can be used to remove T-cells which cause Graft Versus Host Disease (GVHD). As a result, postoperative survival rate is significantly increased. [26]

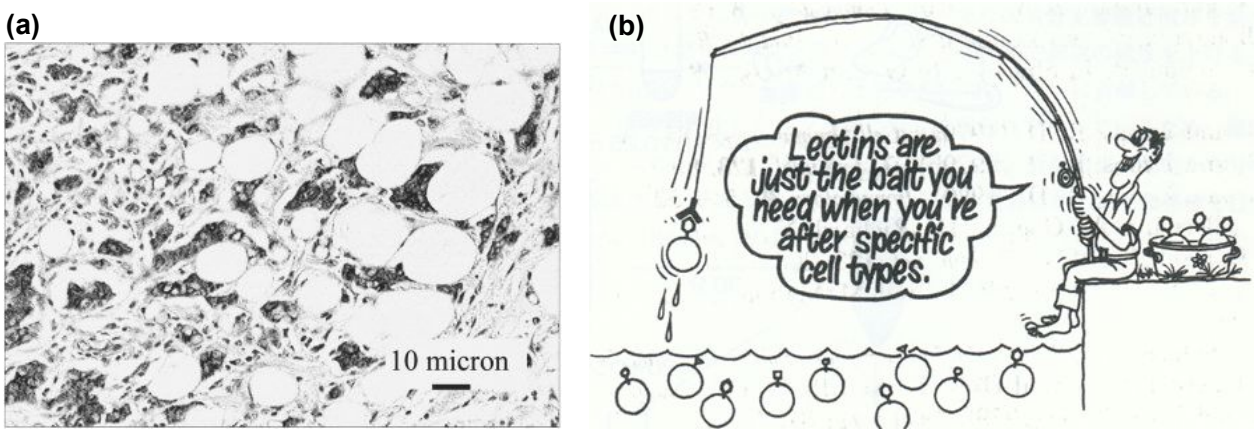


Figure 15 (a) Ductal carcinoma of the breast infiltrating fat. Specimen stained with HPA.[25] (b) The application for the cell separation. [26]

Moreover, the lectin-sugar interaction is also used in drug-delivery systems. As shown in Fig.16a, drugs attached to lectins can be delivered to the cells of interest through specific interaction between lectin and sugar-chains presenting on target cell surface.[27] By using this method, Wheat Germ Agglutinin (WGA) modified anti-cancer agent was successfully introduced into human colonic adenocarcinoma cells (Caco 2). Reversely, drug attached to sugars would also be delivered to cells that have lectins binding to those specific drug-modified sugar chains (Fig.16b). It has been reported that the asialoglycoprotein receptor and a macrophage lectin are used as a target in the above drug-delivery system.

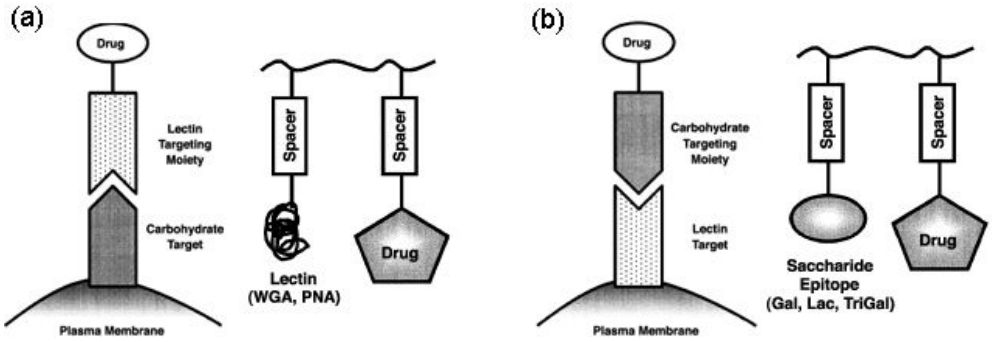


Figure 16. Drug delivery system by direct (a) and reverse (b) lectin targeting.[27a]

Recognition of glyco-conjugates

Lectins are used to analyze not only sugars on cell surfaces but also glyco-conjugates. "Lectin blotting" is a well-known method for analyzing glycoproteins (Fig.17).[28] In this method, the labeled-lectins specifically bind to a glycoprotein of interest that is blotted on the membrane. Upon the treatment of glycosidase on the membrane, the target glycoproteins would be cleaved off for further study and analysis.

"Lectin affinity column" is known as conventional purification method of glycoproteins. Particular glycoproteins are able to be purified upon binding to the types of lectins immobilized to column. As an example, Hirabayashi and coworkers immobilized lectins to Frontal Affinity Chromatography (FAC).[29] By using FAC, the sample-lectin interactions could be quantitatively analyzed according to column exit velocity. They systematized this method and have succeeded in analyzing their lectin-sugar pairs.

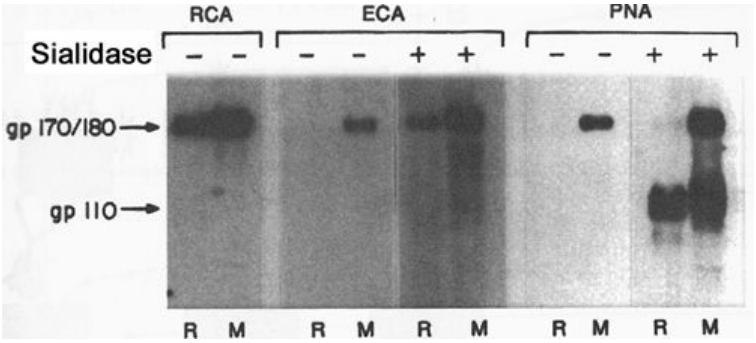


Figure 17. Lectin blotting of the thymus homogenate of Rat (R) or Mouse (M).[28]

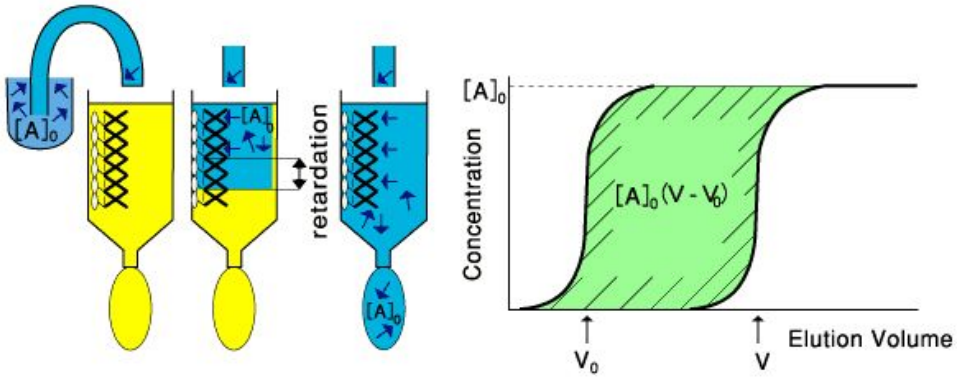


Figure 18. Principle of FAC: "A" (e.g., sugar) is continuously applied to a column, where affinity ligand "B" (e.g., lectin) is immobilized. If "A" has affinity to "B," elution front of "A" is retarded.[29a]

Summary of this thesis

As mentioned above, lectins play important roles not only in physiological activities but also in glyco-research as a sugar recognition tool. In recent glyco-biology and glyco-science, the importance of lectins is increasing. Thus, the methodologies to modify and use lectins, "lectin engineering", are essential in facilitating glyco-science. In this thesis, I developed new chemical methods for functionalizing lectins, which can apply to the biological and chemical researches of sugar-derivatives. As protein engineering is new valuably used in many fields, it is expected that lectin engineering well contributes not only to fundamental researches, but also to all fields involving sugar such as medicine and industry.

This thesis consists three chapters. Each chapter is summarized as follows.

Chapter 1

A new luminescent sugar biosensor was constructed on the basis of lectin by incorporating the lanthanide-based luminescence. By using luminescent resonance energy transfer (LRET) between the coordinated terbium ion and the fluorophore attached on lectin, I successfully detected sugars and its derivatives. Thank to the benefit of long-lived luminescence, not only sensing the sugars in the presence of fluorescent impurities but also the luminescent monitoring trimming process of the glycoprotein were carried out.

Chapter 2

A method for selective chemical labeling based on a specific protein-ligand interaction is useful for the exploration of new receptor proteins and the functionalization of proteins. I developed a new method for lectin-selective covalent labeling using sugar-tethered acyl transfer catalyst. This method was successfully applied not only to a purified lectin, but also to crude mixtures, such as lysates of recombinant bacterial cells or mammalian tissue expressing the target lectin.

Chapter 3

This chapter described a new microarray system to detect sugar derivatives. By immobilizing fluorophore labeled-lectin in a matrix of supramolecular hydrogel, a fluorescent lectin array was successfully constructed. This lectin microarray enabled not only sensing simple saccharides but also analyzing glycoproteins and profiling of several types of cell lysates based on the kinds of sugar included in the cell lines.

Reference

1) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. *Essentials of Glycobiology;* Cold Spring Harbor Laboratory Press: New York, 1999.

2) Sharon, N.; Lis, H. *Lectins (second edition)*; Springer: Netherlands, 2003

3) Nishimoto, N. *Cell Technology* , **2007**, *26*, 250-253. (b) Casadevall, A.; Dadachova, E.; Pirofski, L.-a. *Nature Rev. Microbiology* **2004**, *6*, 695-703.

4) (a) Dijk M. A.; Winkel, J. *Current Opinion in Chem. Biol.* **2001**, *5*, 368-374. (b) Kazuma, T.; Hitoshi, Y.; Yochimi, K.; Ishida, I. *Cell Technology* , **2007**, *26*, 258-262.

5) Milenic, D. E.; Brady, E. D.; Brechbiel, M. W. *Nature Rev. Drug Discovery* **2004**, *3*, 488-498.

6) Wu, A. M.; Senter, P.D. *Nature Biotech.* **2005**, *23*, 1137-1146.

7) Trail, P. A.; Willner, D.; Lasch, S. J.; Henderson, A. J.; Hofstead, S.; Casazza, A. M.; Firestone, R. A.; Hellstrom, I.; Hellstrom, K. E. *Science* **1993**, *261*, 212-215.

8) (a) Veronese, F. M.; Pasut, G. *Drug Discovery Today* **2005**, *10*, 1451-1458. (b) Harris, J. M.; Chess, R. B. *Nature Rev. Drugdiscovery* **2003**, *2*, 214-221.

9) (a) Tsien, R. Y. *FEBES Lett.* **2005**, *579*, 927-932. (b) Miyawaki, A. *Protein Nucleic acid and Enzyme*, **2007**, *52*, 1558-1562.

10) (a) Ando, R.; Hama H, Yamamoto-Hino, M.; Mizuno, H.; Miyawaki, A. *Proc. Natl. Acad. Sci. USA* **2002**, *99*,12651-12656. (b) Mizuno, H.; Mal, T. K.; Tong, K. I.; Ando, R.; Furuta, T.; Ikura, M.; Miyawaki, A. *Mol. Cell.* **2003**, *12*, 1051-1058. (c) Hayashi, I.; Mizuno, H.; Tong, K. I.; Furuta, T.; Tanaka, F.; Yoshimura, M.; Miyawaki, A.; Ikura, M. *J. Mol. Biol.* **2007**, *28*, 918-926.

11) Pfleger, K. D. G.; Eidne, K. A. *Biochem. J.* **2005**, *385*, 625-637.

12) Nakatsu, T.; Ichiyama, S.; Hiratake, J.; Saldanha, A.; Kobashi, N.; Sakata, K. *Nature,* **2006**, *440*, 372-376.

13) Svendsen, A. *Biochimica et Biophysica Acta,* **2000**, *1543*, 223-238.

14) Hashida, M.; and Bisgaard-Frantzen, H., *Trends in Glycoscience and Glycotechnology,* **2000**, *12,* 389–401.

15) (a) Tanaka, S.; Kohashigawa, K.;, Miura, K.;, Nishimiya, K.; Miura, A.; Tuda S. *Seibutsu Butsuri,* **2003**, *43*, 130-135. (b) Davies, P. L.; Baardsnes, J.; Kuiper, M. J.; Walker, V. K. *Phil. Trans. R. Soc. Lond. B* **2002**, *357*, 927–935. (c) Harding, M. M.; Anderberg, P. I.; Haymet, A. D. J. *Eur. J. Biochem.* **2003**, *270*, 1381–1392.

16) Tachibana, Y.; Fletcher, G. L.; Fujitani, N. *Angew. Chem. Int. Ed.* **2004**, *43*, 856 –856.

17) (a) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357. (b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.

18) Hirsch, A. H. *Current Opinion in Plant Biology* **1999**, *2*, 320-326.

19) Peumans, W. J.; Van Damme, E. J. M. *Plant Physiol.* **1995**, *109*, 347-352.

20) (a) Spiess, M. *Biochemistry* **1990**, *29*, 10018-10022. (b) Baenzinger, J. U.; Maynard, Y. *J. Biol. Chem.* **1979**, *255*, 4607-4613. (c) Stocker, R. J.; Morell, A. G. *J. Biol. Chem.* **1990**, *265*, 1841-1846. (d) http://www.glycoforum.gr.jp/science/word/lectin/LEA06J.html

21) (a) Garred, P.; Larsen, F.; Madsen, H. O.; Koch, C. *Molecular Immunology* **2003**, *40*, 73-84. (b)

Holmskov, U.; Malhotra, R.; Sim, R. B.; Jensenius, J. C. *Immunology Today* **1994**, *15*, 67-74. (c) Lu, J.; Thiel, S.; Wiedemann, H.; Timple, R.; Reid, K. B. M. *J. Immunol.* **1990**, *144*, 2287-2294. (d) http://www.glycoforum.gr.jp/science/word/lectin/LEB02J.html

22) Ando, T.; Ando, H.; Kiso, M. *Trends in Glycoscience and Glycotechnology* **2001**, *13*, 573–586.

23) (a) James, T. D.; Samankumara Sandanayake, K. R. A.; Shinkai, S. *Ang*ew. *Chem. Int. Ed.* **1996**, *35*, 1910-1922. (b) Ferrand, Y.; Crump, M. P.; Davis, A. P. *Science* **2007**, *318*, 619-622.

24) (a) Bouckaert, J.; Hamelryck, T. W.; Wyns L.; Loris, R. *J. Biol. Chem.* **1999**, *274*, 29188-29195. (b) Loris, R.; Stas, P. P. G.; Wyns, L. *J. Biol. Chem.* **1994**, *269*, 26722-26733.

25) Dweck, M. V.; Ross, H. A.; Street, A. J.; Brooks, S. A.; Adam, E.; Titcomb,; Woodside J. V.; Schumacher, U.; Leathe, A. J., *Int. J. Cancer* **2001,** *95*, 79–85.

26) Reisner, Y.; Kapoor, N.; Kirkpatrick, D.; Pollack, M. S.; Cunningham-Rundles, S.; Dupont, B.; Hodes, M. Z.; Good, R. A.; O'Reilly, R. J. *Blood* **1983**, *61,* 341-348.

27) (a) Minko, T. *Advanced Drug Delivery Reviews* **2004**, *56*, 491–509. (b) Yamazaki,N.; Jigami,Y.; Gabius,H.-J.; Kojima,S. *Trends in Glycoscience and Glycotechnology*, **2001**, 13, 319–329. (c) Yamazakia, N.; Kojimab, S.; Bovinc, N. V.; Andre´d, S.; Gabiuse, S.: Gabiusd, H.-J. *Advanced Drug Delivery Reviews* **2000**, *43*, 225–244.

28) Gershoni, J. *Methods in Cell Biology* **1994**, *3*, 323-331.

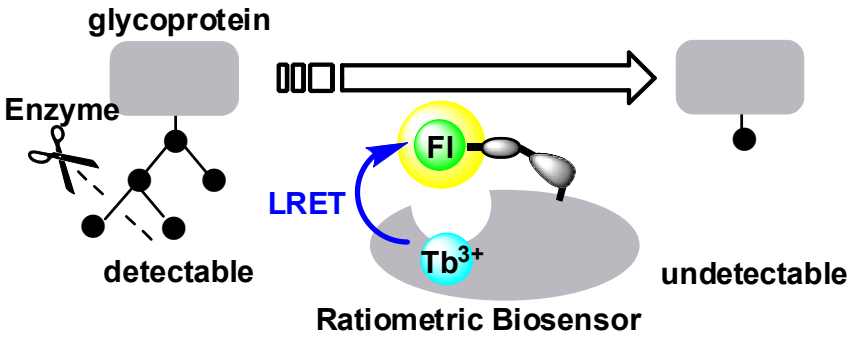
29) (a) http://www.glycoforum.gr.jp/science/word/glycotechnology/GT-C07J.html (b) Hirabayashi, J. *Glycoconjugate J.*, **2004**, *21*, 35–40. (c) Hirabayashi J.; Arata Y.; Kasai, Y. *J. Chromatogr. A*, **2000,** *890*, 261-271.

Chapter 1

Long-lived Luminescent Sugar Biosensor by using Luminescent Resonance Energy Transfer (LRET) on Protein Surface

Abstract

A new luminescent biosensor for complicated glyco-conjugates was engineered on the basis of a lanthanide-complexed sugar-binding protein (lectin) modified with a fluorophore. Using luminescence resonance energy transfer (LRET) within the engineered protein, ratiometric luminescent sensing can be carried out and furthermore, it is successfully applied to a luminescent assay for an enzymatic trimming toward a glycoprotein.



1-1. Introduction

1-1-1. Biosensors based on fluorescent proteins such as GFP

Fluorescent biosensors using fluorescent proteins have been actively studied. In a pioneering work, Miyawaki and coworkers developed a new fluorescent indicator for Ca^{2+} (*Cameleon*) based on a Ca^{2+} binding protein domain (*Calmodulin*).[1] The Ca^{2+}-induced conformational changes of *calmodulin* were utilized to detect Ca^{2+} by Fluorescent Resonance Energy Transfer (FRET) between the tethered fluorescent proteins such as YFP and CFP (Fig.1a). The rational fusion of fluorescent proteins with a ligand-binding protein domain is a representative of strategies to construct fluorescent biosensors. On the basis of this strategy, the fluorescent indicators for small molecules such as cGMP (Fig.1c)[2] and the fluorescent probes for post-translational modification such as phosphorylation (Fig.1b)[3] were reported.

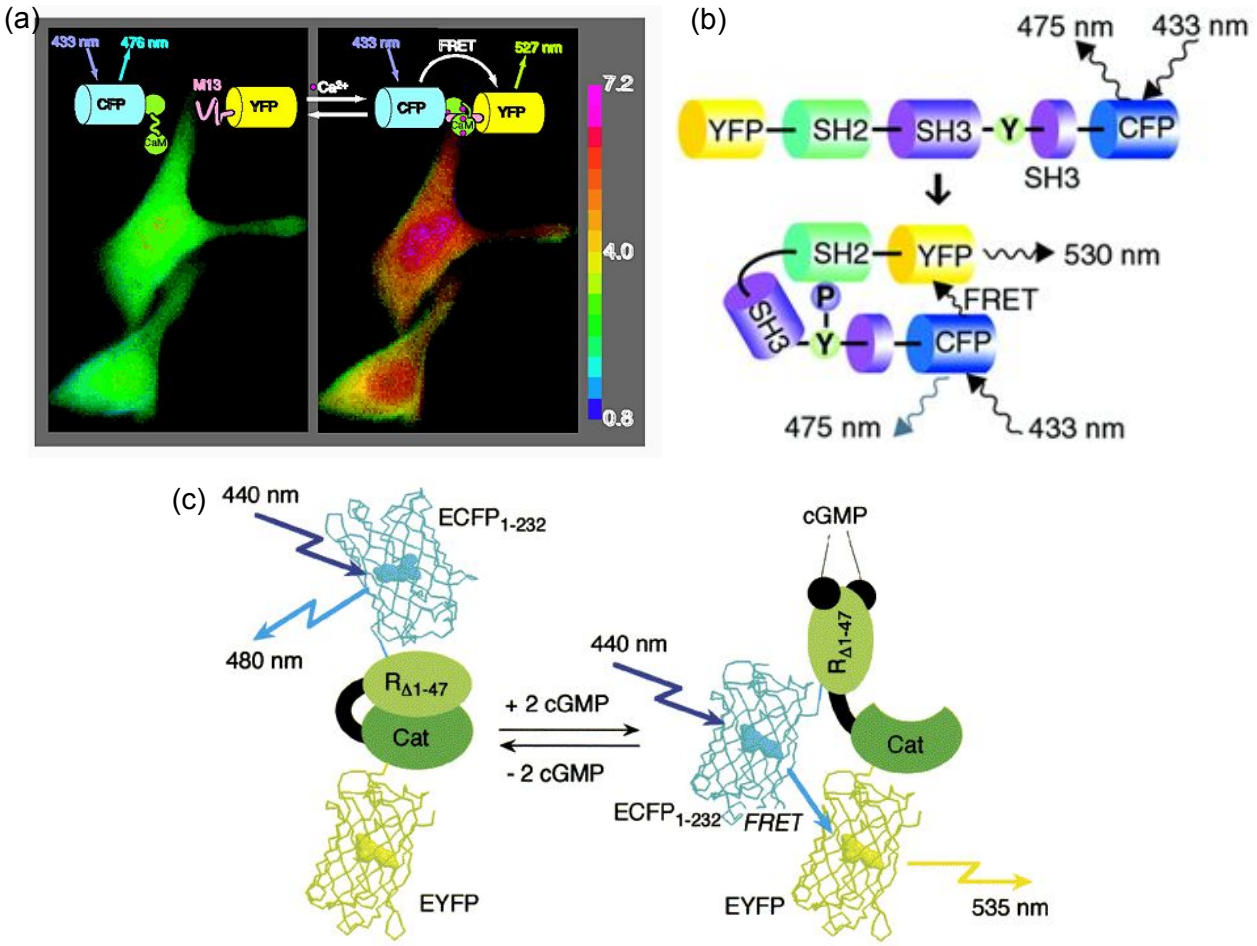


Figure 1. Schematic illustration of protein based fluorescent sensor. (a) Ca^{2+} indicator (Cameleon) [1], (b) Phosphorylation probe[3] and (c) cGMP biosensor. [2]

1-1-2. Biosensors based on ligand binding protein coupled with artificial molecular

The site-specific introduction of fluorescent dyes to a ligand binding protein is an alternative method to construct protein-based fluorescent biosensors. Compared with the fluorescent proteins, the molecular size of fluorescent dyes is obviously smaller, and it is easy to select a suitable wavelength. Furthermore, an original sensing mechanism distinct from FRET is expected.

For example, Hellinga and co-workers developed a fluorescent maltose biosensor based on maltose binding protein (MBP) as a scaffold which is one of periplasmic binding proteins.[4] The mutant MBP to which cysteine was site-specifically introduced as a chemo-selective tag was modified with an environmentally sensitive fluorophore as the transducer (Fig.2a). Applying this method to other periplasmic binding proteins, they succeeded in the development of many fluorescent biosensors for various ligands. Morii and co-workers applied the same strategy to the inositol-triphosphate (IP_3) binding domain (*PLCδ–PH* domain), and they constructed a fluorescent IP_3 biosensor (Fig. 2b).[5] The labeling site was selected based on the crystal structure of the complex of *PLCδ–PH* domain with IP_3, and the environmentally sensitive fluorophore was introduced to this position. By using the fluorescent IP_3 biosensor, they successfully monitored the intracellular dynamics of IP_3. Hirose and coworkers developed a glutamate optical sensor (EOS) based on the glutamate receptor (Fig.2c).[6] By immobilizing EOS on a cell surface of neuron cell, they succeeded in the *in situ* spatial mapping of glutamate released from synapse(Fig.2d).

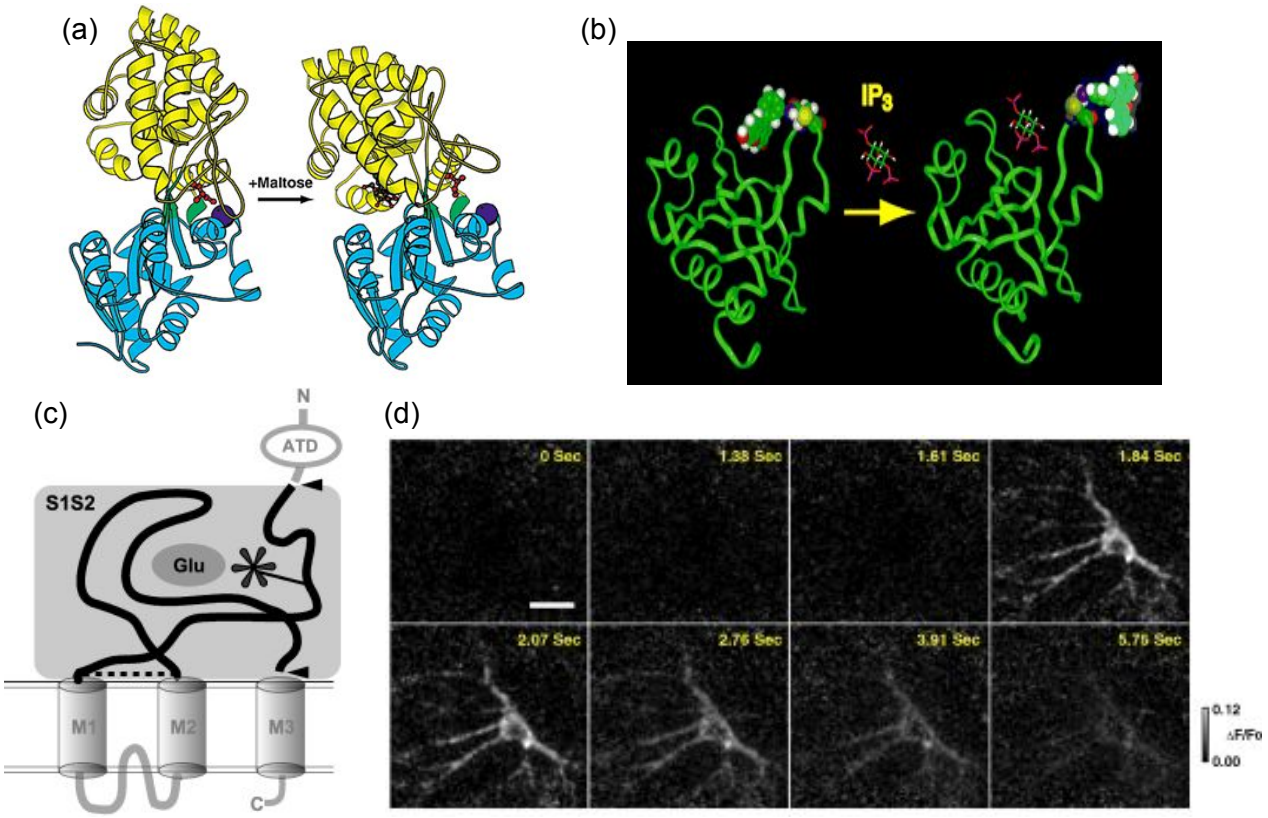


Figure 2. Schematic illustration of ligand mediated conformational changes in fluorescent biosensor for maltose (a)[4], IP3 (b)[5] and glutamate (c)[6]. (d) Visualization of the glutamate release form neuron cells using the glutamate biosensor.[6]

1-1-3. Post-Photo Affinity Labeled Modification (P-PALM) as a construction method of biosensors

As mentioned above, the site specific introduction of fluorescent dyes to proteins is a useful method to construct the fluorescent biosensors. However, it is often difficult to select a suitable site to introduce an artificial fluorophore. As a unique alternative strategy, Hamachi and Nagase developed a method to site-specifically modify the proximity of the protein active site, so-called "Post-Photoaffinity Labeling Modification (P-PALM)". In this method, we initially designed a P-PALM reagent. It has three functionalities, that is, a high affinity ligand to bind to protein, a photoreactive group to label the lectin by photoirradiation, and a cleavage site to remove the ligand and to concurrently produce a chemoselective modification site. The P-PALM method is briefly illustrated in Fig.3a.

1) When the P-PALM reagent was bound to protein, the photolabeling to the proximity of the sugar binding pocket was carried out by photoirradiation.
2) By appropriate reductants, the disulfide bond was cleaved and the reactive thiol moiety was produced.
3) Various functional molecules were attached to the thiol moiety.

They applied this method to *Concanavalin A* (Con A) which is a lectin from *Canavalia ensiformis*. The photolabeled site was characterized by the peptide-mapping using protease digestion and tandem MS/MS analysis, to be assigned to only Tyr100, a proximal amino acid in the sugar binding pocket of Con A (Fig. 3b). The fluorescence intensity of the fluorescent Con A, in which environmental sensitive fluorophores such as fluorescein and dansyl derivatives were tethered, decreased by the addition of saccharides which can bind to native Con A (Fig.3c). The association constants for saccharides that were calculated by these fluorescence changes were almost comparable to those of native Con A, indicating that the modified fluorescent Con A retained the natural sugar selectivities and affinities. In this way, they succeed in the construction of the fluorescent sugar sensor by using P-PALM for Con A.

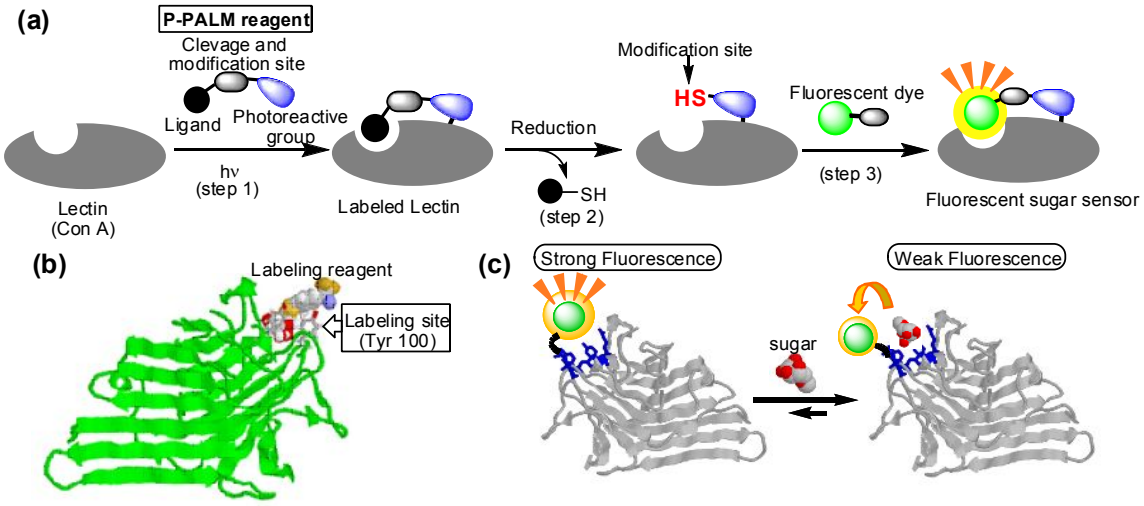


Figure 3. (a) Illustration of P-PALM scheme for semisynthesis of fluorescent biosensor.
(b) Molecularly modeled structure of the complex of Con A with P-PALM reagent.
(c) Schematic illustration of the present sensing mode of fluorescent sugar sensor.

Recently, Hamachi and Nakata successfully improved the sensing capability by incorporation of the ratiometric sensing mode into the fluorescent Con A. The ratiometric sensing by comparing two fluorescent intensities in two different wavelengths is useful to quantitative bio-imaging because it is not influenced by subtle experimental conditions such as cell thickness. The ratiometric sensing was achieved by the double modification of Con A with two distinct fluorophores in the site specific manner, using the site specific acylation to ε-amino group of Lys 114 (Fig.4a), as well as P-PALM.

The preparation of doubly modified Con A is briefly illustrated in Fig.4c. As the fluorescent dye, aminocoumarin (AMCA) and fluorescein (FL) were selected to be modified by the amino-group modification and P-PALM, respectively. A fluorescent resonance energy transfer (FRET) occurred between AMCA and Fl, so that the single wavelength excitation / the double wavelength detection was carried out by this doubly modified Con A. By the benefit of the ratiometric sensing, the detection of the intracellular glucose was carried out (Fig.4b).

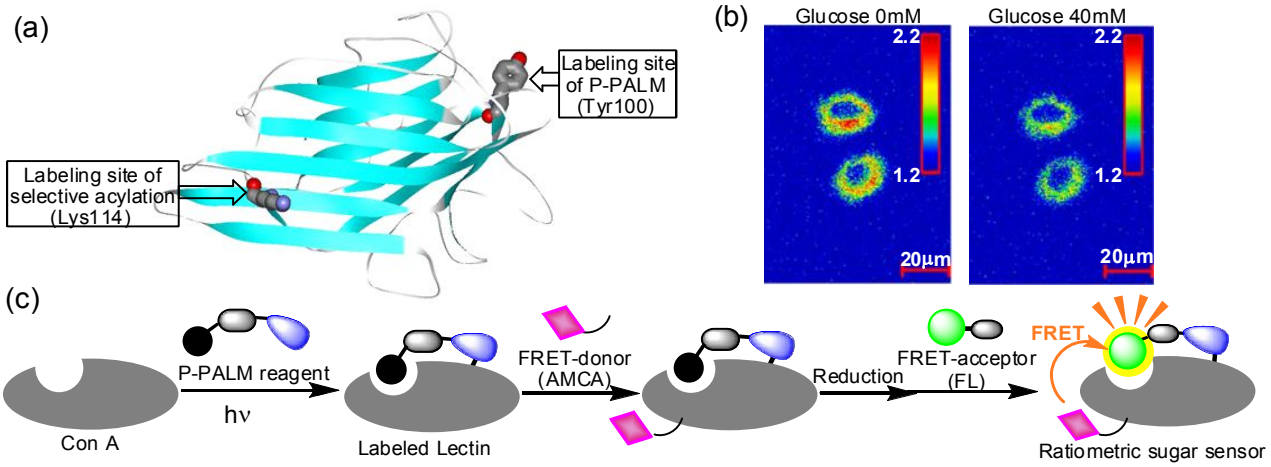


Figure 4 (a) The crystal structure of Con A including the amino acids of labeling site (Tyr100, Lys114).
(b) Intracellular ratiometric image of HepG2 cells probed by AMCA-Fl-Con A.
(c) Scheme for double modification of Con A by coupling P-PALM and the selective acylation to afford semisynthetic ratiometric sugar sensor.

1.1.3 Biosensors by using long-lived luminescence

In biosensing, in spite of many advantages, fluorescent biosensors have some drawbacks. In particular, an intrinsic fluorescence is a big drawback of the intracellular fluorescent sensing. On the other hands, it is considered that the lanthanide-based luminescence has advantages over the fluorescence on several points such as the unique longer lifetime of emission, being scarcely disturbed with scattering of the excited light, and with fluorescent impurities such as an intrinsic fluorescence of biological samples (Fig.5a).[9] Thanks to these benefits, a luminescence mode sensing system is anticipated to be promising particularly for application to complicated mixtures of biological samples.

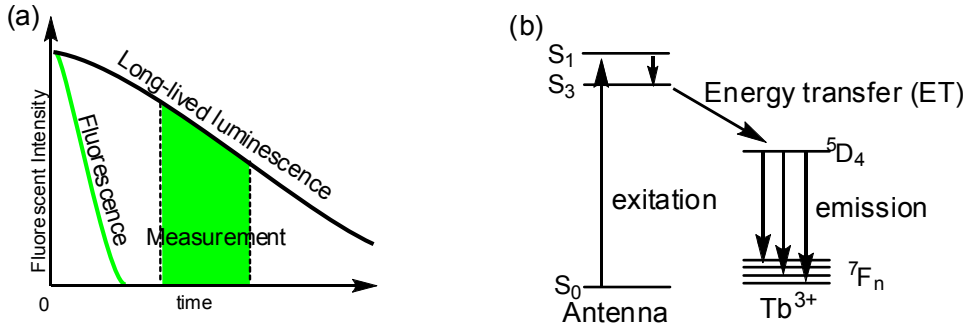


Figure 5 (a) The advantage of long-lived luminescence. (b) Energy level of Tb^{3+}.

In the lanthanide ions, it was known that terbium ion (Tb^{3+}) and europium ion (Eu^{3+}) show the strong luminescence. However, the direct excitation for these ions is difficult, so that Energy Transfer (ET) from a proximal antenna molecule was used (Fig.5b). Kikuchi and coworkers developed a new Tb^{3+}-based luminescent indicator for zinc ion (Zn^{2+}) by using Zn^{2+}-modulated ET (Fig.6a).[10] This indicator enabled the time-resolved observation of the intracellular dynamics of Zn^{2+}. Matsumoto and coworkers detected a specific DNA sequence by Luminescent Resonance Energy Transfer (LRET).[11] The hybridization of a target DNA and the complimentary oligo-DNAs labeled with lanthanide complex and fluorophore was efficiently monitored by LRET. By the benefit of the long-lived luminescence, the highly sensitive detection, compared to the conventional methods by FRET, was achieved.

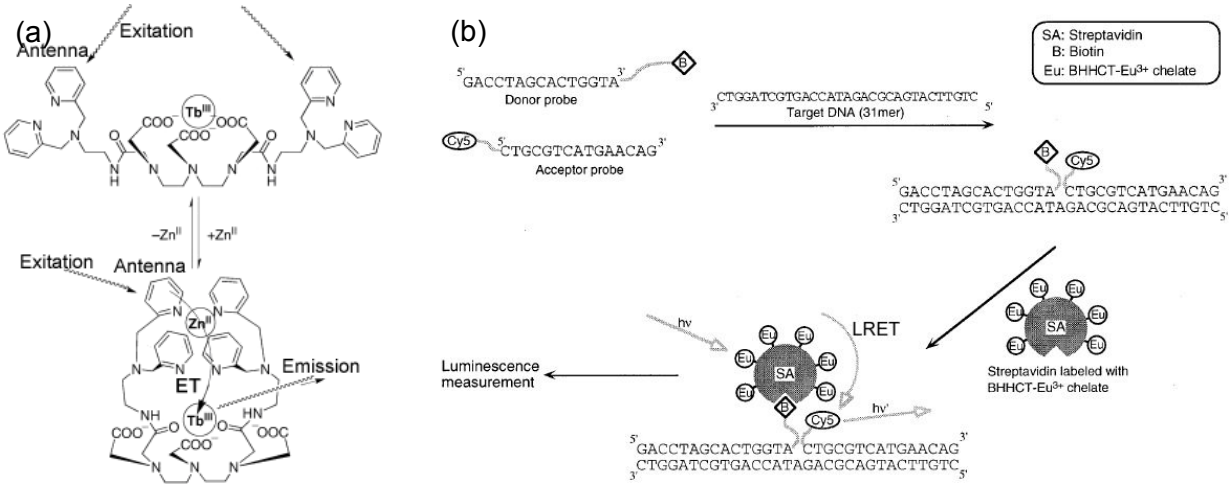


Figure 6 (a) Molecular structure of Zn^{2+} sensor.[10] (b) DNA hybridization assay using LRET.[11]

These examples showed that the application of lanthanide-based luminescence is a useful method to improve the capability of biosensors. However, there are no reports on natural protein having long-lived luminescence. Thus, in this chapter, I described to develop a luminescent biosensor by the incorporation of long-lived luminescent lanthanide complex into a protein scaffold.

1-2. Result

1-2-1. Construction of luminescent sugar biosensor having a long lifetime

Site-specific introduction of Fluorescein (FL) to *Concanavalin A* (Con A), a mannose-binding protein, was conducted by post-photoaffinity labeling modification (P-PALM). That is, the benzylthiol group newly incorporated into the Tyr100 site of the Con A surface was labeled with Fluorescein-5-maleimide (Fig.7a) to yield Fluorescein-modified Con A (FL-Con A, Fig.7b). It is reported that there is the binding site of metal ion such as manganese ion (Mn^{2+}) or terbium ion (Tb^{3+}) in the proximity of the sugar binding pocket of Con A (Fig.8a).[12] We used this binding site in order to introduce Tb^{3+} into Con A. In fact, when terbium chloride ($TbCl_3$) was added to native Con A, mainly three luminescence peaks (490, 544 and 583 nm) appeared under the time-resolved luminescence mode (delay time = 0.1 ms, gate time = 0.5 ms, Fig.8b). But the luminescence of Tb^{3+} was not observed in the absence of Con A. These luminescence due to the Tb^{3+} emission sensitized by fluorescent amino acids (Trp, Tyr etc) of Con A are a direct evidence of the introduction of Tb^{3+} into Con A (Con A/Tb) because the close distance (within 3 Å) is required for the sensitization.[13]

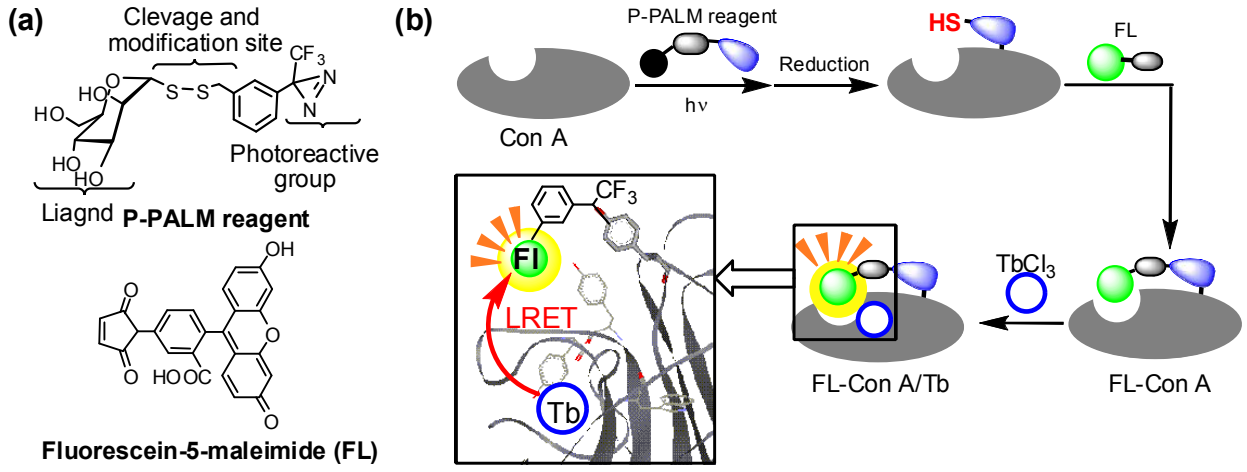


Figure 7 (a) Molecular structures of P-PALM reagent and FL. (b) P-PALM scheme for semisynthesis of the long-lived luminescent sugar biosensor and the mechanism of LRET on Con A.

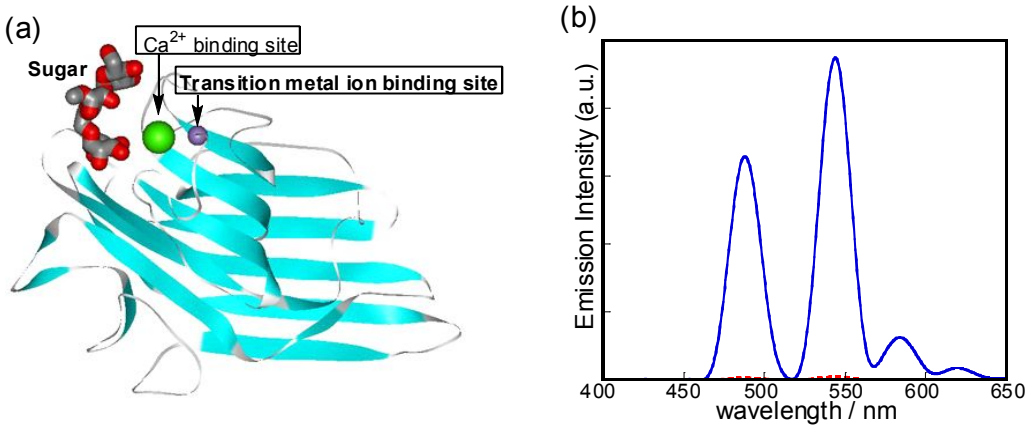


Figure 8 (a) Crystal structure of Con A including metal ions. (b) Time resolved luminescence spectra (delay 0.1 ms, gate 0.5 ms) of 50 μM $TbCl_3$ in the presence (blue flat line) and in the absence (red dotted line) of 3 μM Con A. Conditions: 10 mM HEPES buffer (pH 7.5), 5 mM $CaCl_2$, 0.1 M NaCl, λ_{ex} = 280nm.

Next, we introduced $TbCl_3$ into FL-Con A, and we observed the luminescent change under the time-resolved luminescence mode. As shown in Fig. 9a, an additional peak at 513 nm was newly observed by the addition of $TbCl_3$ into FL-Con A. Figure 9b shows that the FL-Con A excitation spectrum obtained by the fluorescence mode was sufficiently overlapped with one of the Tb^{3+} emission peak (490nm) and the emission fluorescent peak is coincident with the newly appeared peak (513nm). In addition, the excitation maximum of the luminescence of FL-Con A monitored at the FL emission peak (513 nm) appeared at 280 nm in the presence of Tb^{3+}, the value of which is identical with the excitation spectrum monitored at the Tb^{3+} emission peak (544 nm, Fig. 9c). Thus, it is clear that the Tb^{3+}-complex was formed with FL-Con A (FL-Con A/Tb) and the Fl emission appeared in the rather long time range by a LRET from Tb^{3+} (Fig.7b).

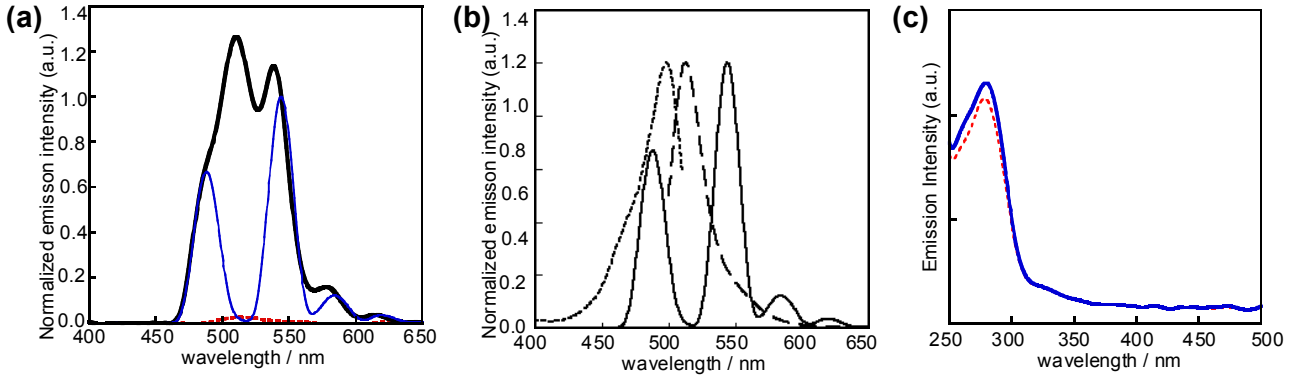


Figure 9. (a) Time resolved luminescence spectra of 3μM native-Con A with 50μM $TbCl_3$ (blue light line) and 1μM FL-Con A in the absence (red dashed line) and presence (black heavy line) of 50μM $TbCl_3$ (λ_{ex} = 280 nm). (b) Excitation (dot line) and emission (dash line) spectra of 1μM FL-Con A in the fluorescence mode, and the emission spectrum of 3μM native-Con A with 50μM $TbCl_3$ (solid line) in the phosphorescence mode. (c) Excitation spectrum of FL-Con A/Tb monitored at the FL emission peak (blue solid line) and monitored at the Tb^{3+} emission peak (red dot line) by the phosphorescence mode. Conditions: 10 mM HEPES buffer (pH 7.5), 5 mM $CaCl_2$, 0.1 M NaCl.

1-2-2. Sugars sensing by long-lived luminescence

Owing to the LRET emission of FL, two distinct luminescent fluorophores (i.e., Tb^{3+} and FL) that are essential to ratiometric luminescence analysis are apparently equipped within one protein scaffold (FL-Con A/Tb). We conducted luminescent titration experiments of FL-Con A/Tb with several saccharide derivatives such as monosaccharides and oligosaccharides. Figure 10a shows a typical luminescent spectral change of FL-Con A/Tb by addition of a branched mannopentaose (Man-5), an essential fragment of glycoproteins surface. Clearly, a LRET peak at 513 nm decreases relative to the Tb luminescence at 544 nm with increasing Man-5 concentration. The titration curve for Man-5 obeys a typical saturation as shown in Fig. 10b and gives us a binding constant of 1.0×10^5 M^{-1}. The value is comparable to that of native Con A. Figure 10b also displays other titration curves with different saccharide derivatives monitored by the ratiometric luminescence. The ratio change occurs for all of mannose derivatives with corresponding sensitivity, whereas

Me-α-galactose and a β-linked glucose derivative such as cellobiose are not sensed luminescently. The selectivity and binding affinity thus evaluated (Table 1: Man-5, -4, -3 > Man-2, Man-1 > Cel, >> Me-α-Gal) is almost identical with literature values of native Con A determined by ITC (isothermal titration calorimetry), indicating that FL-Con A/Tb is a luminescent biosensor retaining a natural selectivity in spite of the double modification.

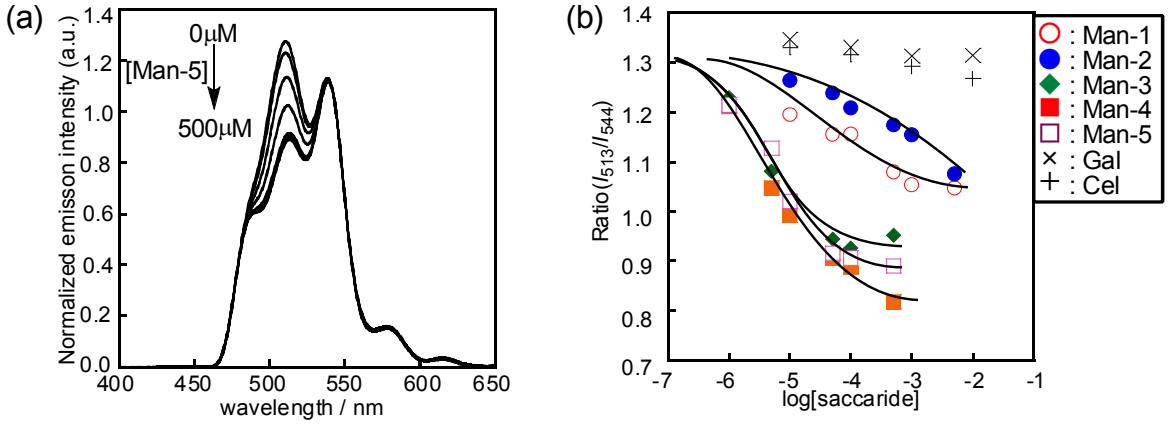


Figure 10 (a) Time resolved luminescent spectral changes of FL-Con A/Tb (1μM Fl-Con A, 50μM $TbCl_3$) upon addition of Man-5 (0→500μM). (b) Luminescent titration plots of the ratiometric intensity (I_{513}/I_{544}) versus saccharide concentration (log [saccharide]): Man-1 (○), Man-2 (●), Man-3 (◆), Man-4 (▲), Man-5 (□), Gal (×), Cel (+), in 10mM HEPES buffer (pH 7.5), 5mM $CaCl_2$, 0.1M NaCl, λ_{ex} = 280 nm.

Table 1. Comparison of the association constants of FL-Con A/Tb with those of native-Con A to various saccharide and glycoproteins.

Saccharide	K / M^{-1}	
	FL-ConA / Tb	native-ConA[b]
Man-1	1.6×10^4	1.1×10^4
Man-2	4.3×10^3	1.3×10^4
Man-3	2.4×10^5	2.5×10^5
Man-4	2.2×10^5	2.0×10^5
Man-5	1.0×10^5	2.0×10^5
Gal	- [a]	- [a]
Cel	- [a]	- [a]
Ribo B (Man-5 derivative)	1.8×10^5	- [c]
Ribo A (Man-5 nonderivative)	- [a]	- [c]

[a] Precise values cannot be determined because of the low affinity. [b] Determined by ITC [14]. [c] No data were reported in previous literatures.

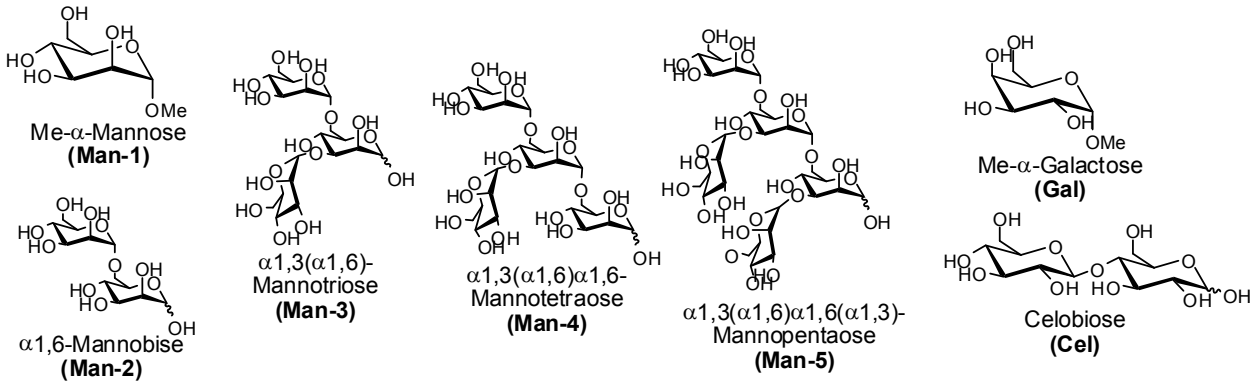


Figure 11. Saccharide structure used in this titration study.

Because a long-lived luminescence is advantageous over the short-lived fluorescence, we can detect sugar in the presence of the fluorescent impurities by FL-Con A / Tb. In fact, the fluorescent impurities such as another contaminated fluorophore (DMACA, Fig12a), a protein emission and the scattered light due to the excitation, all of which were observed in the fluorescence spectra (Fig.12b) can be cancelled in its luminescence spectra so as to obtain the simplified spectrum (Fig.12c). In addition, the ratiometric luminescent changes were observed by the addition of Man-5.

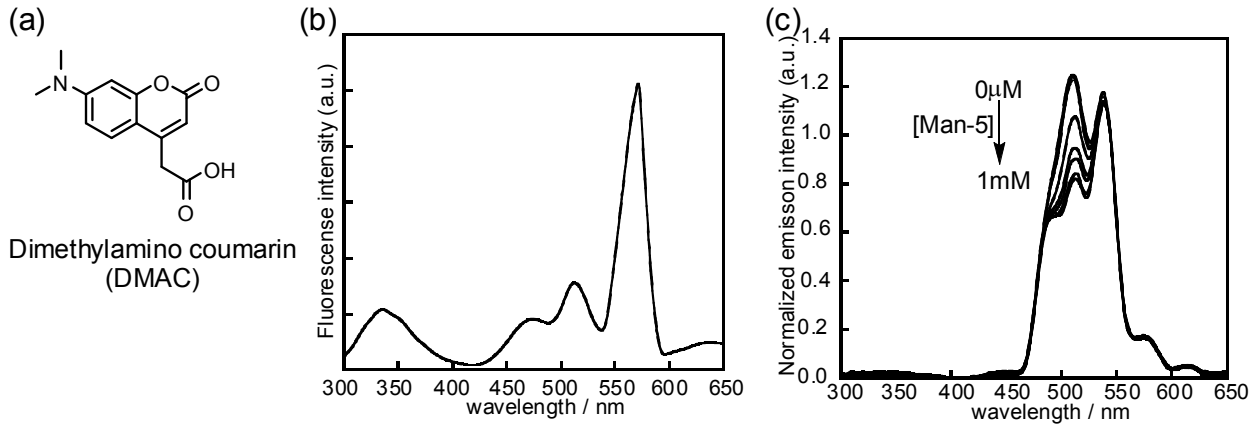


Figure 12 (a) Molecular structures of DMAC. (b) Fluorescence spectra of Fl-Con A/Tb in the presence of 20 μM DMAC. (c) Time-resolved luminescent spectral changes of FL-Con A/Tb upon addition of Man-5 (0→1 mM) in the presence of 20 μM DMAC. Conditions: 10 mM HEPES buffer (pH 7.5), 5 mM $CaCl_2$, 0.1 M NaCl, λ_{ex} = 280 nm.

<u>1-2-3. Enzyme Assay by luminescent biosensor</u>

In addition to sugars, a glycoprotein sensing can be also carried out by LRET of FL-Con A/Tb with a microM sensitivity ($K = 1.8 \times 10^5 M^{-1}$ for ribonuclease B (Ribo-B)[15], an enzyme modified with a Man-5 derivative), while mannose and the nonglycosylated protein (Ribo-A) were less sensitively detected (Fig.13). Using this selectivity, a luminescent assay for an enzymatic trimming reaction on a glycoprotein surface can be designed. In a proof-of-principle experiment, α-mannosidase, an exoglycosidase [16], and Ribo-B were employed as a model enzyme and a substrate, respectively. Figure 14a shows a typical luminescence change of FL-Con A/Tb depending on the reaction time

and the time course monitored by the ratiometric intensity was plotted in Fig.14b. Clearly, the LRET peak increases upon the α-mannosidase injection and the change gradually levels off. The observed change is reasonably explained as follows: the LRET peak intensity was suppressed due to the Ribo-B complexation with FL-Con A/Tb at the initial stage, and it was recovered by the concentration decrease of the complexed species due to the hydrolytic cleavage of Man-5 moiety of Ribo-B (Fig.14c) because the produced mannose has the much lower affinity to FL-Con A/Tb. It should be noted that this simple method needs neither a tedious labeling process of saccharides nor fully equipped mass spectrometers.

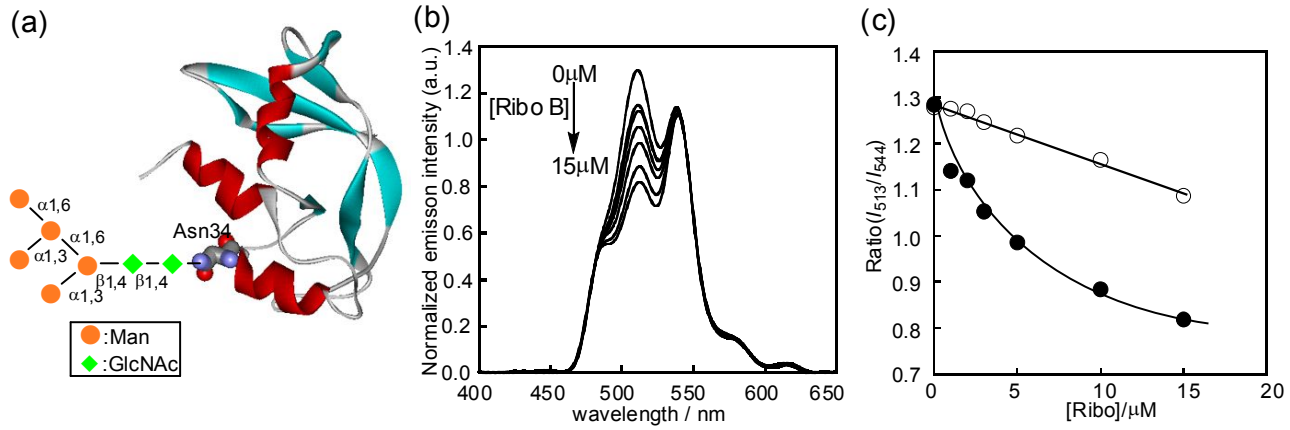


Figure 13. (a) Sugar chain structure of Ribo B. (b) Time-resolved luminescent spectral changes FL-Con A/Tb upon addition of Ribo B (0→15μM). (c) Luminescent titration plots of the ratiometric intensity (I_{513}/I_{544}) versus saccharide concentration (log [protein]): Ribo-B (●), Ribo A (❍). Conditions: 10 mM HEPES buffer (pH 7.5), 5 mM $CaCl_2$, 0.1 M NaCl, λ_{ex} = 280 nm.

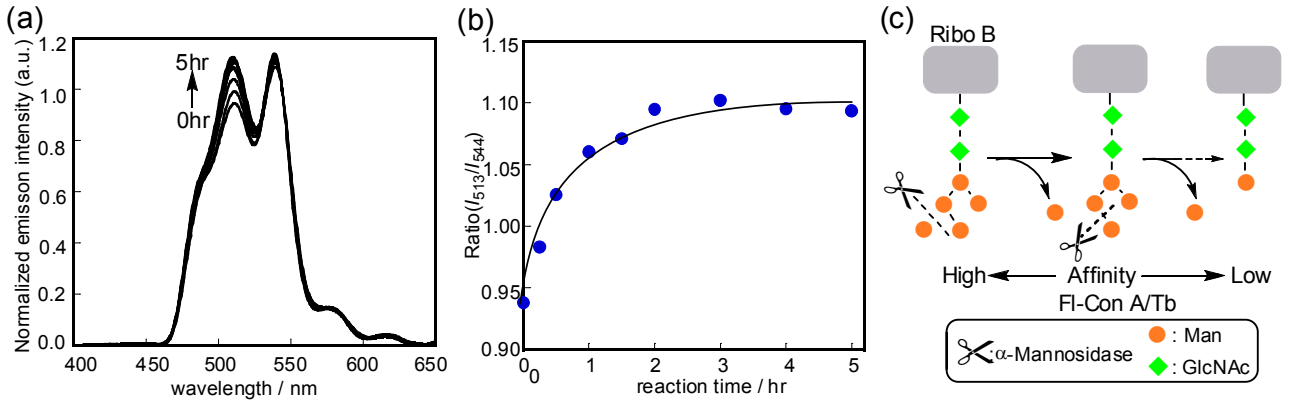


Figure 14. (a) Time dependent spectral change of the emission of FL-Con A/Tb in the presence of Ribo B by addition of α-mannosidase (0→5hr). (b) Time profile of the ratiometric intensity (I_{513}/I_{544}). (c) Illustration of luminescent sensing scheme of the branched mannoside on Ribo B by α-mannosidase.

1-3. Discussion

1-3-1. Sugar sensing mechanism

The long lived-luminescent sugar biosensor (FL-Con A/Tb) was constructed by coupling of the fluorescent sugar sensor (FL-Con A) with the Tb^{3+} luminescence which worked as internal standard. The fluorescence intensity of FL-Con A decreased by the addition of Man-5 (Fig.15a), whereas, response was observed neither for a randomly-modified FITC-Con A/Tb (Fig.15b) nor for a native Con A/Tb (Fig.15c). Therefore, it is assumed that the FL attached to the proximity of the sugar-binding pocket can sense a micro-environmental change such as micro-pH caused upon the saccharide binding. The bulkier saccharide kicked FL out from the sugar binding cavity of Con A, so that the fluorescence change is induced. In fact, Nakata previously confirmed that the absorption spectral change of FL-Con A upon the addition of Man-3 was similar to the absorption change induced by the pH shift (Fig.16).

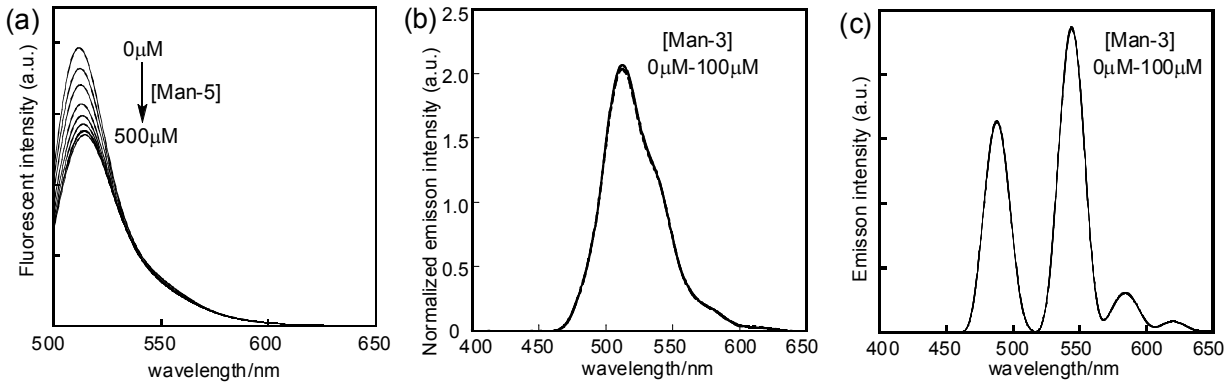


Figure 15 (a) Fluorescent spectral changes of FL-Con A upon addition of Man-5 (0→500 μM). (b) Time resolved luminescent spectral changes in the emission of FITC-Con A / Tb ((b), FITC-Con A 1□M and 50μM $TbCl_3$) and native Con A / Tb ((c), native Con A 1μM and 50μM $TbCl_3$) upon addition of Man-3 (0→100μM).

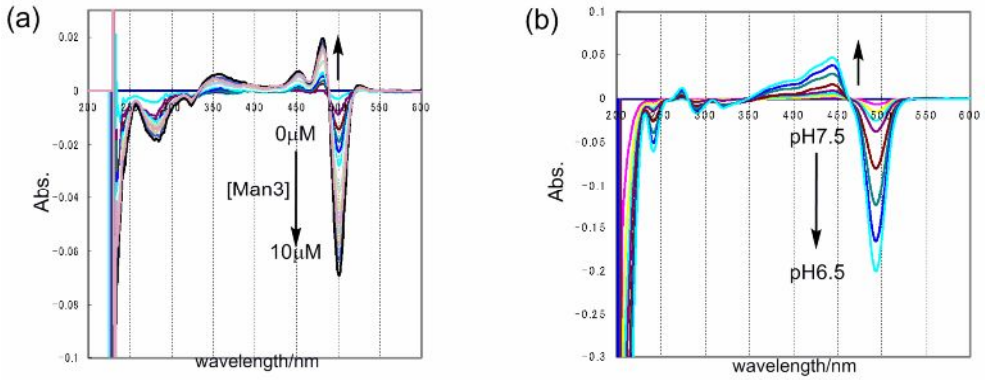


Figure 16. Absorbance difference spectrum of FL-Con A by the addition of Man-3 (0→10 μM) (a) or by the pH shift (pH 7.5→pH6.5) (b).

1-3-2. Method for introduction of lanthanide ion to proteins

In this method, we introduced Tb^{3+} to the metal ion binding site of Con A. It was reported that not only Con A but also other legume lectins such as *Lens Culinaris Agglutinin* (LCA) and *Peanut*

Agglutinin (PNA) have such metal ion binding sites. Thus, it is expected to construct the long-lived luminescent sugar biosensors which can detect different types of saccharides by applying the same method to other legume lectins.

On the other hand, the low binding affinity of Tb^{3+} (K_d=60μM) to lectins is a drawback of this method. This may be solved by the chemical labeling of an appropriate chelator for lanthanide ion. Various chelators were already reported (Fig.17a).[17] For instance, Imperiali and coworkers developed an efficient lanthanide binding tag (LBT) based on a Ca^{2+} biding domain (Fig.17b).[18] By a fusion of LBT to the protein of interest, lanthanide ion can be easily fixed to the target protein.

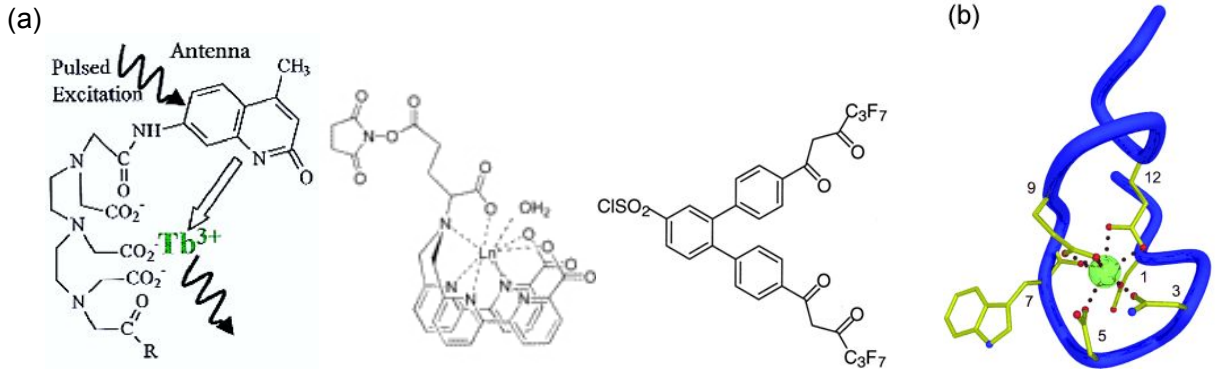


Figure17. (a) Molecular structures of chelators for protein labeling.[17] (b) Structure of the LBT-Tb^{3+} complex.[18]

1-3-3. The utility of LRET to various bioanalysis

Sensing with long-lived luminescence was achieved by Luminescent Resonance Energy Transfer (LRET) on the Con A in this study. LRET is now considered to be an excellent method to improve the capability of protein-based biosensors, and therefore, some examples were recently reported. For example, Imperiali and coworkers developed interesting probes for analyzing protein-peptide interaction. This probe was based on phosphorylated peptide binding domain (SH2 domain), and LRET was used to monitor the interaction between LBT-fused SH2 domain and a fluorescent dye-modified peptide (Fig.18a).[19] Selvin and coworkers applied LRET to analyze the dynamics of a voltage gated potassium channel (Fig.18b).[20] They succeeded in monitoring a voltage-induced conformational change by LRET between Tb^{3+}/chelator labeled channel and fluorophore-tethered ligand (scorpion toxin).

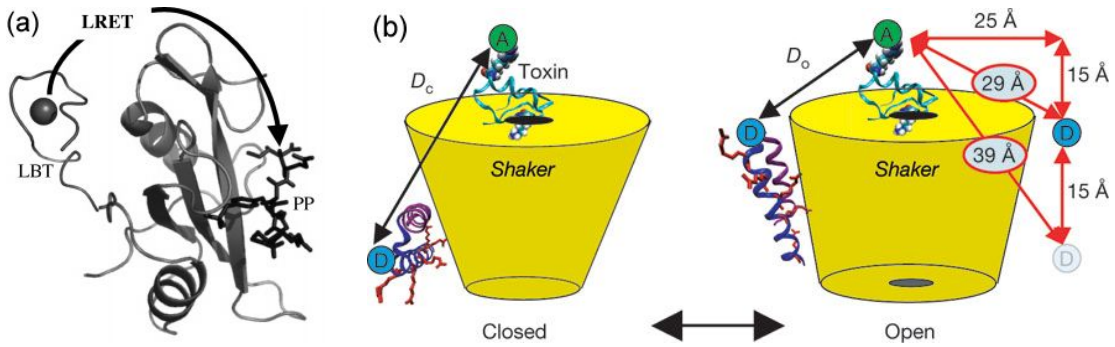


Figure 18. (a) Graphic of SH2-LBT with a bound phosphopeptide.[19] (b) Diagram of lanthanide complex modified K^+ channel and fluorophore labeled toxin.[20]

1-4. Conclusion

In this chapter, I successfully constructed a luminescent sugar biosensor with the longer life-time by incorporating the lanthanide-based luminescence into a protein-based biosensor. Thank to the benefit of long-lived Tb^{3+} luminescence, not only ratiometric sensing to complicated saccharide derivatives but also the luminescent monitoring the glycoprotein trimming process were carried out. The present result is expected to extend the validity of luminescent sensing in biological systems, and this biosensor may be applicable to other enzymatic processes of glyco-conjugates.

1-5. Experimental Section

General Methods

MALDI-TOF MS were recorded on PE Voyger DE-RP, where SA was used as a matrix. UV-vis spectra were obtained on a Shimazu UV-3500 spectrometer. Luminescence spectra were recorded on Perkin Elmer LS 55 spectrometer. Con A and randomly-modified FITC-Con A were purchased from Funakoshi and used without further purification. Fluorescein-5-maleimide and 7-dimetylaminocomarin-4-acetic acid were purchased from Molecular Probes. α-Mannosidase (from jack bean) was purchased from Seikagaku Corporation. Ribonuclease B was purchased from Sigma and used after purification by HPLC. The other chemical reagents were purchased from Aldrich, Tokyo Chemical Industry, or Seikagaku Corporation.

Preparation of FL-Con A

SH-Con A was prepared according to our previous report.[7] Fluorescein-maleimide (56.0μM, Molecular Probe) in 10mM phosphate buffer (pH7.5) was slowly added to a solution of SH-Con A (11.2μM, 11.5ml) in 10mM phosphate buffer (pH7.5) at 4 ˚C, and the reaction mixture was incubated for 12h at r.t. in the dark. After incubation, the solution was submitted to gel filtration chromatography (TOYO PEARL, 1 x 15 cm), the eluted protein fraction (10mM acetate buffer (pH 5.0)) was collected, concentrated by ultrafiltration, and dialyzed against 10mM HEPES buffer solution (pH 7.5) containing 5mM $CaCl_2$, and 0.1M NaCl, to afford Fl-Con A (3.2μM, 30.0ml). The labeling efficiency was estimated from the ratio of [FL]/([Con A]/2) to be 0.66-0.80. Fl-Con A was identified by MALDI-TOF-MS (SA): *m/z* calcd: 26,232; found: 26,226±9. UV spectra: ε_{492}=80,000$M^{-1}cm^{-1}$, ε_{280}=23,500$M^{-1}cm^{-1}$ (pH 7.5). Con A: ε_{280}=35,000$M^{-1}cm^{-1}$, excitation and emission spectra : λ_{ex} = 488 nm, λ_{em} = 513 nm.

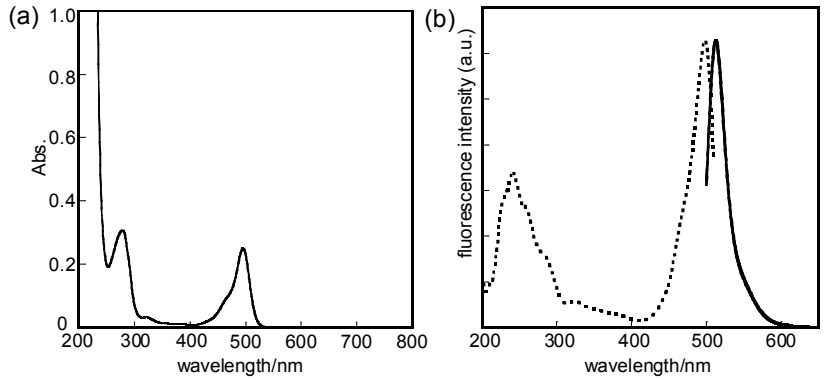


Figure 19. a)UV-visible spectroscopy of the FL-Con A. b) Excitation (dot line) and emission (solid line) spectra of FL-Con A. λ_{ex} = 488 nm, λ_{em} = 513 nm.

Time-resolved luminescence study

All stock solutions (FL-Con A, $TbCl_3$, saccharide) were bubbling under argon atmosphere prior to usage. Luminescence spectrum (Perkin Elmer LS 55) was measured with 0.5 msec gate after a 0.1 msec delay. The slit widths for the excitation and emission were set to 10nm and 20nm, respectively. The excitation wavelength was 280nm.

Luminescence titration of saccharides

Saccharide solution was added dropwise to a 1μM FL-Con A in the presence of 50μM $TbCl_3$ in 10mM HEPES buffer (pH 7.5) containing 5mM $CaCl_2$, 0.1M NaCl at 20 ˚C, and the luminescence spectrum was measured. Titration curves thus obtained were analyzed with the nonlinear least-squares curve-fitting method or the Benesi-Hildebrandt plot to give the association constants for various saccharides.

Fluorescence titration of saccharide

Saccharide solution was added dropwise to a 1μM Fl-Con A in 10mM HEPES buffer (pH 7.5) containing 5mM $CaCl_2$, 0.1M NaCl at 20 ˚C, and the fluorescence spectrum was measured. λ_{ex} = 488 nm.

Enzyme Assay

After 0.05 μmol of ribonuclease B (Ribo B) in 0.2 mL of pH 6.5 100 mM Tris / bis-Tris buffer (0.25 mM) was mixed with 1 unit of α-mannosidase in distilled water, the reaction mixture was incubated at 37 °C. At a corresponding reaction time, the reaction solution was collected and mixed with to a 1μM FL-Con A in the presence of 50μM $TbCl_3$ (the final concentration of Ribo B 10 μM, 10 mM HEPES buffer (pH 7.5) containing 5 mM $CaCl_2$ and 0.1 M NaCl), and the time-resolved luminescent spectrum was measured. The trimming process of Ribo B was also monitored by MALDI-TOF MS.

1-6. Reference

1) Miyawaki, A.; Llopis, J.; Heim, R.; McCaffery, J. M.; Adams, J. A.; Ikura, M.; Tsien R. Y. *Nature* **1997**, *388*, 882-887.

2) (a) Sato, M.; Hida, N.; Ozawa, T.; Umezawa, Y. *Anal. Chem.* **2000**, *72*, 5918-5924. (b) Honda, A.; Adams, S. R.; Sawer, C. L.; Lev-Ram, V.; Tsien, R. Y.; Dostman, W. R. G. *Proc. Natl. Acad. Sci. USA* **2000**, *98*, 2437-2442.

3) (a) Kurosawa, K.; Mochizuki, N.; Ohba, Y.; Mizuno, H.; Miyawaki, A.; Matsuda, M. *J. Biol. Chem.* **2001**, *276*, 31305-31310. (b) Ting, A. Y.; Kain, K. H.; Klemke, R. L.; Tsien, R. Y. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 15003-15008.

4) (a) Delorimier, R. M.; Smith, J. J.; Dwyer, M. A.; Looger, L. L.; Sal, K.M.; Paavola, C. D.; Rizk, S. S.; Sadigov, S.; Conrad, D. W.; Loew, L.; Hellinga, H. W. *Protein Science* **2002**, *11*, 2655-2675. (b) Marvin, J. S.; Hellinga, H. W. *Nature Struc. Biol.* **2001**, *8*, 795-798.

5) (a) Morii, T.; Sugimoto, K.; Makino, K.; Otuka, M.; Imoto, K.; Mori, Y. *J. Ame. Chem. Soc.* **2002**, *124,* 1138-1139. (b) Sugimoto, K.; Nishida, M.; Otsuka, M.; Makino, K.; Ohkubo, K.; Mori, Y.; Morii, T. *Chem. Biol.* **2004**, *11*, 475-485.

6) Namiki, S.; Sakamoto, H.; Iinuma, S.; Iino, M.; Hirose, K. *Eur. J. Neurosci.* **2007**, *25*, 2249-2259.

7) (a) Nagase, T.; Nakata, E.; Shinkai, S.; Hamachi, I. *Chem. Eur.-J.* **2003**, *9*, 3660-3669. (b) Hamachi, I.; Nagase, T.; Shinkai, S. *J. Am. Chem. Soc.* **2000**, *122*, 12065-12066. (c) Koshi, Y.; Nakata, E.; Hamachi, I. *Trends in Glycoscience and Glycotechnology* **2007**, *19*, 121-131.

8) Nakata, E.; Koshi, Y.; Koga, E.; Katayama, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2005**, *127,* 13253-13261.

9) Richardson, F. S. *Chem. Rev.* **1982**, *82*, 541.

10) Hanaoka, K.; Kikuchi, K.; Kojima, H.; Urano, Y.; Nagano. T. *J. Am. Chem. Soc.* **2004**, *126,* 12470-12476.

11) (a) Sueda, S.; Yuan, J.; Matsumoto, K. *Bioconjugate Chem.* **2000**, *11,* 827-831. (b) Sueda, S.; Yuan, J.; Matsumoto, K. *Bioconjugate Chem.* **2002**, *13,* 200-205.

12) (a) Barber, B. H.; Fuhr, B.; Carver, J. P. *Biochemistry* **1975**, *14*, 4075. (b) Sherry, A. D.; Cottam, G. L. *Arch. Biochem. Biophys.* **1973**, *156*, 665.

13) a) W. D. Horrocks, Jr., W. E. Cillier, *J. Am. Chem. Soc.* **1981**, *103*, 2856-2862; b) J. Bruno, W. D. Horrocks, Jr., R. Zauhar, *J. Biochemistry* **1992**, *31*, 7016-7026.

14) Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387.

15) Williams, R. L.; Greene, S. M.; McPherson, A. *J. Biol. Chem.* **1987**, *262*, 16020.

16) a) Y. T. Li, *J. Biol. Chem.* **1966**, *241*, 1010-1012; b) A. L. Tarentino, T. H. Plummer, Jr., F. J. Maley, *J. Biol. Chem.* **1970**, *245*, 4150-4157.

17) Weibel, N.; Charbonniere, L. J.; Guardigli, M.; Roda, A.; Ziessel, R. *J. Am. Chem. Soc.* **2004**, *126*, 4888.

18) (a) Nitz, M.; Sherawat, M.; Franz, K. J.; Peisach, E.; Allen, K. N.; Imperiali, B. *Angew. Chem. Int. Ed.* **2004**, *43*, 3682-3685. (b) Franz, K. J.; Nitz, M.; Imperiali, B. *ChemBioChem* **2003**, *4*, 265.

19) Sculimbrene, B. R.; Imperilai, B. *J. Am. Chem. Soc.* **2006**, *128*, 7436-7352.

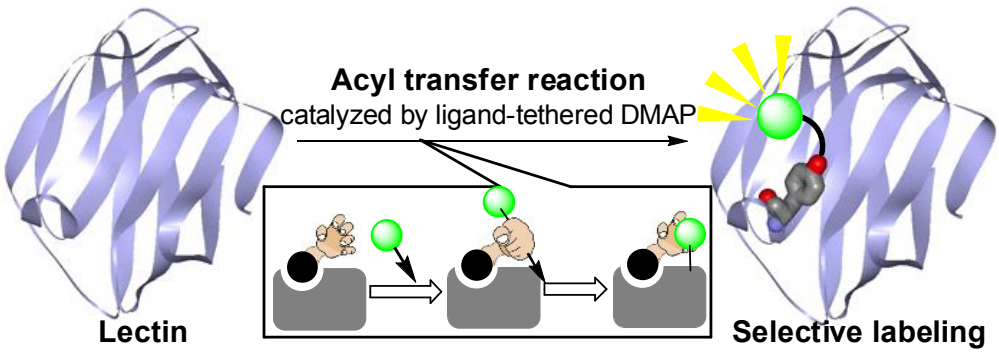
20) Posson, D. J.; Ge, P.; Miller, C.; Bezanlla, F.; Selvin, P. R. *Nature* **2005**, *436*, 848-851.

Chapter 2

Target-specific chemical acylation of lectins by ligand-tethered DMAP catalyst.

Abstract

Because sugar-binding proteins, so called lectins, play important roles in many biological phenomena, the lectin selective labeling should be useful for investigating biological processes involving lectins as well as providing molecular tools for analysis of saccharides and these derivatives. We describe herein a new strategy for lectin selective labeling based on an acyl transfer reaction directed by ligand-tethered DMAP (4-dimethylaminopyridine). DMAP is an effective acyl transfer catalyst which can activate an acyl ester for its transfer to a nucleophilic residue. To direct the acyl transfer reaction to a lectin of interest, we attached the DMAP to a saccharide ligand specific for the target lectin. It was clearly demonstrated by biochemical analyses that the target-selective labeling of Congerin II, an animal lectin having selective affinity for Lactose/LacNAc (*N*-acetyllactosamine), was achieved in the presence of Lac-tethered DMAPs and acyl donors containing probes such as fluorescent molecules or biotin. Conventional peptide mapping experiments using HPLC and tandem mass-mass analysis revealed that the acyl transfer reaction site-specifically occurred at Tyr 51 of Cong II. This strategy was successfully extended to other lectins by changing the ligand part of the ligand-tethered DMAP. We also demonstrated that this labeling method is applicable not only to purified lectin in test tubes, but also to crude mixtures such as E. coli lysates or homogenized animal tissue samples expressing Congerin.



2-1. Introduction

2-1-1. Selective chemical labeling of proteins

A method for selective chemical labeling based on a specific protein-ligand interaction, so called "Affinity Labeling" or "Photo-Affinity Labeling", is one of the important technologies in protein science (Fig.1). In the affinity labeling, a labeling reagent consisting of three functionalities, those are a ligand moiety to selectively bind to a protein, a reactive group to label the target protein and a tag to select the labeled one, is usually used. When a labeling reagent is bound to a target protein, the selective labeling to the proximity of the ligand binding site was carried out mainly owing to the proximity effect. In the photo-affinity labeling, a labeling reagent having a photoreactive group is used, and the labeling reaction was triggered by photo irradiation. These methods are useful for the exploration of new receptor proteins and the determination of ligand binding site of the protein.[1]

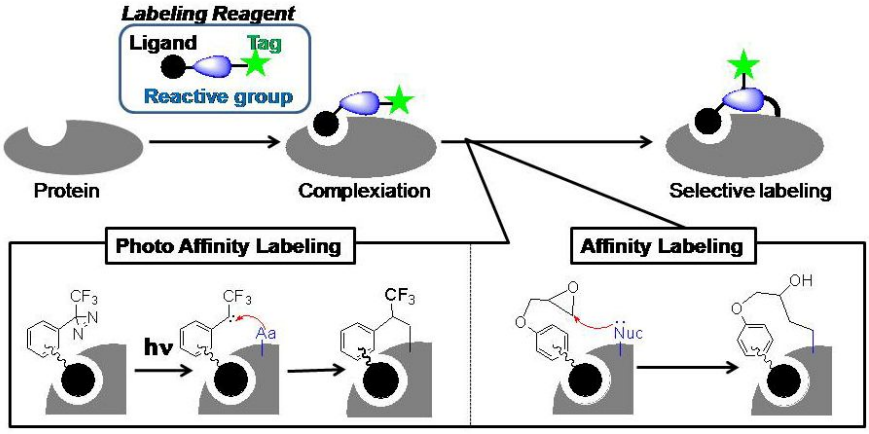


Figure 1. Schematic illustrations of affinity labeling.

Recently, the affinity labeling is elaborately applied to proteomics.[2] For example, Cravatt and coworkers developed a method so-called Activity Based Protein Profiling (ABPP), using active site-directed chemical probes based on appropriate suicide substrates, to monitor an enzyme function in complex biological systems.[3] A variety of ABPP probes have been developed for several enzyme classes, such as serine proteases and cysteine proteases, and have been used to the high-throughput analysis of enzyme activities (Fig.2). In addition, they succeeded in the inhibitor screening by the competitive ABPP (Fig.3).

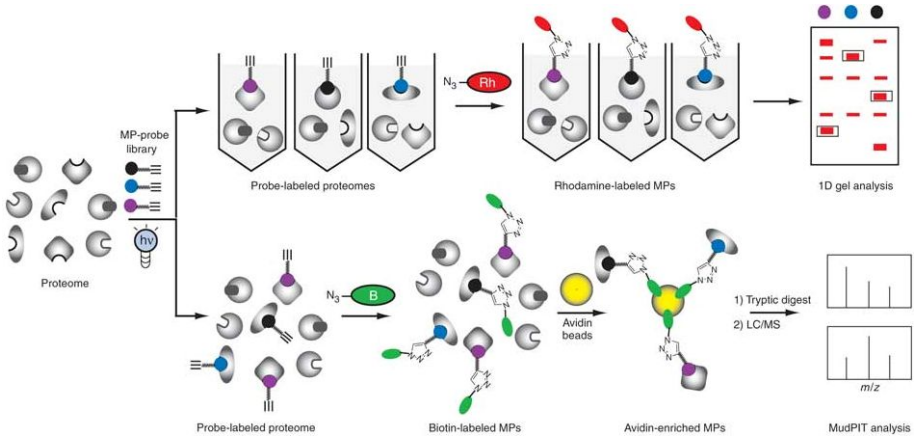


Figure 2. Functional proteomic analysis of enzyme activities using ABPP.[3]

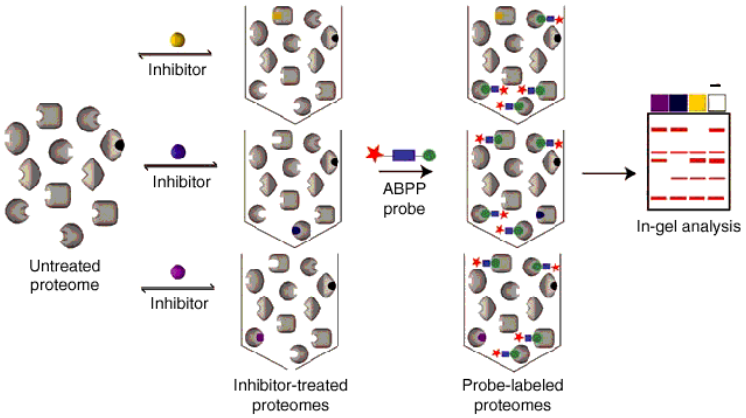


Figure 3. Inhibitor screening by the competitive ABPP.[3]

The affinity labeling for protein can be applied not only for analyzing proteins but also for functionalizing proteins. Takaoka and Hamachi developed Post-Affinity Labeling Modification (P-ALM) based on the affinity labeling (Fig.4).[4] *Human carbonic anhydrase II* (hCAII) was labeled at the proximity of the active site using the epoxide-based P-ALM reagents, and subsequently, the schiff base exchange reaction was carried out in order to cleave the masking ligand. This exchange reaction provided the attachment of the functional molecular to hCAII as well as the recovery of the enzyme activity. As the result, the modified hCAII acted as a fluorescent biosensor estimating the affinity of inhibitors for hCAII.

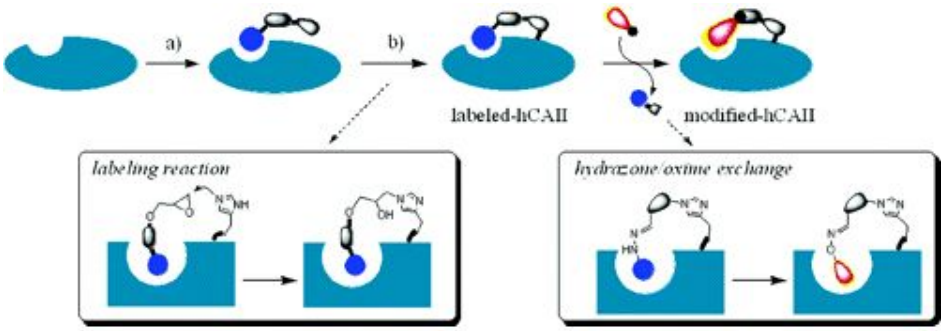


Figure 4. P-ALM Scheme for the selective modification of hCA II. [4]

2-1-2. Selective chemical labeling of lectins

Selective labeling of sugar binding proteins (lectins) should be valuable for investigating the biological roles of lectins, as well as for converting lectins to useful molecular tools for analyzing saccharides. For example, Hatanaka and coworkers succeeded in the photo-affinity labeling of lectins by using a sugar-tethered photo-labeling reagent (Fig.5a).[5] Shin and coworkers synthesized a photo-affinity labeling reagent which had a multivalent sugar ligand. By using this labeling reagent, lectins can be selectively labeled in the presence of biological impurities (Fig.5b).[6] Ilver and coworkers designed a labeling reagent based on a glycoprotein-mimic, and they successfully labeled Adhesin of *Helicobacter pylori* on the cell surface (Fig.6).[7]

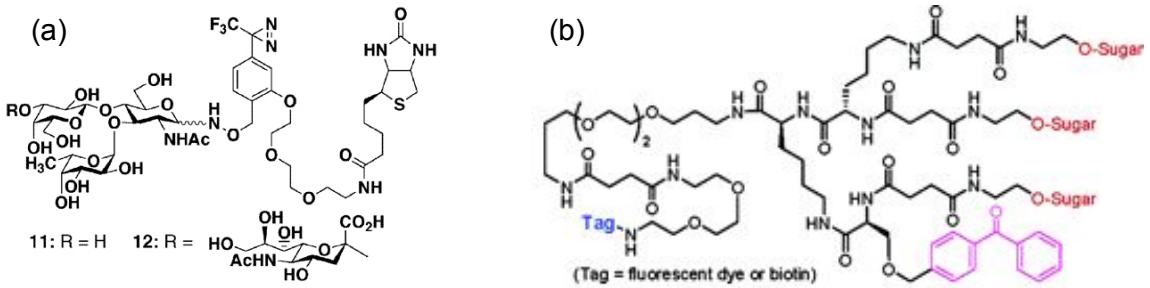


Figure 5. Molecular structures of Photo Affinity Labeling Reagents by Hatanaka (a)[5] and Shin (b).[6]

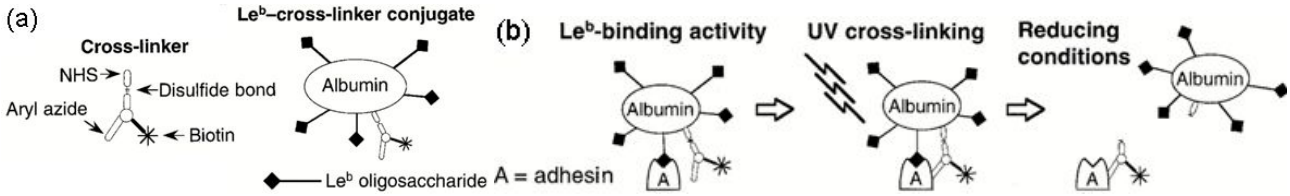


Figure 6 (a) A multifunctional cross-linking agent was conjugated to the Le_b glycoconjugate. (b) The labeling for the adhesin of *H. Pylori*.[7]

As mentioned in chapter 1, Hamachi and coworkers constructed fluorescent sugar biosensors by applying P-PALM to lectins. Furthermore, Hamachi and Nakata incorporated a phenyl boronic acid derivative as a sub-binding site into lectin, and they successfully gave a new sugar selectivity to the lectin by the cooperative recognition of the lectin and the sub-binding site (Fig.7).[8] They demonstrated that the lectin tethering the sub-binding site could act as a fluorescent oligo-saccharide sensor which responded only to the specific oligo-saccharides. By using P-PALM, Toshima and coworkers attached to lectin an anthraquinone derivative which can photodegrade saccharides. The anthraquinone-lectin hybrid cleaved specific saccharides which bind to the lectin by photoirradiation (Fig.8).[9]

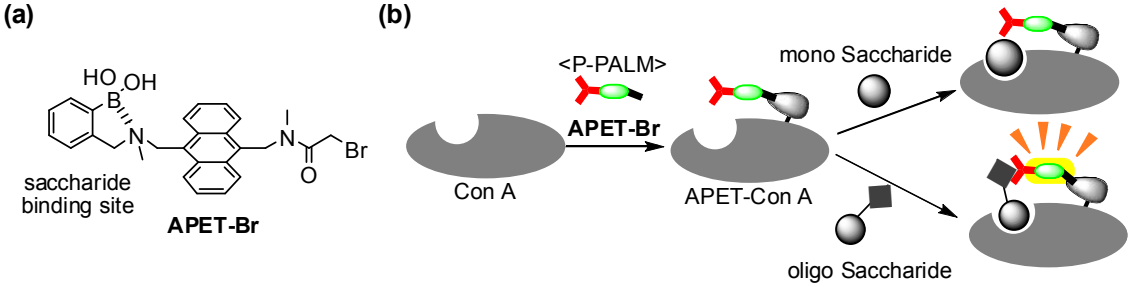


Figure 7. (a) Structures of sub-binding site. (b) Schematic illustration of oligo-saccharide sensor.[8]

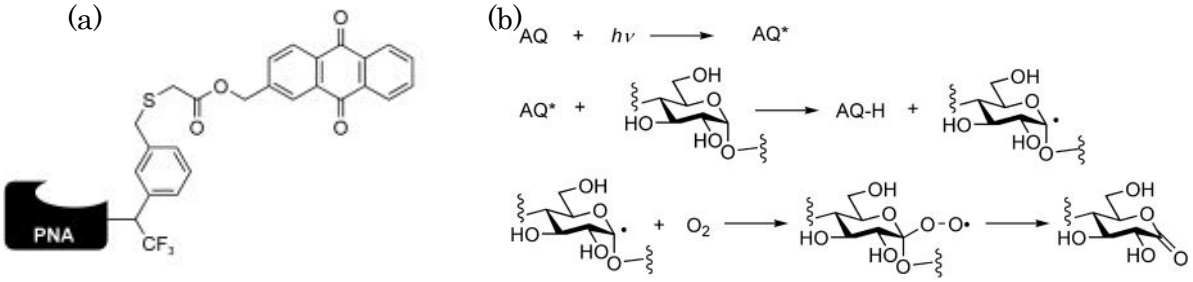


Figure 8. (a) Anthraquinone modified PNA. (b) Scheme of the cleavage of polysaccharide.[9]

2-1-3. A new protein labeling method using ligand-tethered catalyst

As mentioned above, the selective chemical labeling for proteins and lectins is an important technology. But the affinity labeling and the photo-affinity labeling, which are major methods in the selective chemical labeling, have the respective drawbacks such as the low labeling yield and the limitation of applicable proteins. Here, I propose a new strategy for chemical labeling using a ligand-tethered catalyst. This strategy is conceptually illustrated in Fig.9.

(a) When the ligand-tethered catalyst is bound to a target protein, the catalyst unit closes to protein.

(b) Since a reactive species can be generated by the catalyst on the protein surface, the labeling reaction to the proximity of the ligand binding site is carried out.

(c) The ligand-tethered catalyst is removed from the protein.

Here I used 4-Dimethylaminopyridine (DMAP, Fig.10a) as a catalyst unit. DMAP is a well-established acyl transfer catalyst which can activate an acyl ester for transfer to a nucleophilic residue.[10] In order to direct the acyl transfer to a lectin of interest, DMAP was connected to a suitable saccharide ligand having a high affinity to the target lectin. It is reasonably expected that the DMAP unit close to the sugar binding pocket of the lectin according to the sugar-lectin interaction, so that the acyl group activated by the anchored DMAP is transferred to a residue of the target lectin on the proximal lectin surface.

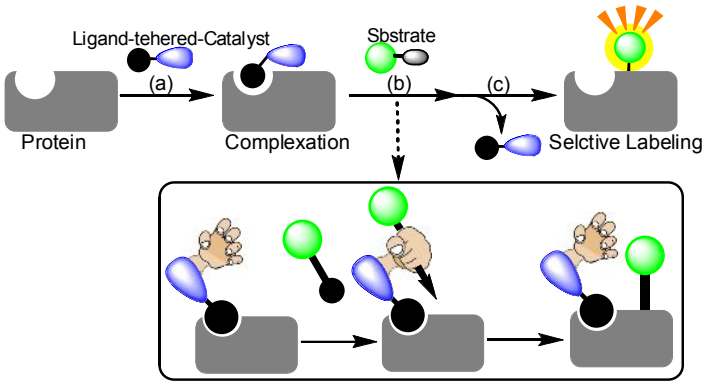


Figure 9. Scheme of selective protein labeling by ligand-tethered catalyst.

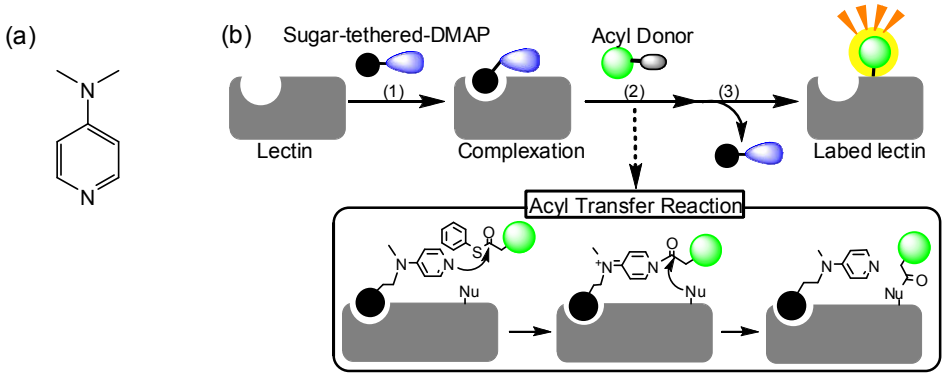


Figure 10. (a) Molecular structure of DMAP. (b) Scheme of selective lectin labeling by sugar-tethered DMAP and acyl donor.

2-2. Results

2-2-1. Site selective labeling of Congerin II in test tube

To test this idea, congerin II (Cong II), an animal lectin having Lactose/LacNAc selectivity was employed.[11] A series of sugar-tethered DMAPs and acyl donors were synthesized (Fig.11). The acceleration of the acyl transfer from the acyl donor **5** to Cong II in the presence of Lac-DMAP **1** was clearly shown by MALDI-Tof mass spectroscopy of the reaction mixture and UV-vis spectroscopy of the purified fluorescein-labeled Cong II (FL-Cong II, Fig. 12). It is considered that hydrolysis of the activated acyl intermediate is a side-reaction highly competitive with the acyl transfer, in aqueous solution. Indeed, we found that in the absence of Cong II, Lac-DMAP **1** induced a significant degree of thioester hydrolysis of **5** (48 %), but hydrolysis scarcely occurred without **1** (3%, Table 1). However, the labeling of Cong II by the acyl transfer reaction was remarkably facilitated by Lac-DMAP **1** to give the FL-Cong II in 35% yield. This can be attributed to the active intermediate (acylated DMAP) being produced predominantly in the vicinity of the lectin surface by virtue of the saccharide-lectin interaction and its rapid reaction with a nucleophilic amino acid residue located on the lectin surface, with suppression of the nonproductive hydrolysis. This is supported by the experimental result that the acyl transfer did not occur (less than 1%) in the presence of excess saccharide ligand (50 mM lactose), because of inhibition of the interaction of Lac-DMAP **1** with Cong II (Fig. 12).

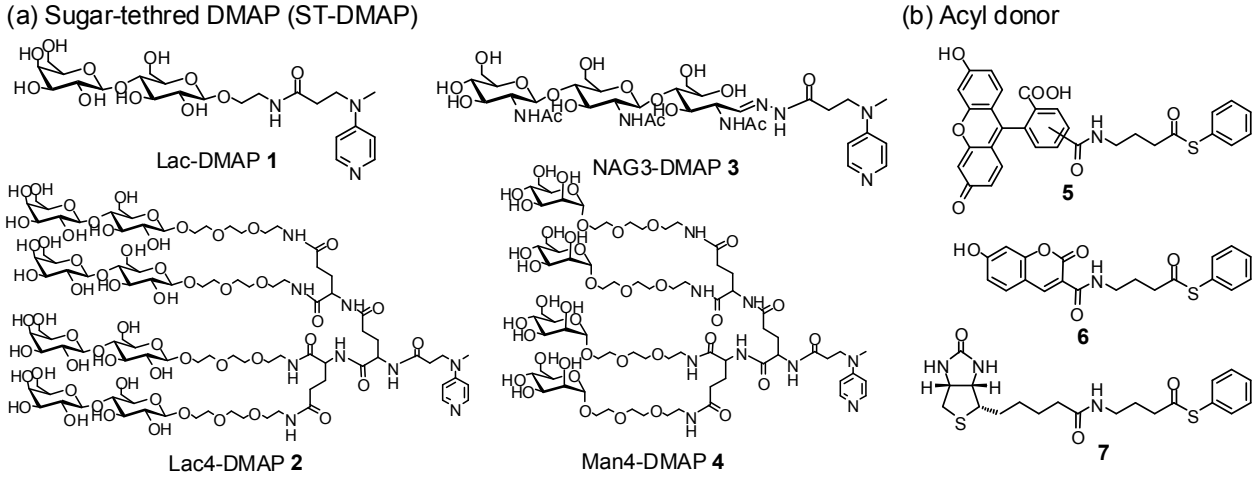


Figure 11. Molecular structures of sugar-tethered DMAPs **1-4** (a) and acyl donors **5-7** (b).

Table 1. Modification yields of lectins and hydrolysis yields of acyl donor **5** in the presence or absence of Lac-DMAP **1**

		Modification yield	Hydrolysis yield
Yields	with Lac-DMAP **1**	35±2%	48%
	without Lac-DMAP **1**	1.6±0.2%	3%
Ratio of yield with / without **1**		22	16

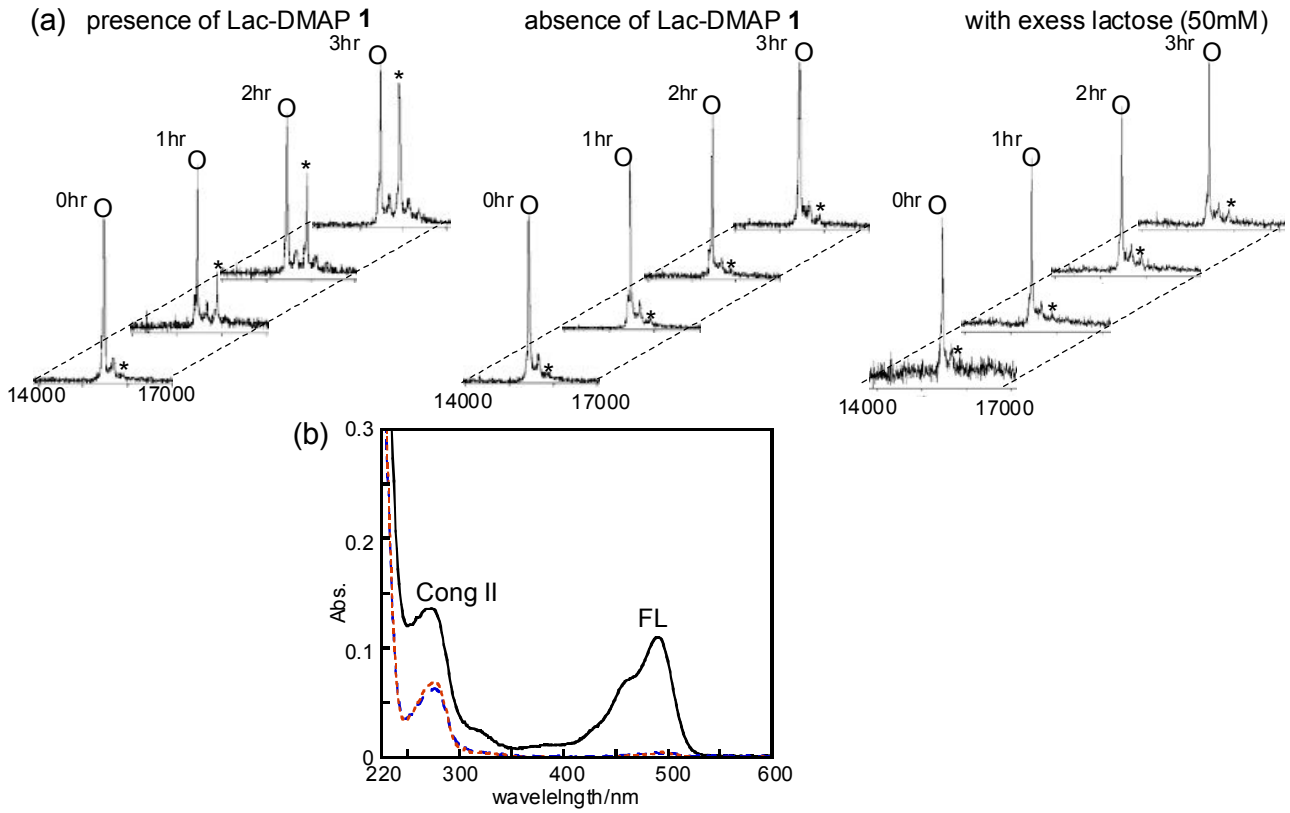


Figure 12. (a) MALDI-TOF mass spectra of reaction mixture of Cong II (10μM) with Lac-DMAP **1** (50μM) and acyl donor **5** (50μM) (O ; native Cong II (M_w 15,330), ＊; FL-Cong II (M_w 15,770)).
(b) Absorption spectra of fluorescein labeled Cong II in the presence (black line), the absence (red dot line) of Lac-DMAP **1** and with excess saccharide ligand (blue dash line).

2-2-2. Determination of the labeling site of Cong II

The labeling site for Cong II using Lac-DMAP **1** and the optimal acyl donor **6** having a coumarin fluorophore was identified by conventional peptide mapping experiments with HPLC and MALDI-Tof analysis. It was confirmed that a single peptide fragment was fluorescent among the Trypsin-digested peptides as monitored by HPLC (Fig.13). The mass/mass experiments for the isolated fluorescent peptide determined that the modified site was Tyr51 (Fig.14). Thus, it is clear that the OH group of Tyr51 was the main acyl acceptor in the present labeling reaction. Crystallographic analysis told us that this Tyr is located near the lactose binding pocket (Fig.15).

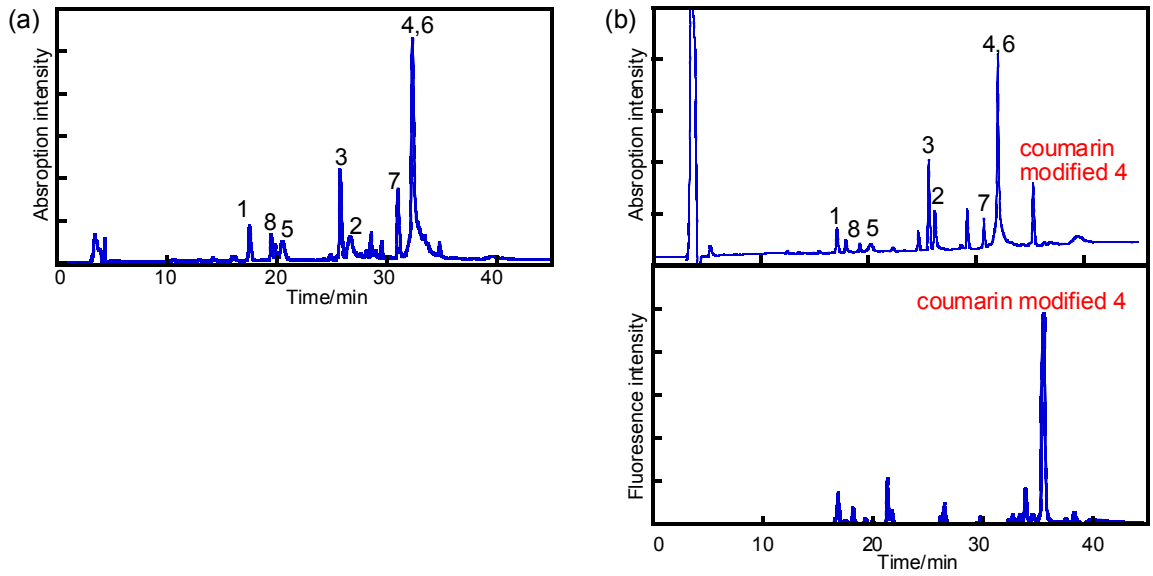


Figure 13. HPLC analysis of native Cong II (a) and coumarin labeled Cong II (b) digests monitored by UV detector (220nm, upper) and fluorescence detector (λ_{ex} = 360 nm, λ_{em} = 410 nm, lower).

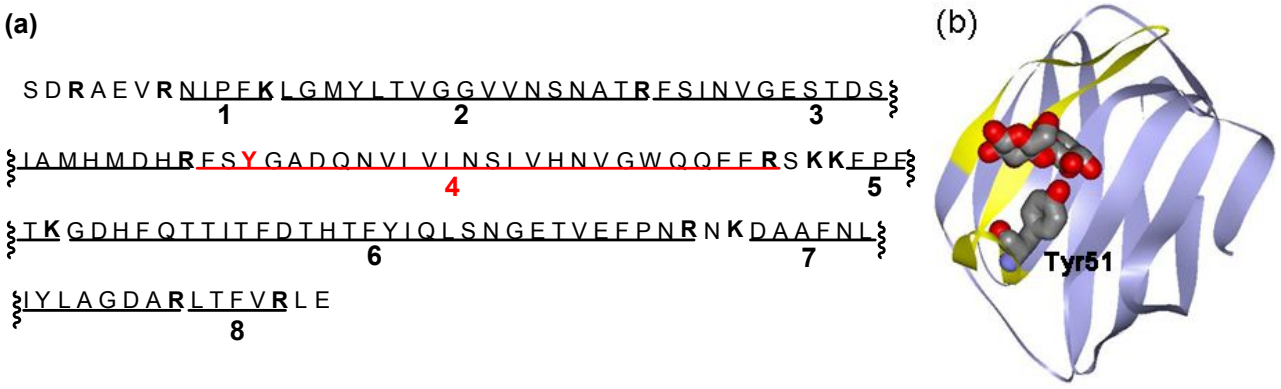


Figure 14. (a) Full amino acid sequence of Cong II. Fragments 1-8 represent the fragments of Cong II hydrolyzed by trypsin. (b) Structural model of the complex of Cong II and lactose (PDB ID: 1is4). The fragment isolated by HPLC : yellow, the modified amino acid (Tyr 51) : stick style.

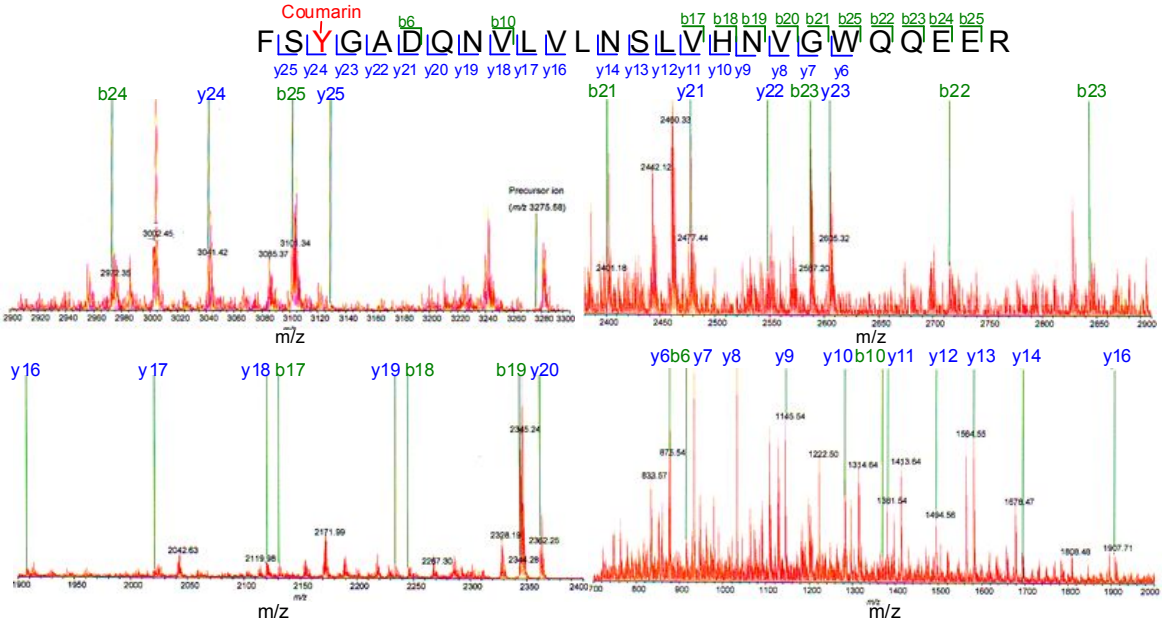


Figure 15. Tandem mass-mass analysis of the coumarin modified fragment.

2-2-3. Sugar binding study of the labeled Cong II.

Restoration of the saccharide binding capability of the modified Cong II was clearly demonstrated by the fluorescent binding assay for various saccharides. The bimolecular fluorescence quenching and recovery technique (BFQR, Fig.16) was used for fluorescein-labeled Cong II (FL-Cong II) prepared by the Lac-DMAP method. Fluorescence of FL-Cong II was initially quenched by dabcyl-appended lactose (Dab-Lac **10**), and then various saccharides were added to the quenched solution. Fluorescence intensity was recovered by the addition of lactose or galactose, which can bind to Cong II, indicating that these sugars replace the Dab-Lac. In contrast, maltose and mannose did not cause the fluorescence recovery because of their lack of binding to Cong II. These results indicate that the binding pocket of Cong II was open to retain almost the same saccharide selectivity after labeling.

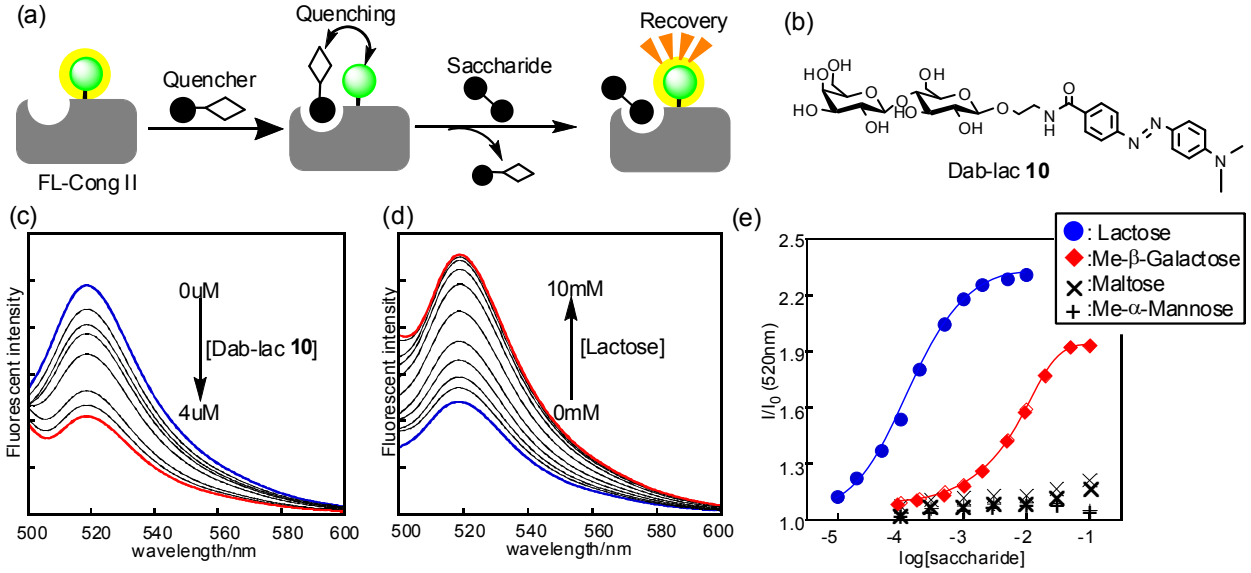


Figure 16 (a) Schematic illustration of Bimolecular Fluorescence Quenching and Recovery (BFQR). (b) Structure of Dab-lac **10**. (c) Fluorescence quenching process of FL-Cong II upon addition of **10**. (d) Fluorescence recovery of FL-Cong II by the addition of Lactose in the presence of **10** (4uM). (e) Fluorescence titration plot of the relative intensity (I/I_0) of FL-Cong II versus the saccharide concentration (log[saccharide]). Titration conditions; 0.2μM FL-Cong II, pH 7.5, 20 °C, λ_{ex} = 488nm.

2-2-4. Labeling other lectins having different sugar selectivities

The present strategy has the advantage that one can flexibly switch the target lectin by changing the saccharide structure of the ligand-tethered DMAP. For example, the acylation of wheat germ agglutinin (WGA, GlcNAc/NeuNAc selective lectin) or concanavalin A (Con A, Man/Glc selective lectin) was successfully carried out using NAG3-DMAP **3**, and Man4-DMAP **4**, respectively (Fig.16, Table 2). The high target-selectivity was clearly confirmed by the experiments with a mixture including three kinds of lectins in the presence of each saccharide-tethered DMAP. As shown in the SDS-PAGE gel images of Fig. 17, the fluorescent band of Cong II was clearly observed

in lane 3 by the acylation of the lectin mixture catalyzed by Lac4-DMAP **2**, whereas the other two bands coming from other lectins were not clearly observed. On the other hand, the fluorescent WGA band (lane 2) or Con A band (lane 1) selectively appeared by the fluorescence imager, where the NAG3- or Man4-tethered DMAP was used as the catalyst, respectively.

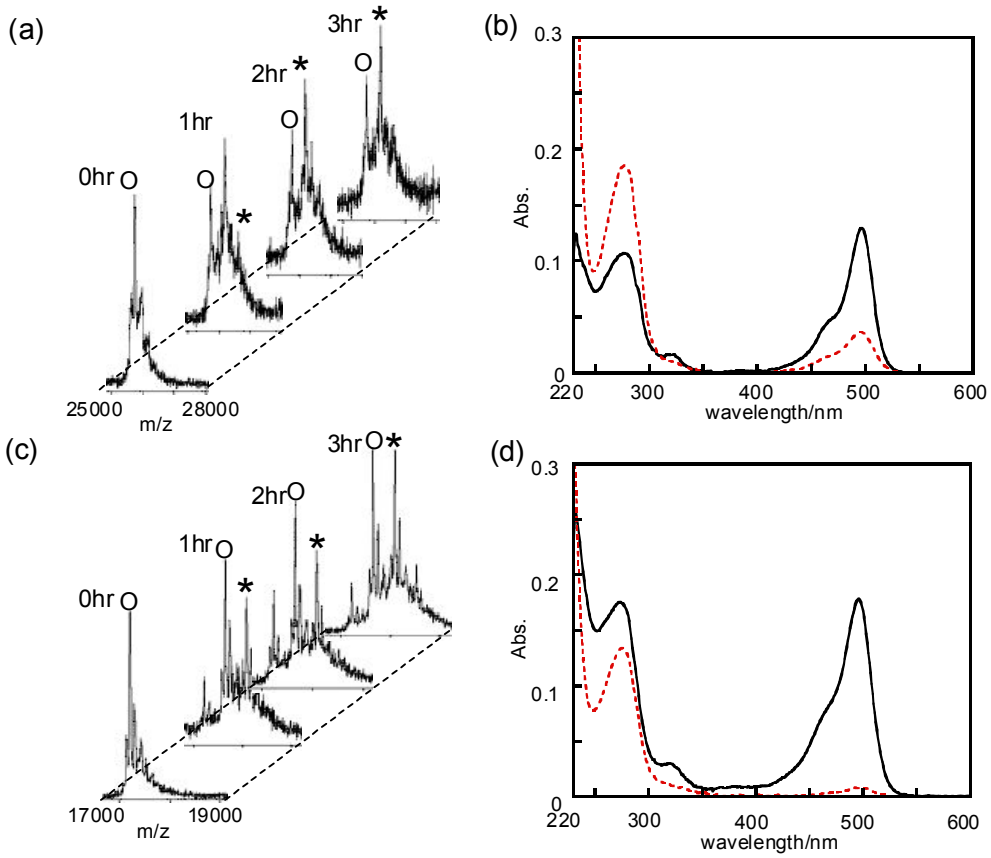


Figure 16. (a) MALDI-TOF mass spectra of reaction mixture of Con A and Man4-DMAP **4** and acyl donor **5** (O ; native Con A (M_w 25,600), ＊; FL-Con A (M_w 26,040)). (b) Absorption spectra of fluorescein labeled Con A in the presence (black line), the absence (red dot line) of **4**.
(c) MALDI-TOF mass spectra of reaction mixture of WGA and NAG3-DMAP **3** and acyl donor **5** (O ; native WGA (M_w 17,100), ＊;FL-WGA (M_w 17,540)). (d) Absorption spectra of fluorescein labeled WGA. in the presence (black line), the absence (red dot line) of **3**

Table 2. Modification yield of lectins in the presence or absence of sugar-tethered DMAPs.

Lectin		Cong II	WGA	Con A
ST-DMAP		**1**	**3**	**4**
Modification yield	with ST-DMAP	35%	62%	74%
	without ST-DMAP	1.6%	2.2%	11%
Ratio of yield with / without ST-DMAP		22	28	6.7

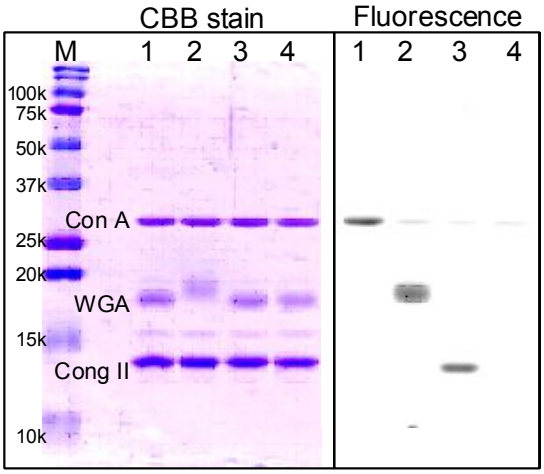


Figure 17. Selective acylation in lectin mixture. Protein modification was detected by using SDS-PAGE: left, CBB staining; right, coumarin fluorescence image. Lane = M: protein marker, 1: Man4-DMAP **4**, 2: NAG3-DMAP **3**, 3: Lac4-DMAP **2**, 4: absence of DMAP. Reaction conditions: 0.1 mg/mL lectins (Con A, WGA, and Cong II), 50μM sugar-tethered DMAPs **2**-**4**, 50μM acyl donor **8,** pH 8.0, 3h, 25 °C.

2-2-5. Selective labeling for a lectin in the crude mixture

This method was also applied to crude mixtures such as cell lysates of *E. coli* cells overexpressing Cong II. As shown in the SDS-PAGE gel image (Fig.18b), the only one band was fluorescent which is ascribed to the coumarin-labeled Cong II, indicating that the labeling using the ligand-tethered DMAP is selective even in the presence of many other proteins. Interestingly, the in-gel fluorescence intensity was greater when catalyzed by Lac4-DMAP **2** which has four Lactose units (lane 3) than with the mono-valent Lac-DMAP **1** (lane 2), suggesting selective labeling is facilitated by multivalent binding, in such a crude sample.[12] Labeling of Cong II did not take place significantly in the presence of excess amounts of saccharide ligand (50 mM lactose, lane 4), or without DMAP ligands (lane 5).

More interestingly, we can selectively label a lectin in an animal tissue lysate using this method. The lysate of the skin mucus of *Conger eel* contains two kinds of galectins (congerin I and II). The congerin was labeled with biotinylated acyl donor **9** in the presence of the catalyst Lac4-DMAP **2**. SDS-page (Fig.18c) demonstrated that the naturally occurring Congerin was labeled (lane 7), which was analyzed by blotting and the subsequent diaminobenzidine (DAB) staining. In contrast, no labeling occurred in the presence of other DMAP derivatives such as Man4-, or NAG3-DMAP (lane 8, 9), in the coexistence of Lac4-DMAP with excess of saccharide ligand (lane 10), or in the absence of sugar-tethered DMAP (lane 11).

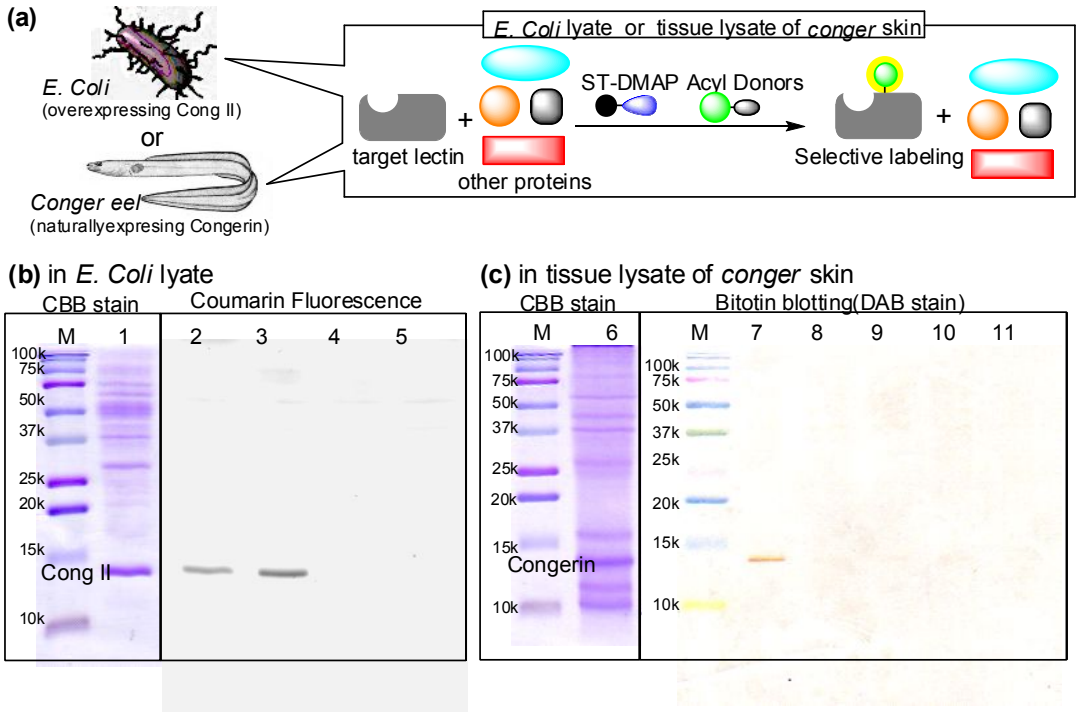


Figure 18. (a) Schematic illustration of selective acylation of lectin in crude mixtures. Protein modification was detected by SDS-PAGE: CBB staining (lane M, 1, 6); coumarin fluorescence (lane 2-5); biotin blotting (lane 7-11). (b) Coumarin labeling in cell lysates of *E. coli* overexpressing Cong II. (c) Biotin labeling in tissue lysates of conger skin mucus. Lane = M: protein marker, 1: whole proteins of *E. coli*, 2: Lac-DMAP **1**, 3: Lac4-DMAP **2**, 4: Lac4-DMAP **2** with 50mM Lactose, 5: in the absence of DMAP, 6: whole proteins of conger skin mucus, 7: Lac4-DMAP **2**, 8: Man4-DMAP **4**, 9: NAG3-DMAP **3**, 10: Lac4-DMAP **2** with 50mM Lactose, 11: in the absence of DMAP.
Reaction conditions: 50μM sugar-tethered DMAPs, 50μM acyl donor, pH 8.0, 3h, 25 °C.

2-3 Discussion

2-3-1. Catalytic property of ST-DMAP.

DMAP is a representative acyl transfer catalyst that is predominantly used in organic solvent. But the acyl transfer reaction using DMAP in water have not been conducted because the active intermediate consisting of DMAP is immediately hydrolyzed by the nucleophilic attack of H_2O (Fig.19b). In fact, ST-DMAP greatly facilitated the hydrolysis of acyl donor in the absence of lectin (Table 1). A free ST-DMAP which is not bound to lectin does not catalyze the acyl transfer reaction, but it only accelerated the hydrolysis of acyl donors. This feature resulted in the fact that a nonspecific labeling of protein was not facilitated by an excess amount of ST-DMAP. This is an advantage for the labeling of proteins which don't have a strong affinity with a ligand such as lectins.

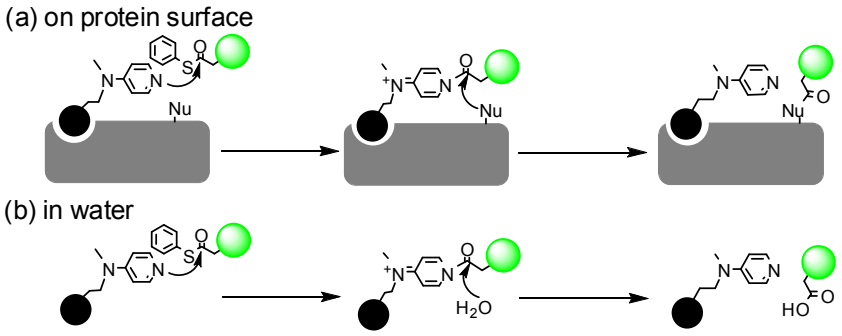


Fig. 19 Reaction scheme of acyl transfer on proteins surface (a) or in water (b).

By careful examination of amino acid residues labeled by the acyl transfer reaction, it was clear that the labeling site of Cong II was Tyr51. It was probably conceivable that a phenol ester was formed at OH group of Tyr. The basic treatment of the labeled Cong II underwent the hydrolytic cleavage of the attached fluorophore, which strongly supported the above conclusion (Fig.20). In the case of the modified Con A, the basic hydrolysis also took place, suggesting that the acylation site of Con A is probably Tyr, some as the case of Cong II. In fact, there is Tyr100 in the proximity of the sugar binding pocket of Con A. On the other hand, the basic hydrolysis of the labeled WGA did not occur. Thus, it is considered that the acylation site of WGA is not Tyr, but Thr or Ser (as an alkyl ester), or Lys (as an amide). These result indicated that the DMAP catalyzed acyl transfer reaction can modify with various nucleophilic amino acids on a protein surface.

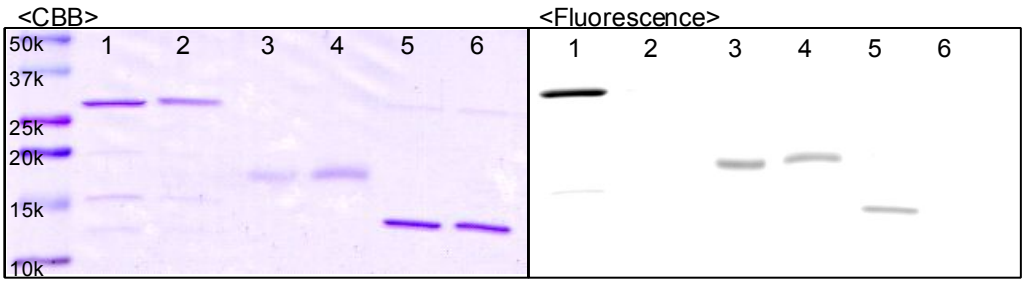


Figure 20. Hydrolytic cleavage of the coumarin labeled lectins by the basic treatment. The labeled lectins were detected by using SDS-PAGE: left, CBB staining; right, coumarin fluorescence image. Lane = M: protein marker, 1, 2: labeled Con A, 3, 4: labeled WGA, 5, 6: labeled Cong II. The labeled lectins of lane 2, 4 and 6 were treated with the basic condition (pH 12.0) at 25 °C for 1 day.

Next, we examined the multivalency effect of the DMAP unit in ST-DMAP. We compared the reactivity of the lactose-tethered DAMP (Lac-DMAP **1**) with DMAP dimer (Lac-2DMAP **11**) and DAMP trimer (Lac-3DMAP **12**) in the acyl transfer reaction for Cong II (Fig.21). As shown in Table 3 and Fig.22, the oligomeric DMAP catalysts (**11**, **12**) accelerated the labeling reaction, relative to the monomeric DMAP catalyst (**1**). Particularly, the reaction using trimeric DMAP (**12**) was fastest (labeling yield was over 60 % at 15min). It is clear that the oligomerization of DMAP unit is useful for accelerating acyl transfer reaction.

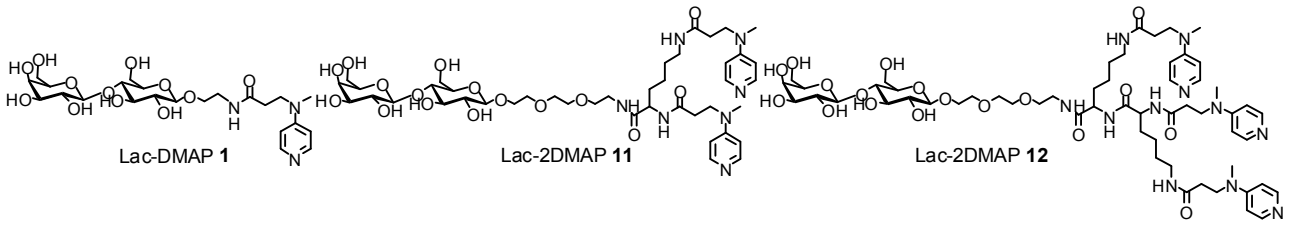


Figure 21. Molecular structures lactose-tethered multi DMAPs **1**, **11**, **12**.

Table 3. Modification yields of Cong II using acyl donor **5** and multi DMAPs **1**, **11**, **12** for 15min

	Lac-DMAP **1**	Lac-2DMAP **11**	Lac-3DMAP **12**	Lac-3DMAP **12** with 30mM Lactose	without DMAP
Modification yield	14±1%	51±1%	64±2%	1.2±0.2%	0.3±0.2%

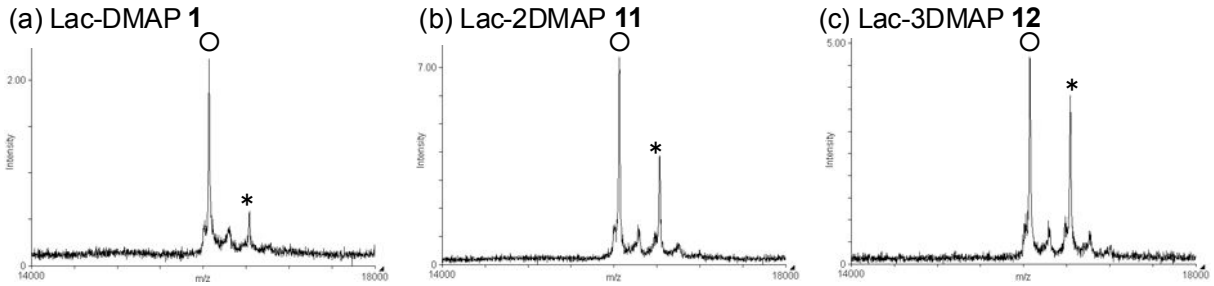


Figure 22. MALDI-TOF mass spectra of labeled Cong II using acyl donor **5** and Lac-DMAP **1** (a), Lac-2DMAP **11** (b) or Lac-3DMAP **12** (c) (○; native Cong II (M_w 15,330), ＊;FL-Cong II (M_w 15,770)).

2-3-2. The study of the reactivity of acyl donor.

The acyl transfer reaction for Cong II directed by Lac-DMAP **1** was optimized by careful selection of the acyl donor structure particularly in the carboxylic ester (acyl group) part and the thiol (leaving group) part (Fig.23). As shown in Table 4 (left side) and Fig.24, acyl donor **7** containing benzyl thioester was not reactive enough to be activated by DMAP (less than 1%), compared to thiophenyl esters. Among the thiophenyl esters, on the other hand, acyl donor **6** which includes an alpha-amino acid structure was much too reactive, so that it caused a considerable amount of nonspecific acylation (11%). We found that acyl donor **5**, containing the gamma-amino acid moiety, showed enhanced specificity by reducing nonspecific reactions. The selectivity, that is the ratio of the yield catalyzed by Lac-DMAP **1** over that of the non-catalyzed reaction, was 22 for the gamma-amino thiophenyl ester **5**, a value 20-fold greater than that for the alpha-amino thiophenyl ester **6**.

These results were confirmed by the hydrolysis reaction of acyl donors **5**-**7** (Talbe 4, right side). The order of hydrolysis rate without DMAP was in good agreement with the order of the nonspecific modification yield of lectins. Acyl donor **6** was so reactive that it did not need the activation by DMAP in the hydrolysis under the present conditions. This may cause a considerable amount of the nonspecific acylation for lectin. The reactivity of acyl donor **7** was very low, so that the activation with DMAP hardly underwent.

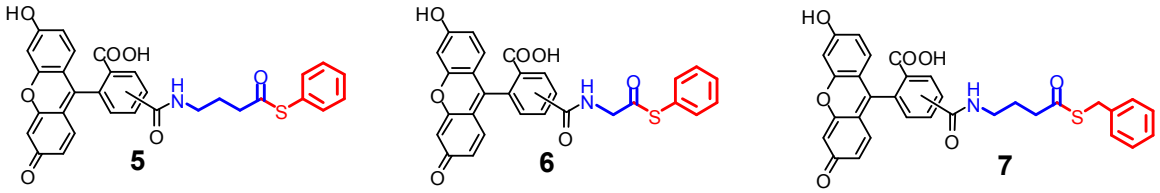


Figure 23. Molecular structures of acyl donors **5**-**7**.

Table 4. Modification yields for Cong II and hydrolysis yields of acyl donors **5**-**7**.

		Modification yield for Cong II			Hydrolysis yield		
Acyl donor		**5**	**6**	**7**	**5**	**6**	**7**
Yields	with Lac-DMAP **1**	35±2%	12%	<1%	48%	95%	<1%
	without Lac-DMAP **1**	1.6±0.2%	11%	<1%	3%	93%	<1%
Ratio of yield with / without **1**		22	1.1	-	16	1.0	-

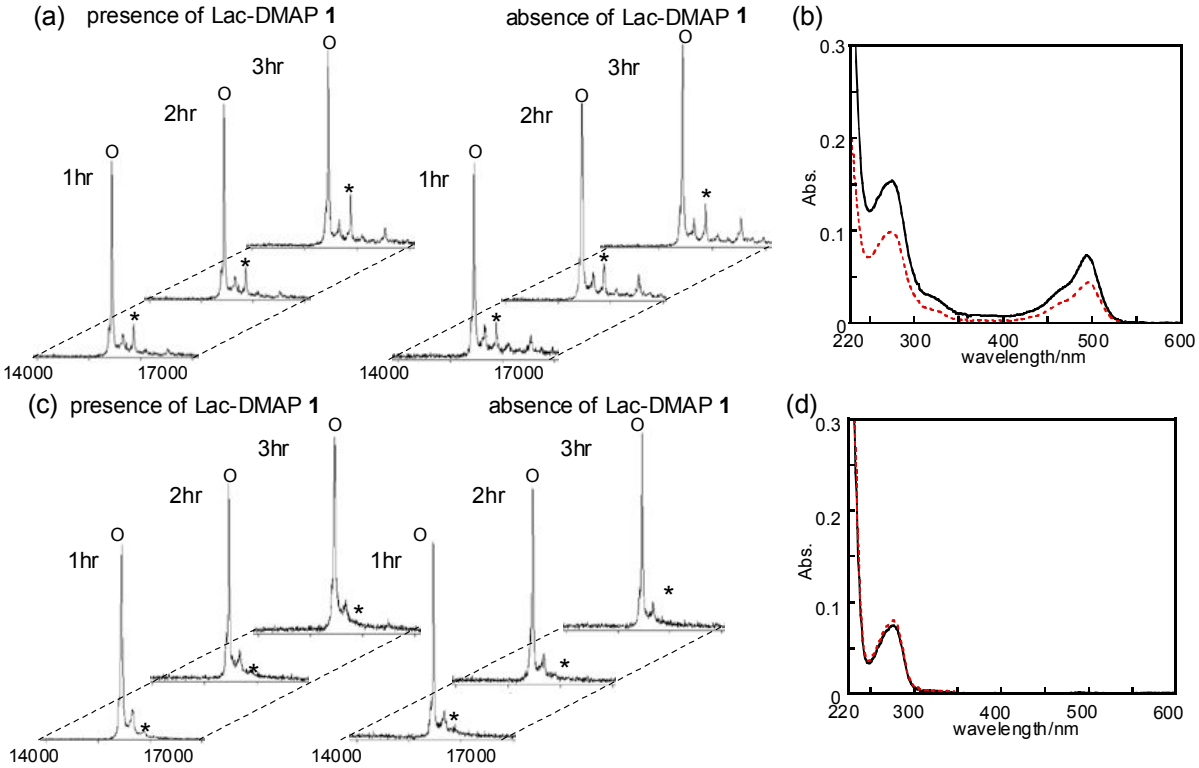


Figure 24. MALDI-TOF mass spectra of reaction mixture of Cong II with acyl donor **6** (a) or **7** (c) (○ ; native Cong II (M_w 15,330), ＊ ; FL-Cong II (M_w 15,770)). Absorption spectra of FL-Cong II labeled by acyl donor **6** (b) or **7** (d) in the presence (black line), the absence (red dot line) of Lac-DMAP **1**.

On the basis of the optimal acyl donor structure (γ-aminobutyric acid - thiophenol ester), we can easily design various acyl donors having other functional molecules. It was confirmed that the specific labeling occurred in the case of the acyl donor **8** having a coumarin fluorophore or acyl donor **9** having a biotin tag (Fig.25).

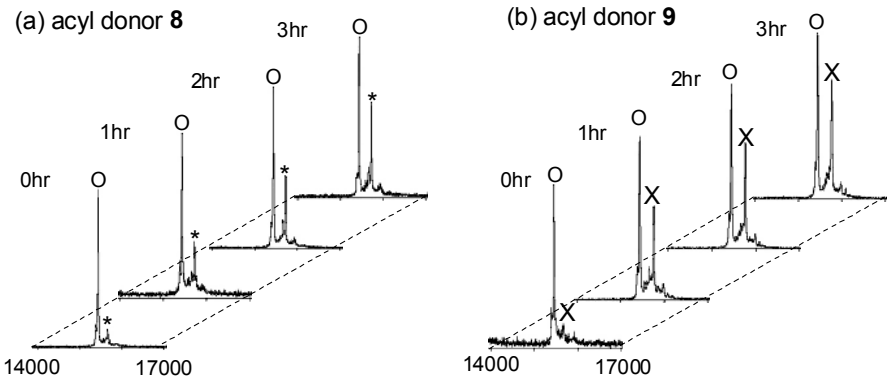


Figure 25. MALDI-Tof mass spectra of reaction mixtures containing Cong II, Lac-DMAP **1**, and acyl donor **8** (a) or **9** (b). (○; native Cong II (M_w 15,330), ＊; Coumarin modified Cong II (M_w 15,600), ×; Biotin modified Cong II (M_w 15,640)).

Because thioester is not perfectly stable, the degradation of acyl donor was concerned during the reaction. In fact, the acyl donor **8** was rapidly hydrolyzed in samples of biological origin such as conger skin lysate (Fig.26b). I found that the hydrolysis of the acyl donor can be inhibited by the addition of 4-(2-aminoethyl) benzenesulfonyl fluoride (ABSF, Fig.26a), an irreversible inhibitor for serine proteases. In the presence of 1 mM ABSF, the hydrolysis was scarcely observed (Fig.26c). Using this condition, the acyl transfer reaction can be carried out in lysate sample. But the instability of the acyl donors let us tell that the ST-DMAP catalyzed protein labeling in cell or in vivo is impossible at present.

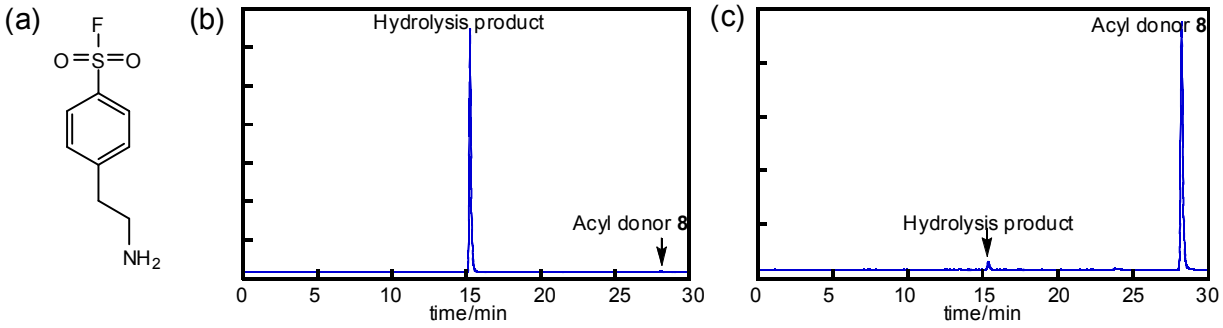


Figure 26. (a) Molecular structure of ABSF. HPLC analysis of acyl donor **8** in conger skin lyste with (c) or without (b) 1mM ABSF.

2-3-3. The advantage of this labeling method.

The affinity labeling and photo-affinity labeling have been established and wildly used as selective chemical labeling methods. In order to understand the advantages and disadvantages of the present acyl transfer reaction, I compared this method with them (Table 5).

(Generality) In the affinity labeling, target amino acids are limited by the types of reactive groups of labeling reagents. In the acyl transfer reaction and the photo-affinity labeling, broad kinds of amino acids can be modified.

(Efficiency) The labeling yield of the photo-affinity labeling (less than 20 %) is generally lower than that of the affinity labeling. As shown in table 2, the labeling yield of acyl transfer reaction is about 50 % in this study.

(Removing ligand) In the affinity labeling and the photo-affinity labeling, the ligand moiety is covalently bound to the target protein after labeling reaction. In acyl transfer reaction, in contrast, the ligand moiety is noncovalently bound to the target protein, so that it can be easily removed. Thus, the acylated proteins retain the activity, which is advantageous for some cases such as the construction of biosensors.

In another point, the acyl transfer is powerful for the labeling of proteins which don't have strong affinity with ligand such as lectins. In summary, since these methods have advantages and disadvantages each other, we need to select the suitable one corresponding to its purpose.

Table 5. The comparison of the futures in these labeling methods

	Affinity lableing	Photo affinity lableing	Acy transfer reaction
Genelarity for amino acid	Low	High	High
Labeling yield	<100%	<20%	35%～75%
Reamoval of liand module	Difficult (covalent bonding)	Difficult (covalent bonding)	Easy (noncovalent bonding)

3-4. Conclusion

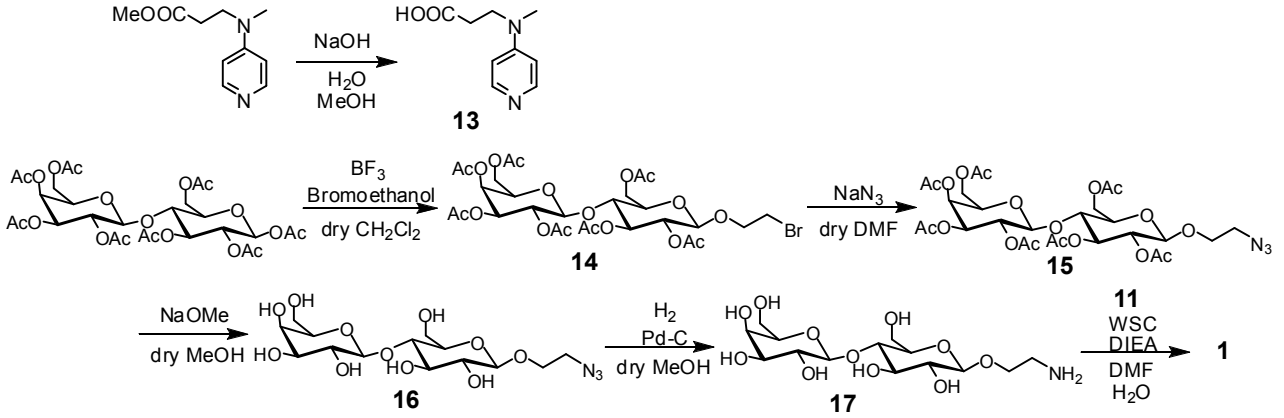
We have developed a novel method for lectin-selective covalent labeling using ligand-directed acyl transfer catalyzed by DMAP. The method was successfully applied not only to a purified lectin, but also to crude mixtures, such as lysates of recombinant bacterial cells or animal tissue expressing the target lectin. Thus, it is expected that this method can be used for profiling lectin-like proteins in cells or in tissues without corresponding antibodies. In addition, fluorescent lectins prepared by this method are useful for the scaffold of biosensor detecting various saccharides. This chemical strategy may be generally applicable to receptor proteins other than lectin, by changing the ligand part for targeting a protein of interest and/or by replacing the catalyst part for other chemical reactions available.

2-5. Experimental section

General Comments for Synthesis.

All chemical reagents were purchased from commercial suppliers (Aldrich, TCI, or Wako) and used without further purification. All air- or moisture-sensitive reactions were performed with distilled solvents (DMF, dichloromethane, or MeOH) under a nitrogen or argon atmosphere.

Synthesis of Lac-DMAP 2-1



Compound (**13**):

3-(Methyl-4-pyridylamino) propionic acid methyl ester was synthesized as previously reported.[13] To a solution of 3-(Methyl-4-pyridylamino) propionic acid methyl ester (240 mg, 1.23 mmol) in 2 mL of methanol was added 3 mL of 1*N* NaOH solution. After stirring for 8 h at room temperature, the reaction mixture was neutralized with HCl solution. The resulting solution was lyophilized to yield a white solid. Methanol was added to dissolve the crude product, and the insoluble salts were removed by filtration. The filtrate was evaporated under vacuum to afford **13** as a white solid (280 mg, 1.09 mmol, 89 %). ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 2.49 (t, J_H = 7.2 Hz, 2H), 3.20 (s, 3H), 3.84 (t, J_H = 7.2 Hz, 2H), 7.00 (d, J_H = 7.2 Hz, 2H), 8.07 (t, J_H = 7.2 Hz, 2H).

Compound (**14**):

2-Bromoethanol (1.1 mL, 15.0 mmol) and molecular sieves 3A (2.00 g) were added to a solution of peracetyl lactose (5.11 g, 7.50 mmol) in 20 mL of CH_2Cl_2. After stirring for 30 min at room temperature, boron trifluoride-diethyl ether (4.70 mL, 37.5 mmol) was added to the above solution. The reaction mixture was then stirred for 12 h at room temperature and diluted with 200 mL of ethyl acetate. The organic layer was washed with deionized water and dried over anhydrous Na_2SO_4. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexanes 3:2) to give **14** as a white solid (2.17g, 2.92 mmol, 39%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.98-2.15 (m, 21H), 3.42-3.46 (m, 2H), 3.62-3.64 (m, 1H), 3.78-3.82 (m, 2H), 3.88 (t, J_H = 7.2 Hz,1H), 4.08-4.13 (m, 4H), 4.48-4.54 (m, 2H), 4.90-4.98 (m, 3H), 5.12 (dd, J_H = 7.2, 3.6 Hz, 1H), 5.23 (d, J_H = 7.2 Hz, 1H), 5.35 (d, J_H = 3.6 Hz, 1H).

Compound (**15**):

To a solution of **14** (2.17 g, 2.92 mmol) in 80 mL of DMF was added sodium azide (949 mg, 14.6 mmol). The reaction mixture was stirred for 3 h at 80 °C. After filtration, the filtrate was evaporated under vacuum and diluted with 200 mL of ethyl acetate. The organic layer was washed with brine and dried over anhydrous Na_2SO_4. The solution was filtered and concentrated under vacuum to afford **15** as a white solid (2.06 g, 2.17 mmol, 100%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.97-2.15 (m, 21H), 3.29-3.47 (m, 2H), 3.62-3.70 (m, 2H), 3.80-3.88 (m, 2H), 4.06-4.14 (m, 4H), 4.49-4.57 (m, 2H), 4.90-4.97 (m, 3H), 5.11 (dd, J_H = 7.2, 3.6 Hz, 1H), 5.21 (d, J_H = 7.2 Hz, 1H), 5.36 (d, J_H = 3.6 Hz, 1H).

Compound (**16**):

To a solution of **15** (2.06 g, 2.17 mmol) in 12 mL of methanol was added 181 μL of methanol solution of sodium methoxide (217 μmol). After stirring for 4 h at room temperature, the reaction was quenched by adding ion exchange resin (Amberlite IRC-50, 2 g). After filtration, the filtrate was removed under vacuum to afford **16** as a white solid (880 mg, 2.14 mmol, 99%). ^{1}H-NMR (400 MHz, CD_3OD); 3.39-3.94 (m, 16H), 4.30 (d, J_H = 7.2 Hz, 1H), 4.34 (d, J_H = 7.2 Hz, 1H).

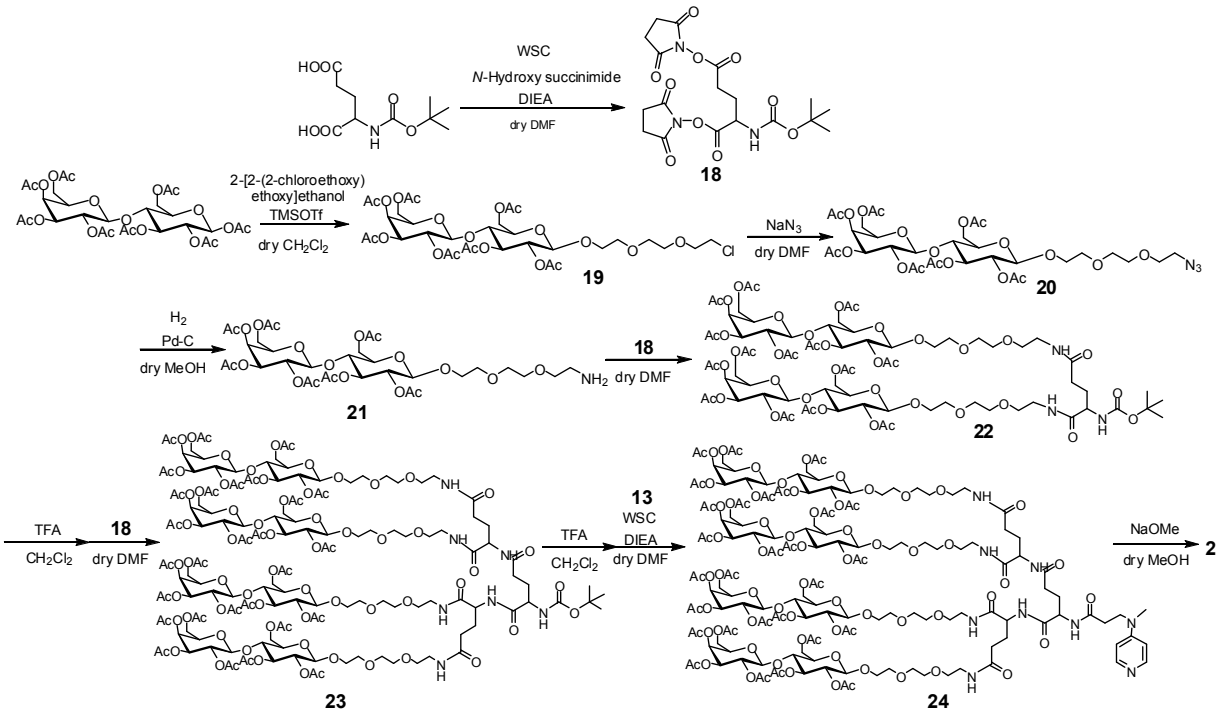
Compound (**17**):

To a solution of **16** (880 mg, 2.14 mmol) in 100 mL of methanol was added palladium-carbon (10%, 100 mg). Under H_2 atmosphere, the reaction mixture was stirred for 20 h at room temperature. After the palladium-carbon was removed by filtration, the filtrate was evaporated under vacuum to afford **17** as a white solid (709 mg, 1.84 mmol, 86%). ^{1}H-NMR (400 MHz, CD_3OD); 3.38-4.02 (m, 16H), 4.35-4.37 (m, 2H).

Lac-DMAP (**1**):

To a solution of **17** (32mg, 168 μmol) in 200 μL of DMF were added WSC (16 mg, 84 μmol) and DIEA (20 μL, 168 μmol). A solution of 3-(Methyl-4-pyridylamino) propionic acid **13** (10 mg, 56 μmol) in 60 μL of distilled water was added to the above solution. The reaction mixture was then stirred for 10 h at room temperature and the solution was purified by reversed-phase HPLC. The eluent was lyophilized and redissolved in distilled water to obtain Lac-DMAP **1** as a 300 μL aqueous solution (9.88 mM, 2.96 μmol, 5%). The RP-HPLC was performed using an ODS-A column (YMC, 250 × 10 mm) using a linear gradient of CH_3CN and 10 mM aqueous ammonium acetate from 2 / 98 to 10 / 90 over 30min: flow rate = 3 mL / min, detection at 280 nm. HRMS (FAB) calcd for [M+H$^+$] ($C_{23}H_{38}N_3O_{12}$): 548.2450; found: 548.2455.

Synthesis of Lac4-DMAP 2



Compound (**18**):

A solution of *N*-(tert-butoxycarbonyl)-L-glutamic acid (1.00 g, 4.05 mmol) in 20 mL DMF containing WSC (2.33 g, 12.2 mmol), *N*-hydroxysuccinimide (1.40 g, 12.2 mmol) and DIEA (3.40 mL, 20.3 mmol) was stirred for 12 h at room temperature. After evaporation, the oily residue obtained was dissolved in 300 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid, sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. Removal of the solvent under reduced pressure afforded the crude product as a colorless oil. Purification by flash column chromatography on silica gel (elution with $CHCl_3$/methanol 9:1) gave **18** as a white solid (830 mg, 1.88 mmol, 46%). ^{1}H NMR (400 MHz, $CDCl_3$) δ/ppm: 1.45 (s, 9H), 2.26-2.33 (m, 1H), 2.41 (br, 1H), 2.84 (br, 10H), 4.78 (br, 1H), 5.29 (br, 1H).

Compound (**19**):

2-[2-(2-chloroethoxy)ethoxy]ethanol (642 μL, 4.42 mmol) and molecular sieves 4A (1.00 g) were added to a solution of peracetyl lactose (1.00 g, 1.47 mmol) in 10 mL of CH_2Cl_2. After stirring for 1 h at room temperature, trimethylsilyl trifluoromethanesulfonate (654 μL, 2.94 mmol) was added to the above solution. The reaction mixture was then stirred for 4 h at 0 °C and diluted with 200 mL of ethyl acetate. The resulting solution was washed with sat. aqueous $NaHCO_3$ and brine, and the organic layer was dried over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution with ethyl acetate/hexanes 2:1) to give **19** as a white solid (492 mg, 626 μmol, 43%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.96-2.14 (m, 21H), 3.61-3.65 (m, 10H), 3.75 (t, J_H = 6.0 Hz, 2H), 3.84-3.92 (m, 2H), 4.05-4.15 (m, 4H), 4.47-4.50 (m, 2H), 4.56 (d, J_H = 8.8 Hz, 1H), 4.89 (t, J_H = 8.8 Hz, 1H), 4.94 (dd, J_H = 9.2, 3.2 Hz, 1H), 5.10 (t, J_H = 8.8 Hz, 1H), 5.19 (t, J_H = 9.2 Hz, 1H), 5.34 (d, J_H = 3.2 Hz, 1H).

Compound (**20**):

To a solution of **19** (480 mg, 610 μmol) in 10 mL of DMF was added sodium azide (198 mg, 3.05 mmol). The reaction mixture was stirred for 4 h at 80 °C, evaporated under vacuum, and diluted with 200 mL of ethyl acetate. The organic layer was washed with deionized water and brine, and the organic layer was dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **20** as a white solid (484 mg, 610 μmol, 100%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.96-2.14 (m, 21H), 3.39 (t, J_H = 8.2 Hz, 2H), 3.61-3.65 (m, 10H), 3.84-3.92 (m, 2H), 4.05-4.13 (m, 4H), 4.46-4.50 (m, 2H), 4.56 (d, J_H = 8.8 Hz, 1H), 4.89 (t, J_H = 8.8 Hz, 1H), 4.94 (dd, J_H = 9.2, 3.2 Hz, 1H), 5.10 (t, J_H = 8.8 Hz, 1H), 5.19 (t, J_H = 9.2 Hz, 1H), 5.33 (d, J_H = 3.2 Hz, 1H).

Compound (**21**):

To a solution of **20** (484 mg, 610 μmol) in 20 mL of methanol was added palladium-carbon (10%, 35 mg). Under H_2 atmosphere, the reaction mixture was stirred for 1 h at room temperature. After the palladium-carbon was removed by filtration, the filtrate was evaporated under vacuum to afford **21** as a white solid (464 mg, 604 μmol, 99%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.96-2.12 (m, 21H), 3.11-3.13 (m, 2H), 3.63-3.68 (m, 10H), 3.90-3.97 (m, 2H), 4.07-4.24 (m, 4H), 4.48-4.60 (m, 3H), 4.92 (t, J_H = 8.8 Hz, 1H), 4.97 (dd, J_H = 9.2, 3.2 Hz, 1H), 5.10 (t, J_H = 8.8 Hz, 1H), 5.20 (t, J_H = 9.2 Hz, 1H), 5.35 (d, J_H = 3.2 Hz, 1H).

Compound (**22**):

To a solution of **21** (310 mg, 404 μmol) in 2mL of CH_2Cl_2 containing DIEA (66.0 μL, 404 μmol) was added **18** (80.2 mg, 181 μmol). The reaction mixture was stirred overnight at room temperature, and the solution was diluted with 200 mL of CH_2Cl_2. The organic layer was washed with 5% aqueous citric acid, sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **22** as a white solid (324 mg, 168 μmol, 93% for **18**). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.42 (s, 9H), 1.96-2.12 (m, 42H), 2.25-2.30 (m, 2H), 3.37-3.43 (m, 4H), 3.56-3.60 (m, 22H), 3.86-3.93 (m, 4H), 4.05-4.15 (m, 8H), 4.48-4.56 (m, 6H), 4.89 (t, J_H = 8.8 Hz, 2H), 4.95 (dd, J_H = 9.2, 3.2 Hz, 2H), 5.10 (t, J_H = 8.8 Hz, 2H), 5.17 (t, J_H = 9.2 Hz, 2H), 5.34 (d, J_H = 3.2 Hz, 2H), 5.63 (br, 1H), 6.61 (br, 1H), 7.02 (br, 1H). MALDI-Tof mass; calcd for [M+Na^+] ($C_{74}H_{111}N_3O_{44}Na$): 1768.65; found: 1768.73.

Compound (**23**):

22 (305 mg, 175 μmol) was dissolved in 10 mL of CH_2Cl_2 containing 20% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature and evaporated under vacuum. The residue was redissolved in 10 mL of CH_2Cl_2 containing DIEA (130 μL, 780 μmol) and **18** (34.8 mg, 78.8 μmol). The reaction mixture was stirred for 3h at room temperature and the solution was diluted with 300 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid, sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **23** as a white solid (260 mg, 74.2 μmol, 94% for **18**). ^{1}H-NMR

(400 MHz, $CDCl_3$); δ/ppm: 1.43 (s, 9H), 1.96-2.15 (m, 84H), 2.25-2.34 (m, 6H), 3.37-3.43 (m, 8H), 3.55-3.69 (m, 46H), 3.87-3.93 (m, 8H), 4.05-4.14 (m, 16H), 4.48-4.55 (m, 12H), 4.89 (t, J_H = 8.8 Hz, 4H), 4.95 (dd, J_H = 9.2, 3.2 Hz, 4H), 5.10 (t, J_H = 8.8 Hz, 4H), 5.17 (t, J_H = 9.2 Hz, 4H), 5.35 (d, J_H = 3.2 Hz, 4H). MALDI-Tof mass; calcd for [M+Na^+] ($C_{148}H_{219}N_7O_{88}Na$): 3525.28; found: 3521.43.

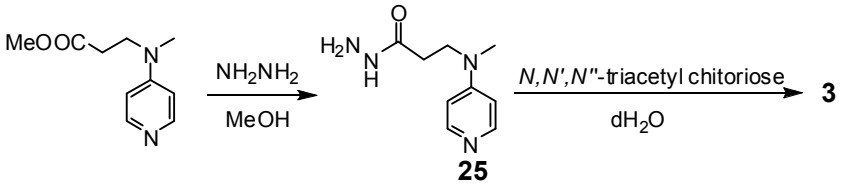
Compound (**24**):

23 (110 mg, 31.4 μmol) was dissolved in 3 mL of CH_2Cl_2 containing 20% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature and evaporated under vacuum. The oily residue was redissolved in 2 mL of DMF containing WSC (28.0 mg, 142 μmol) and DIEA (93 μL, 568 μmol), and a solution of **13** (35.8 mg, 142 μmol) in 1 mL of deionized water was added to this solution. The reaction mixture was stirred for 1h at 60 °C and the solution was diluted with 25 mL of 1% aqueous acetic acid. The aqueous solution was extracted with ethyl acetate, and the extract was washed with sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. Removal of the solvent under reduced pressure afforded **24** as a colorless oil (88.0 mg, 24.7 μmol, 80%). ^{1}H NMR (400 MHz, $CDCl_3$) δ/ppm: 1.96-2.15 (m, 84H), 2.25-2.34 (m, 6H), 2.92 (s, 3H), 3.35-3.43 (m, 8H), 3.50-3.66 (m, 46H), 3.76-3.80 (m, 2H), 3.87-3.93 (m, 8H), 4.06-4.14 (m, 16H), 4.48-4.45 (m, 12H), 4.85-4.90 (m, 4H), 4.95-4.98 (m, 4H), 5.10 (t, J_H = 9.2 Hz), 5.18 (t, J_H = 9.2 Hz, 4H), 5.35 (d, J_H = 3.2 Hz, 4H), 6.50-6.55 (m, 2H), 8.18 (d, J_H = 2.4 Hz, 2H). MALDI-Tof mass; calcd for [M+H^+] ($C_{152}H_{222}N_9O_{87}$): 3565.32; found: 3564.31.

Lac4-DMAP (**2**):

To a solution of **24** (80 mg, 22.4 μmol) in 10 mL of methanol was added 10 μL of methanol solution of sodium methoxide (49 μmol). The reaction mixture was stirred for 1 h at room temperature and the solution was purified by reversed-phase HPLC (YMC-pack ODS-A, 250 x 10 mm). The fraction containing the target compound was lyophilized. Lac4-DMAP **2** was obtained as a white solid (12.0 mg, 3.36 μmol, 15%). The HPLC condition was as follows: CH_3CN/H_2O (both containing 0.1 % TFA) 5/95-20/80 (linear gradient over 30min), flow rate = 3mL /min, detected by UV (280 nm). HRMS (FAB); calcd for [M+H^+] ($C_{96}H_{166}N_9O_{59}$): 2389.0266; found: 2389.0342.

<u>Synthesis of NAG3-DMAP 3</u>



Compound (**25**):

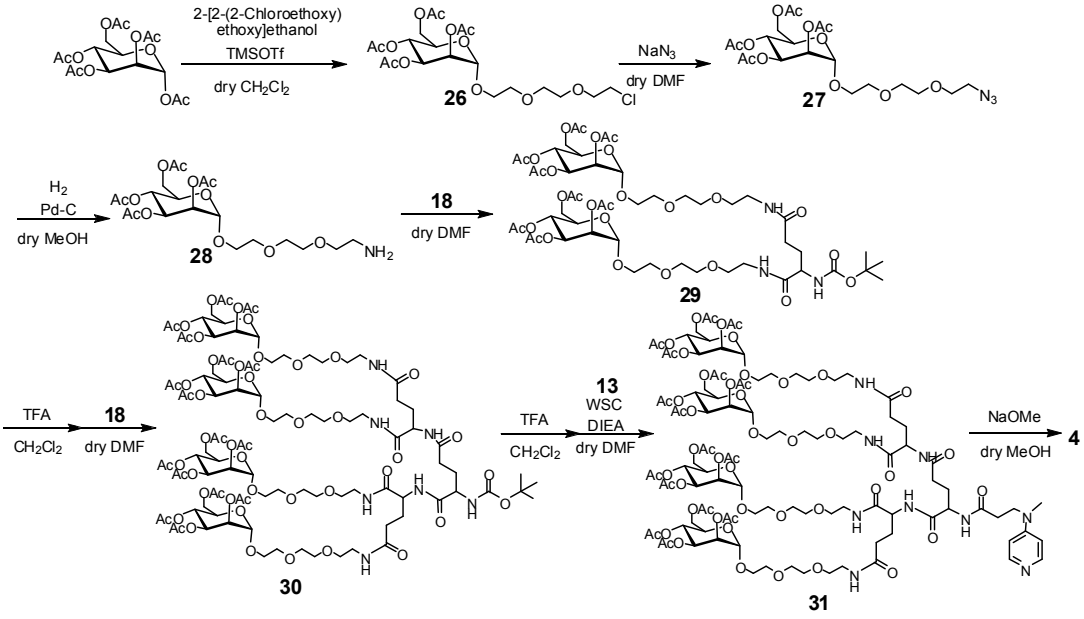
To a solution of 3-(Methyl-4-pyridylamino) propionic acid methyl ester (100 mg, 500 μmol) in 2 mL of methanol was added hydrazine monohydrate (124 μL, 2.50 mmol). The reaction mixture was stirred for 1 day at room temperature and evaporated under vacuum. The residue was precipitated by hexane to give **25** as a white solid (38 mg, 190 μmol, 38%). ^{1}H-NMR (400 MHz,

CD_3OD); δ/ppm: 2.42 (t, J_H = 7.2 Hz, 2H), 3.00 (s, 3H), 3.71 (t, J_H = 7.2 Hz, 2H), 6.68 (d, J_H = 6.8 Hz, 2H), 8.06 (t, J_H = 6.8 Hz, 2H).

NAG3-DMAP (**3**):

25 (3.00 mg, 15.0 μmol) and *N,N',N''*-triacetyl chitoriose (20.0 mg, 30.0 μmol) was dissolved in 1 mL of deionized water. A pH of the resulting solution was adjusted to 5.0, and the reaction mixture was stirred for 3 days at 70 °C and purified by reversed-phase HPLC (YMC-pack ODS-A, 250 x 10 mm). The desired fraction was lyophilized and dissolved in deionized water to obtain NAG3-DMAP **3** as a 200 μL aqueous solution (5.50 mM, 1.10 μmol, 7%). The HPLC condition was as follows: CH_3CN/10 mM aqueous ammonium acetate 2/98-10/90 (linear gradient over 30min), flow rate = 3mL /min, detected by UV (280 nm). HRMS (FAB); calcd for [M+H$^+$] ($C_{33}H_{54}N_7O_{16}$): 804.3622; found: 804.3649.

Synthesis of Man4-DMAP 4



Compound (**26**):

2-[2-(2-chloroethoxy)ethoxy]ethanol (1.13 mL, 7.68 mmol) and molecular sieves 4A (1.50 g) were added to a solution of peracetyl mannose (1.00 g, 2.56 mmol) in 5 mL of CH_2Cl_2. After stirring for 1 h at room temperature, trimethylsilyl trifluoromethanesulfonate (1.14 mL, 5.12 mmol) was added to above solution. The reaction mixture was stirred for 5 h at room temperature and diluted with 100 mL of ethyl acetate. The organic solution was washed with sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution with ethyl acetate/hexanes 2:1) to give **26** as a colorless oil (565 mg, 1.14 mmol, 46%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.98-2.14 (m, 12H), 3.62-3.68 (m, 10H), 3.76 (t, J_H = 5.6 Hz, 2H), 4.04-4.08 (m, 1H), 4.11 (d, J_H = 2.4 Hz, 1H), 4.28 (dd, J_H = 12.0, 2.4 Hz, 1H), 4.86 (d, J_H = 1.6 Hz, 1H), 5.26 (d, J_H = 3.6, 1.6 Hz, 1H), 5.29 (d, J_H = 10.0 Hz, 1H), 5.35 (dd, J_H = 10.0, 3.6 Hz, 1H).

Compound (**27**):

To a solution of **26** (560 mg, 1.12 mmol) in 10 mL of DMF was added sodium azide (198 mg, 3.05 mmol). The reaction mixture was stirred for 14 h at 70 °C, evaporated under vacuum and diluted with 150 mL of ethyl acetate. The organic layer was washed with deionized water and brine, and dried over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **27** as a colorless oil (449 mg, 890 μmol, 79%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.98-2.15 (m, 12H), 3.39 (t, J_H = 4.8 Hz, 2H), 3.66-3.70 (m, 10H), 4.04-4.08 (m, 1H), 4.11 (d, J_H = 2.4 Hz, 1H), 4.28 (dd, J_H = 12.0, 2.4 Hz, 1H), 4.87 (d, J_H = 2.0 Hz, 1H), 5.26 (d, J_H = 3.2, 2.0 Hz, 1H), 5.29 (d, J_H = 10.0 Hz, 1H), 5.36 (dd, J_H = 10.0, 3.2 Hz, 1H).

Compound (**28**):

To a solution of **27** (947 mg, 1.87 mmol) in 20 mL of methanol was added palladium-carbon (10 %, 50 mg). Under H_2 atmosphere, the reaction mixture was stirred for 3 h at room temperature and the solution was filtered to remove palladium-carbon. The filtrate was evaporated under vacuum to afford **28** as a yellow solid (830 mg, 1.64 mmol, 88%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 2.00-2.16 (m, 12H), 3.23-3.26 (m, 2H), 3.64-3.75 (m, 10H), 4.01-4.03 (m, 1H), 4.13 (d, J_H = 12.4, 2.4 Hz, 1H), 4.27 (dd, J_H = 12.0, 4.8 Hz, 1H), 4.91 (d, J_H = 0.8 Hz, 1H), 5.26-5.31 (m, 3H).

Compound (**29**):

To a solution of **28** (50.0 mg, 104 μmol) in 2mL of CH_2Cl_2 containing DIEA (15.0 μL, 104 μmol) was added **18** (17.3 mg, 41.6 μmol). The reaction mixture was stirred for 3 h at room temperature, and the solution was diluted with 50 mL of CH_2Cl_2. The organic layer was washed with 5% aqueous citric acid, sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The solution was filtered and evaporated under vacuum to afford **29** as a white solid (39.0 mg, 33.3 μmol, 80% for **18**). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.41 (s, 9H), 1.98-2.14 (m, 24H), 2.24-2.34 (m, 2H), 3.36-3.47 (m, 4H), 3.54-3.65 (m, 24H), 4.03-4.11 (m, 4H), 4.26 (dd, J_H = 12.0, 4.8 Hz, 2H), 4.86 (s, 2H), 5.25-5.34 (m, 6H), 5.65 (br, 1H), 6.64 (br, 1H), 7.08 (br, 1H). MALDI-Tof mass; calcd for $[M+Na^+]$ ($C_{50}H_{79}N_3O_{28}Na$): 1192.42; found: 1195.72.

Compound (**30**):

29 (232 mg, 198 μmol) was dissolved in 5 mL of CH_2Cl_2 containing 20% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature evaporated under vacuum. The residue was dissolved in 10 mL of CH_2Cl_2 containing DIEA (50.0 μL, 297 μmol) and **18** (30.6 mg, 69.3 μmol). The reaction mixture was stirred for 10 h at room temperature and diluted with 350 mL of CH_2Cl_2. The organic layer was washed with 5% aqueous citric acid, sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The solution was filtered and evaporated under vacuum to afford **30** as a white solid (162 mg, 69.3 μmol, 100% for **18**). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.40 (s, 9H), 1.96-2.13 (m, 48H), 2.29 (br, 4H), 2.56 (br, 2H), 3.35-3.48 (m, 10H), 3.54-3.64 (m, 48H), 4.03-4.09 (m,

8H), 4.26 (dd, J_H = 12.4, 4.0 Hz, 4H), 4.85 (s, 4H), 5.24-5.33 (m, 12H), 5.62 (br, 1H), 6.60-6.81 (br, 2H), 7.71 (br, 1H), 7.88 (br, 1H), 8.58 (br, 1H), 8.78 (br, 1H). MALDI-Tof mass; calcd for [M+Na$^+$] ($C_{100}H_{155}N_7O_{56}$): 2372.95; found: 2376.16.

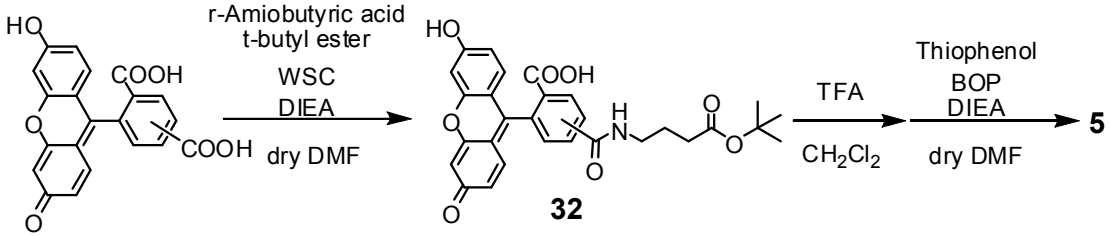
Compound (**31**):

30 (160 mg, 69.3 μmol) was dissolved in 4 mL of CH_2Cl_2 containing 50% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature and evaporated under vacuum. The oily residue was dissolved in 2 mL of DMF containing WSC (38.0 mg, 208 μmol) and DIEA (114 μL, 693 μmol), and a solution of **13** (53.0 mg, 308 μmol) in 1 mL of deionized water was added to this solution. The reaction mixture was stirred for 1h at 60 °C and diluted 25 mL of 1% aqueous acetic acid. The aqueous solution was extracted with ethyl acetate, and the organic layer was washed with sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The solution was filtered and evaporated under vacuum to afford **31** as a white solid (110 mg, 45.5 μmol, 66%). ^{1}H NMR (400 MHz, $CDCl_3$) δ/ppm: 1.95-2.12 (m, 84H), 2.28-2.34 (m, 6H), 2.92 (s, 3H), 3.34-3.46 (m, 8H), 3.51-3.67 (m, 48H), 3.76-3.79 (m, 2H), 4.01-4.08 (m, 6H), 4.22-4.26 (m, 6H), 4.84 (s, 4H), 5.22-5.31 (m, 12H), 6.50-6.54 (m, 2H), 8.14 (d, J_H = 4.4 Hz, 2H). MALDI-Tof mass; calcd for [M+Na$^+$] ($C_{104}H_{158}N_9O_{55}$): 2412.98; found: 2413.71.

Man4-DMAP (**4**):

To a solution of **31** (110 mg, 45.5 μmol) in 10 mL of methanol was added 50 μL of methanol solution of sodium methoxide (245 μmol). The reaction mixture was stirred for 1 h at room temperature and purified by reversed-phase HPLC (YMC-pack ODS-A, 250 x 10 mm). The fraction containing the desired compound was lyophilized. Man4-DMAP **4** was obtained as a white solid (12.0 mg, 3.36 μmol, 15%). The HPLC condition was as follows: CH_3CN/H_2O (both containing 0.1 % TFA) 5/95-20/80 (linear gradient over 30min), flow rate = 3mL /min, detected by UV (280 nm). HRMS (FAB); calcd for [M+H$^+$] ($C_{72}H_{126}N_9O_{39}$): 1740.8147; found:1740.8132.

<u>Synthesis of Acyl donor 5</u>



Compound (**32**):

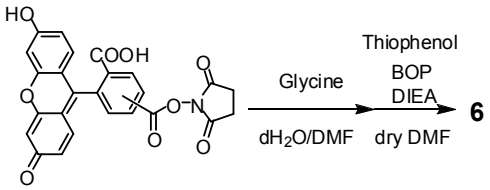
5,6-Carboxyfluorescein was synthesized as previously reported.[14] To a solution of 5,6-carboxyfluorescein (100 mg, 266 μmol) in 5 mL of DMF were added WSC (78.6 mg, 399 μmol), DIEA (153 μL, 931 μmol), and γ-aminobutyric acid *t*-butyl ester (104 mg, 532 μmol). The reaction mixture was stirred for 3 h at room temperature, evaporated under vacuum, and diluted with 200 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid and brine and dried

over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexanes 3:1) to give 3-carboxypropylaminocarbonyl fluorescein *t*-butyl ester **32** as a yellow solid (72 mg, 139 μmol, 59%). 3-carboxypropylaminocarbonyl fluorescein *t*-butyl ester **32** was obtained as the mixture of 5' and 6' isomer. ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 1.37 (s, 9H(6')), 1.44 (s, 9H(5')), 1.79(quintet, J_H = 7.2 Hz, 2H(6')), 1.91 (quintet, J_H = 7.2 Hz, 2H(5')), 2.25 (t, J_H = 7.2 Hz, 2H(6')), 2.35 (t, J_H = 7.2 Hz, 2H(5')), 3.33 (t, J_H = 7.2 Hz, 2H(6')), 3.46 (t, J_H = 7.2 Hz, 2H(5')), 6.51-6.61 (m, 4H(5', 6')), 6.68-6.69 (m, 2H(5', 6')), 7.29 (d, J_H = 8.0 Hz, 1H(5')), 7.60 (d, J_H = 2.0 Hz, 1H(6')), 8.07 (d, J_H = 8.0 Hz, 1H(6')), 8.12 (d, J_H = 8.0 Hz, 1H(6')), 8.18 (d, J_H = 8.0, 1H(5')), 8.42 (s, 1H(5')).

Acyl Donor (**5**):

3-Carboxypropylaminocarbonyl fluorescein *t*-butyl ester **32** was dissolved in 2 mL of CH_2Cl_2 containing 40 % trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature and evaporated under vacuum to afford 3-carboxypropylaminocarbonyl fluorescein as a yellow solid. To a solution of fluorescein aminobutyric acid in 2 mL of DMF were added BOP (20.5 mg, 46.4 μmol), DIEA (6.3 μL, 46.4 μmol), and thiophenol (7.9 μL, 77.4 μmol). The reaction mixture was stirred for 1h at room temperature, evaporated under vacuum, and diluted with 200 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid and brine and dried over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexanes 3:1) to give acyl donor **5** as a yellow solid (17.3 mg, 31.3 μmol, 81%). Acyl donor **5** was obtained as the mixture of 5' and 6' isomer. ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 1.91(quintet, J_H = 7.2 Hz, 2H(6')), 2.06 (quintet, J_H = 7.2 Hz, 2H(5')), 2.69 (t, J_H = 7.2 Hz, 2H(6')), 2.81 (t, J_H = 7.2 Hz, 2H(5')), 3.36 (t, J_H = 7.2 Hz, 2H(6')), 3.49 (t, J_H = 7.2 Hz, 2H(5')), 6.50-6.70 (m, 4H(5', 6')), 6.68-6.69 (m, 2H(5', 6')), 7.26-7.40 (m, 5H(5', 6')+1H(5')), 7.61 (d, J_H = 2.0 Hz, 1H(6')), 8.06 (d, J_H = 8.0 Hz, 1H(6')), 8.12 (d, J_H = 8.0 Hz, 1H(6')), 8.18 (J_H = 8.0 Hz, 1H(5')), 8.41 (s, 1H(5')). HRMS (FAB) calcd for $[M+H^+]$ ($C_{31}H_{24}NO_7S$): 554.1268; found:554.1285.

<u>Synthesis of acyl donors 6</u>

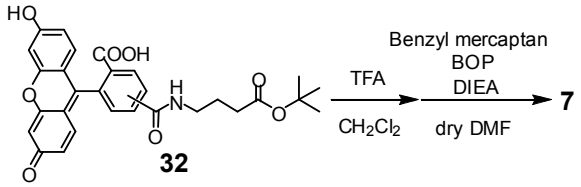


Acyl donor (**6**):

To a solution of fluorescein carboxylic acid succinimide ester (Molecular Probes, 20.0 mg, 42.2 μmol) in 1 mL of DMF containing DIEA (20 μL, 211 μmol) was added a solution of glycine (15.0 mg, 211 μmol) in 5 mL of deionized water. The reaction mixture was stirred for 3h at room temperature and diluted with 25 mL of 5% aqueous citric acid. The aqueous solution was extracted with ethyl acetate, and the organic layer was washed with brine, and dried over anhydrous $MgSO_4$. The solution was filtered and evaporated under vacuum to afford a yellow oil.

To a solution of the yellow in 1 mL of DMF containing BOP (37.0 mg, 84.4 μmol) and DIEA (30.0 μL, 253 μmol) was added thiophenol (23.2 μL, 211 μmol). The reaction mixture was stirred for 3h at room temperature, evaporated under vacuum and diluted with 50 mL of ethyl acetate. The organic layer was washed with brine, and dried over anhydrous Na_2SO_4. The crude residue was purified by flash column chromatography on silica gel (elution with ethyl acetate/hexanes 1:1) to give acyl donor **6** as a yellow solid (10.1 mg, 18.6 μmol, 44%). Acyl donor **6** was obtained as the mixture of 5' and 6' isomer. HRMS (FAB); calcd for $[M+H^+]$ ($C_{29}H_{20}NO_7S$): 526.0955; found: 526.0959. ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 4.29 (s, 2H(6')), 4.43 (s, 2H(5')), 6.53-6.64 (m, 4H(5', 6')), 6.69-6.70 (m, 2H(5', 6')), 7.34-7.44 (m, 5H(5', 6')+1H(5')), 7.54 (d, J_H = 1.6 Hz, 1H(6')), 8.12 (d, J_H = 8.4 Hz, 1H(6')), 8.21 (d, J_H = 8.4 Hz, 1H(6')), 8.28 (d, J_H = 8.0 Hz, 1H(5')), 8.51 (s, 1H(5')).

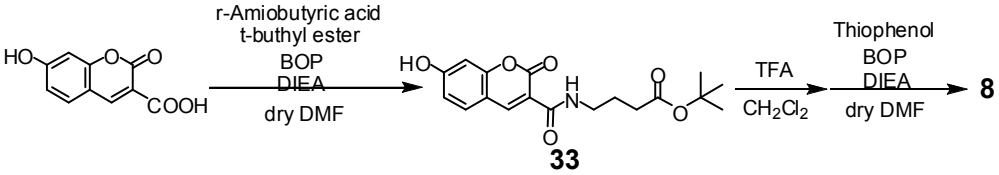
<u>Synthesis of acyl donors 7</u>



Acyl donor (**7**);

3-Carboxypropylaminocarbonyl fluorescein *t*-butyl ester **32** (38.0 mg, 73.4 μmol) was dissolved in 2 mL of CH_2Cl_2 containing 50% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature and evaporated under vacuum to afford 3-carboxypropylaminocarbonyl fluorescein as a yellow solid. To a solution of 3-carboxypropylaminocarbonyl fluorescein in 2 mL of DMF containing BOP (38.9 mg, 88.1 μmol) and DIEA (24.2 μL, 147 μmol) was added benzyl mercaptan (18.2 mg, 147 μmol). The reaction mixture was stirred for 1h at room temperature, evaporated under vacuum and diluted with 100 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid and brine, and dried over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution with ethyl acetate/hexanes 3:1) to give acyl donor **7** as a yellow solid (8.70 mg, 14.8 μmol, 20%). Acyl donor **7** was obtained as the mixture of 5' and 6' isomer. HRMS (FAB) calcd for $[M+H^+]$ ($C_{32}H_{26}NO_7S$): 568.1424; found: 568.1419. ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 1.89(quintet, J_H = 7.2 Hz, 2H(4')), 2.04 (quintet, J_H = 7.2 Hz, 2H(5')), 2.61 (t, J_H = 7.2 Hz, 2H(6')), 2.81 (t, J_H = 7.2 Hz, 2H(5')), 3.33 (t, J_H = 7.2 Hz, 2H(6')), 3.46 (t, J_H = 7.2 Hz, 2H(5')), 3.99 (s, 2H(6')), 4.12 (s, 2H(5')), 6.50-6.70 (m, 4H(5',6')), 6.67-6.68 (m, 2H(5',6')), 7.18-7.30 (m, 5H(5',6')+1H(5')), 7.59 (d, J_H = 2.0 Hz, 1H(6'), Ar-H), 8.10 (d, J_H = 8.0 Hz, 1H(6'), Ar-H), 8.11 (d, J_H = 8.0 Hz, 1H(6'), Ar-H), 8.18 (d, J_H = 8.0 Hz, 1H(5'), Ar-H), 8.41 (s, 1H(5'), Ar-H).

Synthesis of acyl donors 8



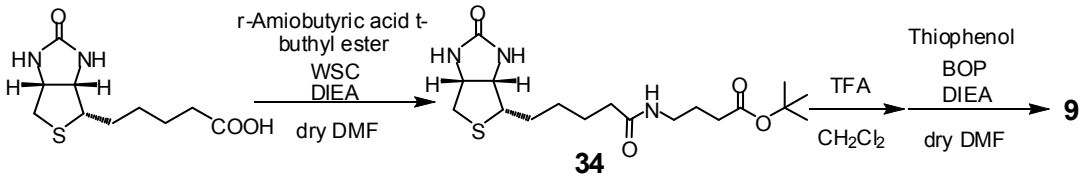
Copmound (**33**):

7-Hydroxycoumarin 3-carboxylic acid was synthesized as previously reported.[15] To a solution of 7-hydroxycoumarin 3-carboxylic acid (72.0 mg, 349 μmol) in 5 mL of DMF containing BOP (185 mg, 419 μmol) and DIEA (92.4 μL, 524 μmol) was added γ-aminobutyric acid *t*-butyl ester (102 mg, 524 μmol). The reaction mixture was stirred for 1 h at room temperature, evaporated under vacuum and diluted with 250 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid and brine, and dried over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution with ethyl acetate/hexanes 2:1) to give **33** as a yellow solid (70.0 mg, 202 μmol, 58%). ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 1.45 (s, 9H), 1.87 (quintet, J_H = 7.2 Hz, 2H), 2.32 (t, J_H = 7.2 Hz, 2H), 3.31-3.47 (m, 2H), 6.76 (d, J_H = 2.0 Hz, 2H), 6.87(dd, J_H = 8.8 Hz, 2.0 Hz, 1H), 7.66 (d, J_H = 8.8 Hz, 1H), 8.74 (s, 1H), 9.04 (br, 1H).

Acyl donor (**8**):

33 (65.0 mg, 187 μmol) was dissolved in 4 mL of CH_2Cl_2 containing 50% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature, and evaporated under vacuum to afford a white solid. To a solution of the white solid in 5 mL of DMF containing BOP (99.3 mg, 225 μmol) and DIEA (66.0 μL, 374 μmol) was added thiophenol (41.2 μL, 374 μmol). The reaction mixture was stirred for 1h at room temperature, the evaporated under vacuum and diluted with 200 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid and brine, and dried over anhydrous $MgSO_4$. The cured residue was purified by flash column chromatography on silica gel (elution with ethyl acetate/hexanes 2:1) to give acyl donor **8** as a white solid (35.0 mg, 91.4 μmol, 49%). HRMS (FAB); calcd for $[M+H^+]$ ($C_{20}H_{18}NO_5S$): 384.0900; found: 384.0921. ^{1}H-NMR (400 MHz, CD_3OD); δ/ppm: 1.99 (quintet, J_H = 7.2 Hz, 2H), 2.78 (t, J_H = 7.2 Hz, 2H), 3.47-3.49 (m, 2H), 6.77 (d, J_H = 2.0 Hz, 2H), 6.88(dd, J_H = 8.8 Hz, 2.0 Hz, 1H), 7.38(s, 5H), 7.67(d, J_H = 8.8 Hz, 1H), 8.76 (s, 1H).

Synthesis of acyl donors 9



Compound (**34**):

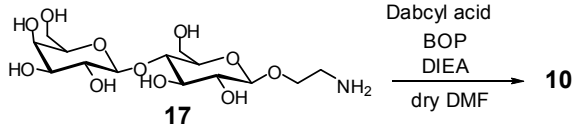
To a solution of (+)-biotin (200 mg, 818 μmol) in 2 mL of DMF containing WSC (188 mg, 982 μmol) and DIEA (323 μL, 1.96 mmol) was added γ-aminobutyric acid *t*-butyl ester (192 mg, 982

μmol). The reaction mixture was stirred overnight at room temperature and evaporated under vacuum. The crude residue was purification by flash column chromatography on silica gel (elution with $CHCl_3$/methanol 6:1) to give **34** as a white solid (170 mg, 441 μmol, 54%). ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.44 (s, 9H), 1.64-1.82 (m, 8H), 2.19 (t, J_H = 7.2 Hz, 2H),2.27 (t, J_H = 7.2 Hz, 2H), 2.73 (d, J_H = 12.8 Hz, 1H), 2.91 (dd, J_H = 12.8, 7.2 Hz, 1H), 3.13-3.15 (m, 1H), 3.29 (q, J_H = 6.0 Hz, 2H), 4.31 (dd, J_H = 6.8, 4.8 Hz), 4.50-4.53 (m, 1H), 5.62 (s, 1H), 6.39 (br, 1H), 6.53 (s, 1H).

Acyl donor (**9**):

34 (160 mg, 73.4 μmol) was dissolved in 4 mL of CH_2Cl_2 containing 50% trifluoroacetic acid. The reaction mixture was stirred for 1h at room temperature and evaporated under vacuum to afford a white solid. To a solution of the white solid in 5 mL of DMF containing BOP (220 mg, 498 μmol) and DIEA (82μL, 498 μmol) was added thiophenol (91.0 μL, 830 μmol). The reaction mixture was stirred for 1h at room temperature, evaporated under vacuum and diluted in 200 mL of $CHCl_3$. The organic layer was washed with 5% aqueous citric acid, sat. aqueous $NaHCO_3$, and brine, and dried over anhydrous $MgSO_4$. The cured residue was purified by flash column chromatography on silica gel (elution with ethyl $CHCl_3$/methanol 10:1) to give acyl donor **9** as a white solid (117 mg, 277 μmol, 67%). HRMS (FAB); calcd for [M+H$^+$] ($C_{20}H_{28}N_3O_3S_2$): 421.1572; found: 421.1585. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.64-1.75 (m, 6H), 1.91 (quintet, J_H = 7.2 Hz, 2H), 2.17-2.21 (m, 4H), 2.68 (d, J_H = 12.8 Hz, 1H), 2.87 (dd, J_H = 12.8, 4.8 Hz, 1H), 3.11-3.23 (m, 1H), 3.29 (q, J_H = 6.0 Hz, 2H), 4.28-4.31 (m, 1H), 4.45-4.49 (m, 1H), 5.43 (s, 1H), 6.36 (s, 1H), 6.45 (s, 1H), 7.38(s, 5H).

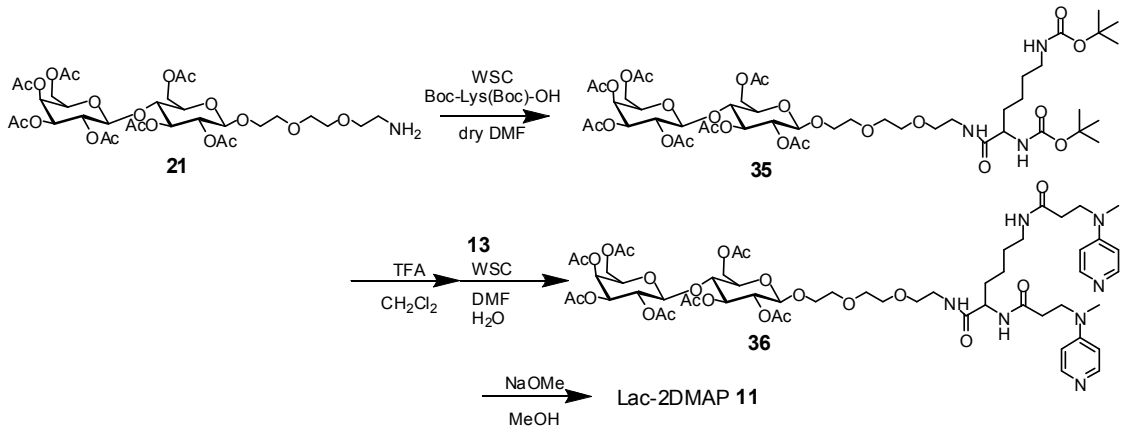
<u>Synthesis of Dab-lac 10</u>



Dab-lac (**10**):

To a solution of **17** (34.3 mg, 89.0 μmol) in 2 mL of DMF containing BOP (39.4 mg, 89.0 μmol) and DIEA (40.0 μL, 226 μmol) was added dabcyl acid (20.0 mg, 74.2 μmol). The reaction mixture was stirred for 2 h at room temperature and evaporated under vacuum, and diluted with deionized water. The aqueous solution was purified by HPLC (YMC-pack ODS-A, 250 x 10 mm). The desired fraction was lyophilized to afford Dab-lac **10** as a red solid (10.1 mg, 15.9 μmol, 18%). The HPLC condition was as follows: CH_3CN/H_2O (containing 10 mM ammonium acetate) 30/70-40/60 (linear gradient over 30min), flow rate = 3mL /min, detected by UV (510 nm). HRMS (FAB); calcd for [M+H$^+$] ($C_{29}H_{41}N_4O_{12}$): 636.2643; found:636.2663.

Synthesis of Lac-3DMAP 11



Compound (**35**):

To a solution of 2-2-1 135 mg (176 μmol) in 10 mL of DMF were added WSC 33.7 mg (176 μmol), HOBt 25.0 mg (176 μmol), Boc-Lys(Boc)-OH 74.3 mg (141 μmol) and DIEA 59 μL (362 μM). The reaction mixtures were stirred 2 h at room temperature, evaporated under vacuum, and diluted with 100 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid, aqueous $NaHCO_3$ and brine and dry over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexane 20:1) to give compound **35** 93 mg (84 μmol, 48 %) as a colorless oil. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.23-1.61 (m, 24H), 1.97-2.42 (m, 21H), 3.10 (d, J_H = 6.4 Hz, 2H), 3.44-3.48 (m, 2H), 3.52-3.64 (m, 10H), 3.79-3.95 (m, 2H), 4.06-4.16 (m, 4H), 4.48-4.51 (m, 2H), 4.56 (d, J_H = 8.0 Hz, 1H), 4.60 (s, 1H), 4.90 (t, J_H = 8.0 Hz, 1H), 4.96 (dd, J_H = 8.8, 3.2 Hz, 1H), 5.10 (t, J_H = 8.0 Hz, 1H), 5.20 (t, J_H = 8.8 Hz, 1H), 5.35 (d, J_H = 3.2 Hz, 1H)

Compound (**36**):

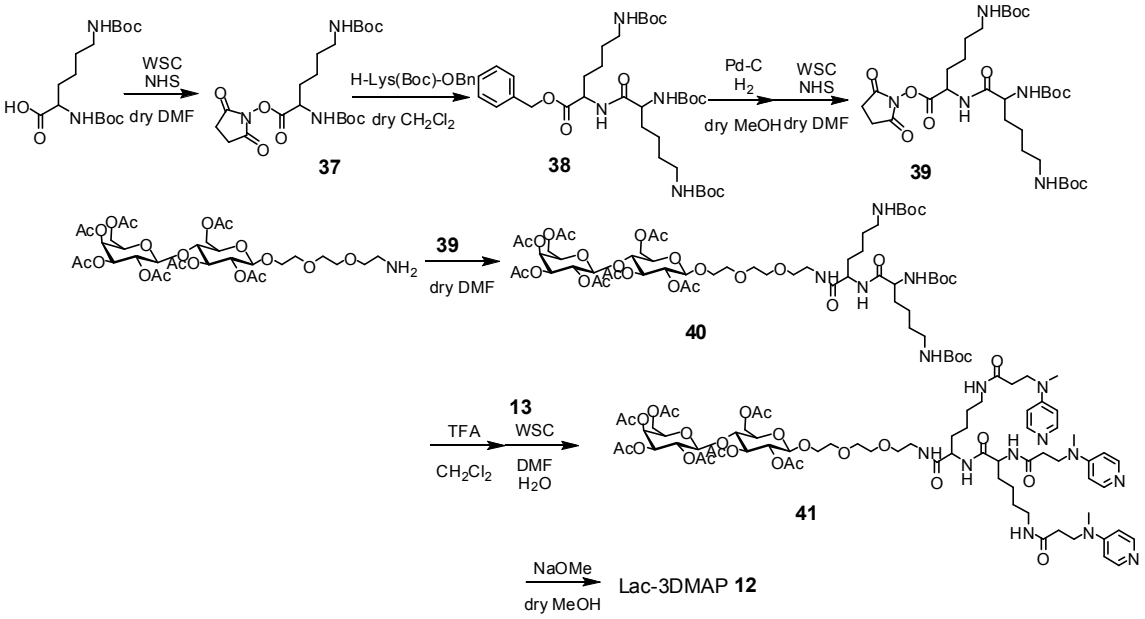
Compound **35** 93 mg (84 μmol) was dissolved in 3 mL of CH_2Cl_2 containing 50 % trifluoroacetic acid. The reaction mixtures were stirred for 1 h at room temperature and evaporated under vacuum. The oily residue was dissolved in 3 mL of DMF containing WSC 48.7 mg (252 μmol) and DIEA 41 μL (252 μmol), and a solution of 3-(methyl-4-pyridylamino) propanic acid 45 mg (252 μmol) in 300 μL of deionized water was added to this solution. The reaction mixtures were stirred 1 h at room temperature, and then evaporated under vacuum. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexane 2:1) to give compound **36** 57 mg (52 μmol, 62 %) as a colorless oil. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.72-1.76 (m, 24H), 1.97-2.14 (m, 21H), 2.59 (t, J_H = 6.8 Hz, 4H), 3.95 (s, 6H), 3.15-3.26 (m, 6H), 3.38-3.41 (m, 2H), 3.52-3.92 (m, 12H), 4.05-4.14 (m, 4H), 4.26-4.28 (m, 1H), 4.48-4.55 (m, 3H), 4.88 (t, J_H = 8.0 Hz, 1H), 4.97 (dd, J_H = 8.8, 3.2 Hz, 1H), 5.10 (t, J_H = 8.0, 1H), 5.19 (t, J_H = 8.8, 1H), 5.35 (d, J_H = 3.2, 1H), 6.67 (t, J_H = 8.0, 4H), 8.10 (t, J_H = 8.0, 4H)

Lac-2DMAP (**11**):

To a solution Compound **16** 25 mg (20.5 μmol) in 3 mL of methanol were added 30 μL of

methanol solution of sodium methoxide. The reaction mixtures were stirred 1 h at room temperature, 1N aqueous hydrogen chloride wad added. The resultant solution was evaporated under vacuum to get Lac-2DMAP **11** (13.1 μmol, 59 %). HRMS (FAB); calcd for [M+H+] ($C_{42}H_{67}N_7O_{16}$): 926.4723; found: 925.4673.

<u>Synthesis of Lac-3DMAP 12</u>



Compound (**37**):

To a solution of Boc-Lys(Boc)-OH 500 mg (947 μmol) in 5 mL of DMF were added WSC 217 mg (1.14 mmol), *N*-hydroxy succinimide 131 mg (1.14 mmol) and DIEA 186 μL (1.14 mmol). The reaction mixtures were stirred over night, evaporated under vacuum, and diluted with 150 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid, aqueous $NaHCO_3$ and brine and dry over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **37** (941 μmol, 99 %) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.27-1.98 (m, 24H), 2.83 (s, 4H), 3.13 (d, J_H = 5.6, 2H), 4.50 (br, 2H), 5.13 (br, 1H)

Compound (**38**):

Compound **37** 202 mg (456 μmol) and H-Lys-OBn 200mg (537 μmol) was dissolved in 5 mL of dichloromethane containing 87.6 μL of DIEA (537 μmol). To the reaction mixtures were added 100 mL of chloroform, and the organic layer was as washed with 5% aqueous citric acid, aqueous $NaHCO_3$ and brine and dry over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **38** 291 mg (438 μmol, 91 %) as colorless oil. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.27-1.86 (m, 39H), 3.04-3.11 (m, 4H), 4.06 (br, 1H), 4.59-4.64 (m, 3H), 5.15 (q, J_H = 1.6 Hz, 2H), 6.62 (br, 1H), 7.26 (s, 5H).

Compound (**39**):

To a solution of **38** 291 mg (438 μmol) in 5 mL of methanol were added 10 % palladium-carbon 15

mg. The reaction mixtures were stirred for 1 h at room temperature under hydrogen atmosphere, filtered and evaporated under vacuum. The residue was dissolved in 10 mL of DMF containing WSC 95 mg (497 μmol), *N*-hydroxy succinimide 57 mg (497 μmol) and 81.1 μL of DIEA (497 μmol). The reaction mixtures were stirred 2 h at room temperature, evaporated under vacuum, and diluted with 100 mL of ethyl acetate. The organic layer was washed with 5% aqueous citric acid, aqueous $NaHCO_3$ and brine and dry over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **39** 280 mg (417 μmol, 95 %) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.26-1.92 (m, 39H), 2.84 (s, 4H), 3.11 (br, 4H), 4.10 (br, 1H), 4.67-4.94 (m, 3H), 6.89 (s, 1H).

Compound **(40)**:

To a solution of 2-aminoethyl lactose 103 mg (134 μmol) in 4 mL of dichloromethane were added **39** 81 mg (121 μmol). The reaction mixture were stirred for 2 h at room temperature, and 50 mL of dichloromethane were added. The organic layer was washed with 5% aqueous citric acid, aqueous $NaHCO_3$ and brine and dry over anhydrous $MgSO_4$. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexane 2:1) to give compound **40** 47 mg (35.5 μmol, 29 %) as a colorless oil. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.25-1.47 (m, 39H), 1.95-2.13 (m, 21H), 3.08 (br, 4H), 3.47-3.51 (m, 2H), 3.58-3.67 (m, 10H), 3.78-3.91(m, 2H), 3.93-4.13 (m, 4H), 4.38-4.40 (m, 1H), 4.49-4.55 (m, 3H), 4.78 (s, 1H), 4.86-4.97 (m, 2H), 5.09 (t, J_H = 8.0 Hz, 1H), 5.19 (t, J_H = 9.2 Hz, 1H), 5.33 (d, J_H = 3.2 Hz, 1H)

Compound **(41)**:

Compound **40** 47 mg (35.5 μmol) was dissolved in 4 mL of CH_2Cl_2 containing 50 % trifluoroacetic acid. The reaction mixtures were stirred for 1 h at room temperature and evaporated under vacuum. The oily residue was dissolved in 3 mL of DMF containing WSC 34 mg (178 μmol) and DIEA 29 μL (178 μmol), and a solution of **13** 32 mg (178 μmol) in 300 μL of deionized water was added to this solution. The reaction mixtures were stirred 2 h at room temperature, and then evaporated under vacuum. The crude residue was purified by gel filtration chromatography (sehadex LH-20, MeOH) to give compound **41** 47 mg (31.1 μmol, 88 %) as a white solid. ^{1}H-NMR (400 MHz, $CDCl_3$); δ/ppm: 1.29-1.75 (m, 12H), 1.93-2.13 (m, 21H), 2.55-2.66 (m, 6H), 3.13 (t, J_H = 6.8 Hz, 4H), 3.32 (s, 9H), 3.52-3.55 (m, 2H), 3.60-3.77 (m, 10H), 3.83-3.91(m, 2H), 3.93-4.13 (m, 4H), 4.22-4.25 (m, 2H), 4.50-4.72 (m, 3H), 4.86-5.03 (m, 2H), 5.12 (dd, J_H = 9.9, 2.4 Hz, 1H), 5.18 (t, J_H = 8.8 Hz, 1H), 5.36 (d, J_H = 3.6 Hz, 1H), 7.03 (d, J_H = 7.2 Hz, 6H), 8.13 (d, J_H = 7.2Hz, 6H)

Lac-3DMAP **(12)**:

To a solution of compound **41** 12 mg (7.96 μmol) in 3 mL of methanol were added 4 μL of methanol solution of sodium methoxide (15.9 μmol l). The reaction mixtures were stirred 1 h at room temperature and the solution was purified by reversed-phase HPLC (YMC-pack ODS-A). The fraction containing the target compound was lyophilized. Lac-3DMAP **12** (1.32 μmol, 17 %) was obtained. HRMS (FAB); calcd for $[M+H^+]$ ($C_{57}H_{89}N_{11}O_{18}$): 1216.6465; found: 1216.6466.

General Comments for Biochemical Assays.

Recombinant congerin II (Cong II) was expressed in *E. coli* strain JM109/pTV-Cong II and purified as previously described.[11] Concanavalin A (Con A) and Wheat Germ Agglutinin (WGA) were purchased from Funakoshi and used without further purification. The concentration of lectins (Cong II; ε_{280nm} = 11,500, Con A; ε_{280nm} = 35,000, and WGA; ε_{280nm} = 25,500 / $M^{-1}cm^{-1}$), sugar-tethered DMAPs 1-4 (ε_{280nm} = 18,000), and acyl donors 5-8 (5-7; ε_{494nm} = 75,000 (pH 9), 8; ε_{419nm} = 36,000 (in MeCN)) [16] was determined spectrophotometrically.

Chemical Acylation of lectins in test tubes.

Acyl transfer reactions were carried out under the following conditions: 10 μM lectin, 50 μM acyl donor, and 50 μM sugar-tethered DMAP in 50 mM HEPES buffer (pH 8.0). The reaction was initiated by the addition of acyl donor, and incubated at 25 °C. For MALDI-Tof mass measurements shown in Fig. 12a, aliquots were taken at every hour, immediately purified by ZipTip (Millipore), and mixed with a matrix solution (sinapinic acid). After 3 h incubation, the reaction solution was subjected to gel filtration (TOYOPEARL HW-40F) using 50 mM acetate buffer (pH 5.0), and the eluent was dialyzed against 50 mM HEPES buffer (pH 7.5). Modification yields shown in Table 1 were calculated by the absorbance at 464 nm to that at 280 nm using the molar extinction coefficients of acyl donors **5-7** (ε_{464nm} = 26,000, ε_{280nm} = 16,000 (pH 7.5)) and lectins.

Inhibition of acyl transfer reaction in the presence of excess saccharide ligand.

Acyl transfer reactions were carried out under the following conditions: 10 μM Cong II, 50 μM acyl donor **5**, and 50 μM Lac-DMAP **1** in 50 mM HEPES buffer (pH 8.0) containing 50 mM lactose. After 3h incubation at room temperature, the reaction solutions were purified and analyzed as described in the main manuscript. The modification yield in the presence of excess lactose was calculated to be less than 1 %.

Investigation of hydrolysis of acyl donors 5-7.

A solution of 50 μM acyl donors **5-7** in 50 mM HEPES buffer (pH 8.0) was incubated at 25 °C in the presence or absence of 50 μM Lac-DMAP **1**. After incubation for 3h, the reaction solution was analysed by reversed-phase HPLC using an ODS-A column (YMC, 250 × 4.6 mm). The HPLC condition used was as follows: CH_3CN / H_2O (both containing 0.1 % TFA) 5 / 95 – 95 / 5 (linear gradient over 45 min), flow rate = 1 mL /min, detection by UV absorption (464 nm). Hydrolysis yields were calculated by determining the ratio of peak areas of acyl donors to that of hydrolysis products.

Determination of modification sites.

The coumarin modified Cong II prepared as described above was diluted in 50 mM HEPES buffer (pH 8.0) containing 2 M urea. Trypsin was added to the solution to be a trypsin/substrate ratio of 1/10 (w/w), and the digestion was allowed to proceed for 3h at 37 °C. The digested peptides

were separated by reversed-phase HPLC using an ODS-A column (YMC, 250 × 4.6 mm), and each fraction was analyzed by MALDI-TOF mass spectrometry. The HPLC condition used was as follows: CH_3CN / H_2O (both containing 0.1 % TFA) 5 / 95 – 50 / 50 (linear gradient over 45 min), flow rate = 1 mL /min, detection by UV absorption (220 nm) and fluorescence (λ_{ex} = 360 nm, λ_{em} = 410 nm). Coumarin modified fragment **4** was characterized by MALDI-Tof mass analysis; calcd for $[M+H^+]$ ($C_{147}H_{212}N_{39}O_{47}$): 3275.54; found: 3275.57. Other peaks observed by HPLC were also assigned to each fragment of Cong II. Coumarin modified fragment **4** was collected, lyophilized, and subjected to mass-mass analysis (Shimadzu, AXIMA QIT (MALDI-QIT-Tof)) to determine the modified amino acid.

BFQR Measurement.

Fluorescence spectra were measured by a Perkin Elmer LS55 fluorescence spectrometer. The slit widths for excitation and emission were set at 15 and 10 nm, respectively. The excitation wavelength was 488 nm. In fluorescence quenching processes, FL-Cong II solution prepared as described above was diluted to 0.2 μM with 50 mM HEPES buffer (pH 7.5) and Dab-lac **10** was added to this solution. In fluorescence recovery processes, to a solution of 0.2 μM FL-Cong II with 4 μM Dab-lac **10** was added various saccharide solutions. The structure of saccharides used for this BFQR study was shown in the Supporting Information.

Selective Acylation of lectins in Lectin mixture.

Acyl transfer reactions in a lectin mixture were carried out in 50 mM HEPES buffer (pH 8.0) containing 0.1 mg/mL of each lectin (Cong II, Con A, and WGA), 50 μM acyl donor 8, and 50 μM sugar-tethered DMAP 2-4 as described above. After 3 h incubation at 25 °C, the reaction mixture was fractionated by 15 % SDS-PAGE in a reducing condition. The gel was analyzed by a fluorescent gel imager ChemiDoc XRS (BioRad), and stained by Coomasie brilliant blue.

Selective Acylation of Cong II in *E. coli* lysates.

E. coli cells expressing Cong II (JM109/pTV-Cong II) were cultured in 2 × YT medium at 37 °C for 20 h, harvested, and lysed by sonication using a Sonifier 450 (Branson). After removing the insoluble materials by centrifugation, the soluble fraction was dialyzed agains 50 mM HEPES buffer (pH 8.0) containing 0.1 M NaCl and incubated with 1 mM *N*-maleoyl-β-alanine to mask free cysteine residues at 25 °C overnight. To the lysate solution were added acyl donor 6 (50 μM) and sugar-tethered DMAP 1 or 2 (50 μM). After 3h incubation at 25 °C, the reaction mixtures were fractionated by 15 % SDS-PAGE in a reducing condition and analyzed as described above.

Congerin Acylation in skin mucous extracts.

Skin mucous extracts of conger myriaster were prepared according to the procedure reported. The extracts were dialyzed against 50 mM HEPES buffer (pH 8.0) containing 0.1 M NaCl, and incubated with 1 mM *N*-maleoyl-β-alanine and 1 mM 4-(2-aminoethyl)benzenesulfonyl fluoride

(protease inhibitor) at 25 °C overnight. To the extract were added acyl donor 7 (50 μM) and each sugar-tethered DMAP 2-4 (50 μM). After 3h incubation at 25 °C, the reaction mixtures were fractionated by 15 % SDS-PAGE as described above and blotted onto a PVDF membrane. The biotinylated products were detected with avidin-peroxidase conjugate (Sigma) using diaminobenzidine (DAB) staining.

Hydrolytic cleavage of the labeled lectins.

The coumarin labeled lectins were prepared by the following conditions: 10 μM lectins (Con A, WGA, Cong II), 50 μM acyl donor **8**, and 50 μM sugar tethred-DMAPs **1, 3, 4** in 50 mM HEPES buffer (pH 8.0). After 3h incubation at 25°C, the reaction mixture was adjusted to pH 12.0, and incubated at 25 °C for 1days. The hydrolytic cleavage was observed by 15 % SDS-PAGE.

2-6. Reference

1) (a) Means, G. E.; Feeney, R. E. *Bioconjugate Chem.* **1990**, *1*, 2-12. (b) Chen, G.; Heim, A.; Riether, D.; Yee, D.; Milgrom, Y.; Gawinowicz, M. A.; Sames, D. *J. Am. Chem. Soc.* **2003**, *125*, 8130-8133. (c) Hayashida, O.; Hamachi, I. *Chem. Let.* **2003**, *32*, 632-633. (d) Dormán, G.; Prestwich, G. D. *Trends Biotechnol.* **2000**, *18*, 64-76. (e) Yasui, N.; Koide, T. *J. Am. Chem. Soc.* **2003**, *125*, 15728-15729.

2) (a) Srinivsan, R.; Huang, X.; Ng, S. L.; Yao. S. Q. *ChemBioChem* **2006**, *7*, 32-36. (b) Chan, E. S. W.; Chapttppaday, S.; Panicker, R. C.; Huangm X.; Yao, S. Q. *J. Am. Chem. Soc.* **2004**, *126*, 14435-14446.

3) (a) Barglow, K. T.; Cravatt, B. F. *Nature Methods* **2007**, *4*, 822-827. (b) Evans, M. J.; Cravatt, B. F. *Chem. Rev.* **2006**, *106*, 3279-3301. (c) Saghatelian, A.; Jessani, N.; Joseph, A.; Humphrey, M.; Cravatt, V. F. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 10000-10005.

4) Takaoka, Y.; Tsutsumi, H.; Kasagi, N.; Nakata, E.; Hamachi, I. *J. Am. Chem. Soc.* **2006**, *128*, 3273-3280.

5) Hatanaka, Y.; Kempin, U.; Jong-Jip, P. *J. Org. Chem.* **2000**, *7*, 5639-5643.

6) Lee, W.-r.; Jung, D.-W.; Williams, D.; Shin, I. *Org. Lett.* **2005**, *7*, 5477-5480.

7) Ilver, D.; Arnqvist, A.; Ogren, J.; Firck, I.-M.; Kersulyte D.; Incecik, E. T.; Berg, D. E.; Covacci, A.; Engstrand, L.; Boren, T. *Science*, **1998**, *279*, 373-377.

8) Nakata, E.; Koshi, Y.; Koga, E.; Katayama, Y.; Hamachi, I. *J. Ame. Chem. Soc.* **2005**, *122*, 13253-13261.

9) Ishi, M.; Matsumura, S.; Toshima, K. *Angew. Chem. Int. Ed.* **2007,** *46*, 8396-8399.

10) Hofle, G.; Steglich, W.; Vorbruggen, H. *Angew. Chem. Int. Ed.* **1978**, *17*, 569-583.

11) (a) Kamiya, H.; Muramoto, K.; Goto, R. *Dev. Comp. Immunol.* **1988**, *12*, 309-318. (b) Ogawa, T.; Ishi, C.; Suda, Y.; Kamiya, H. *Biosci. Biotechnol. Biochem.* **2002**, *66*, 476-480. (c) Shirai, T.; Matsui, Y.; Shionyu-Mitsuyama, C.; Yamane, T.; Kamiya, H.; Ishii, C.; Ogawa, T.; Mramoto, K. *J. Mol. Biol.* **2002,** 321, 879-889. (e) Shionyu-Mitsuyama, C.; Ito, Y.; Konno, A.; Miwa, Y.; Ogawa, T.; Muramoto, K.; Shirai, T. *J. Mol. Biol.* **2005**, *347*, 385-397.

12) (a) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **1998**, *37*, 2754-2794 (b) Gestwick, J. E.; Cairo, C. W.; Strong, L. E., Oetjen, K. A.; Kiessling, L. L. *J. Am. Chem. Soc.* **2002**, *124*, 14922-4933. (c) Hayashida, O.; Takaoka, I.; Hamachi, I. *Tetrahedron Let.* **2005**, *46*, 6589-6592

13) Bhattacharya, S.; Snehalatha, K. *J. Chem. Soc., Perkins Trans.* **1996**, *2*, 2021-2025.

14) Ueno, U.; Jiao, G.-S.; Burgess, K. *Synthesis* **2004**, *15*, 2591-2593.

15) Song, A; Wang, X; Lam, K. S.; *Tetrahedron Lett.*, **2003**, 44, 1755-8.

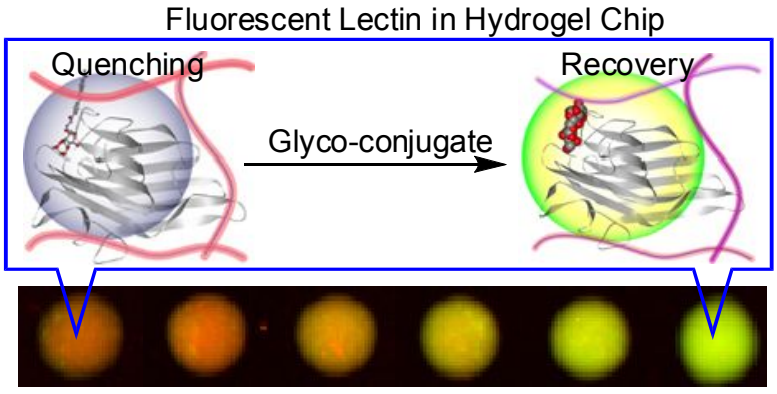
16) Haugland, R. P. *Handbook of Fluorescent Probe and Research Chemicals, Nine Ed.*; Molecular Probes, Inc.: Eugen, OR, 2002

Chapter 3

A fluorescent lectin array using supramolecular hydrogel for simple detection and pattern profiling for various glycoconjugates.

Abstract

Since sugar and its derivatives play important roles in various biological phenomena, the rapid and high throughput analysis of various glyco-conjugates is keenly desirable. We describe herein the construction of a novel fluorescent lectin array for saccharide detection using a supramolecular hydrogel matrix. In this array, the fluorescent lectins were noncovalently fixed under semi-wet conditions in order to suppress the protein denaturation. It is demonstrated by fluorescence titration and fluorescence lifetime experiments that the immobilized lectins act as a molecular recognition scaffold in the hydrogel matrix, similar to that in aqueous solution. That is, a bimolecular fluorescence quenching and recovery (BFQR) method can successfully operate under both conditions. This enables one to fluorescently read-out a series of saccharides on the basis of the recognition selectivity and affinity of the immobilized lectins without tedious washing processes and without labeling the target saccharides. Simple and high-throughput sensing and profiling were carried out using the present lectin array for diverse glyco-conjugates, that not only included a simple glucose, but also oligosaccharides, and glycoproteins, and furthermore, the pattern recognition and profiling of several types of cell lysates were also accomplished.



3-1. Introduction

3-1-1. DNA Microarray. [1]

In post-genome era, "Microarray technology" is one of the most promising approaches that enable the large-scale analysis of whole biological compounds such as nucleic acids and proteins simultaneously and rapidly. Thus, the microarray technologies are currently essential for progress of bio-technologies and bio-science such as bioinformatics and tailored medicine.

DNA microarray is the most representative microarray, for which thousand of oligo-nucleotides are synthesized on a chip by photolithographic methods (Fig.1a) or polymerase chain reaction (PCR) products are spotted onto a chip (Fig.1b). In order to analyze mRNA transcription levels expressed under various conditions, the DNA microarray is hybridized with fluorescent-labeled cDNAs (Fig.1c). This technique, so called "cDMA array", is applied to cancer diagnosis, discovery of biomarkers or drug targets and so on. Other than the gene expression analysis, the DNA microarray is used for single nucleotide polymorphism (SNP) analysis and comparative genomic hybridization (CGH).

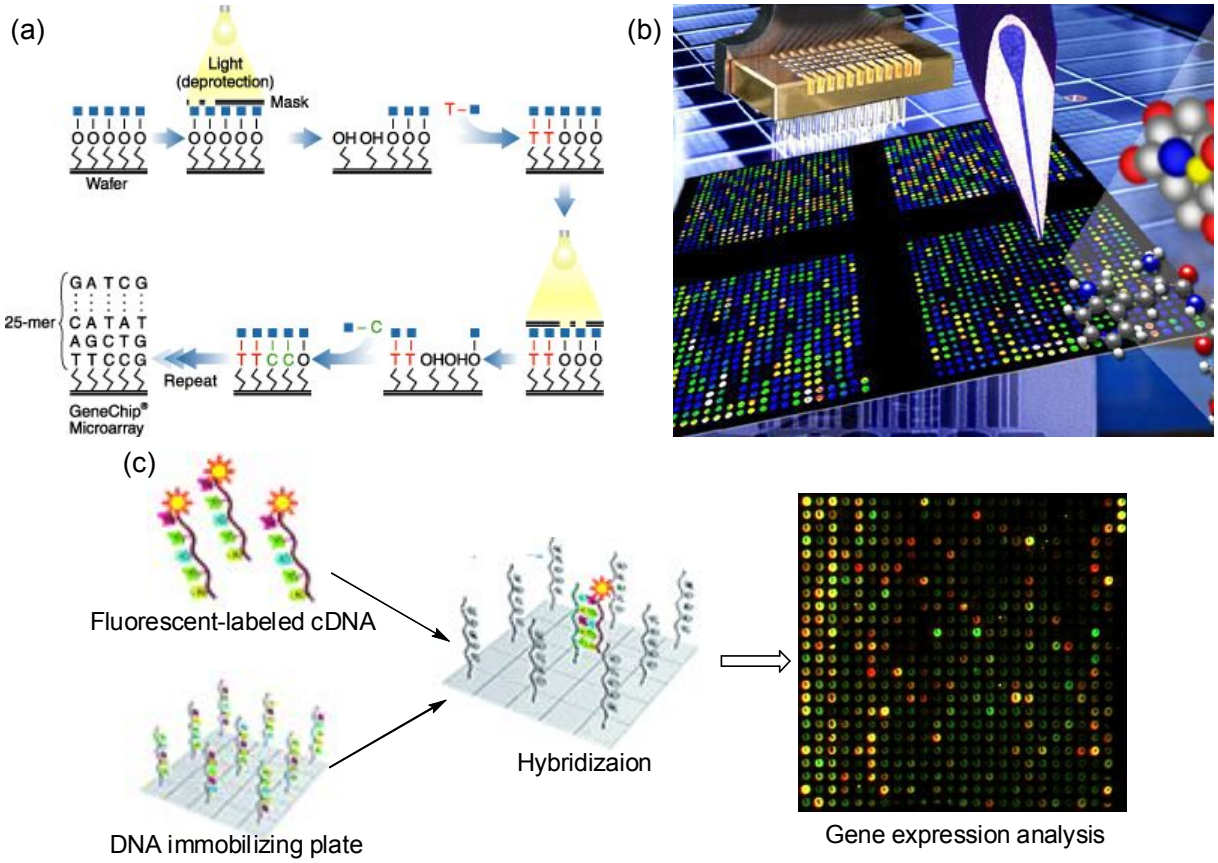


Figure 1. (a) Oligo-nucleotide synthesis on chip using photolithographic methods.[2a] (b) Oligo-nucleotide printing technology. [2b] (c) Schematic illustration of cDNA array [2c]

3-1-2. Protein microarray [3]

In proteomic study, the importance of protein microarray technology is increasing. Depending on their purposes, various capture agents are used in the protein microarray (Fig.2). In order to determine the abundance of protein of interest, antibodies or DNA/RNA aptamers are used as specific capture agents that selectively bind to each target protein. The other purpose is to find out the functions of target proteins, including protein-protein interactions, receptor-ligand interactions and so on. In this case, the molecules interacting with the target proteins such as proteins and small molecules are immobilized on a chip. In order to establish protein microarray, it is necessary to place capture agents in such a way as to maintain their active forms on a solid surface. However, the immobilization of proteins is generally difficult because proteins are easily inactivated.[4] As detection methods analyzing the protein microarray, a sandwich-type immunoassay using labeled antibody is generally used. Surface Plasmon Resonance (SPR) and Quarts Crystal Microbalance (QCM) are known as real-time detection methods.

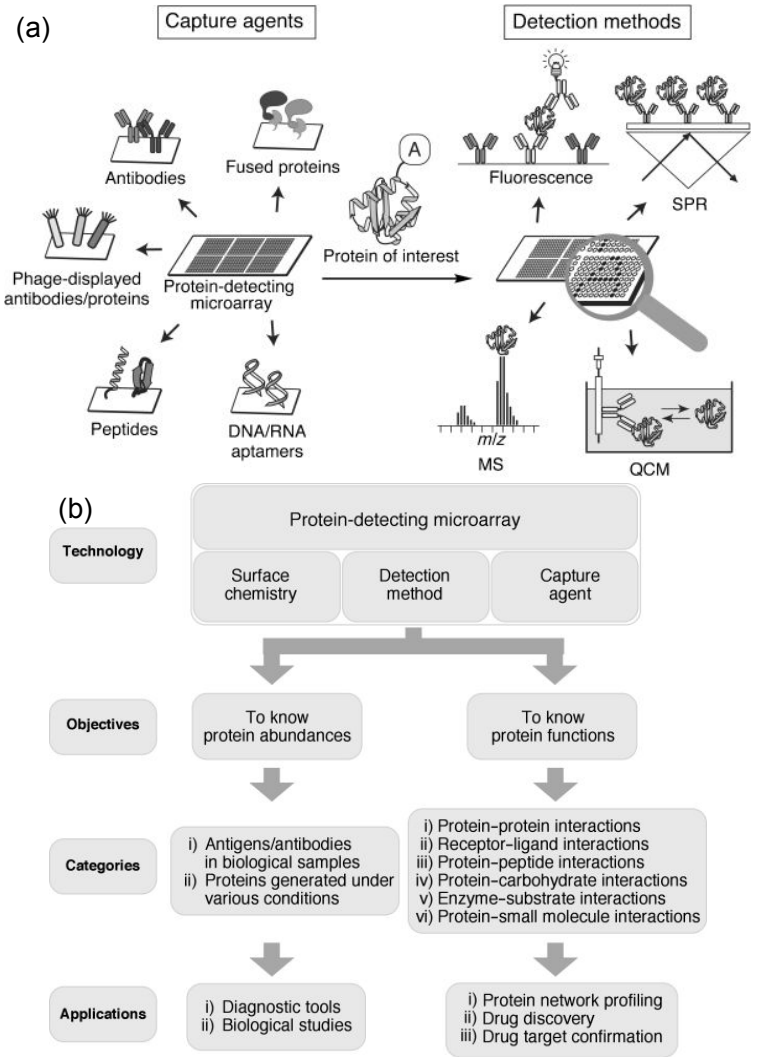


Figure 2. (a) Classification of protein capture agents and diction method.[3a]
(b) Scheme of protein microarray detection. [3a]

3-1-3. Carbohydrate microarray [5]

Recent advances in glyco-biology and glyco-chemistry has revealed that sugar and its derivatives play important roles in various biological phenomena.[6] Therefore, various types of carbohydrate chips, which immobilized carbohydrates on a solid substrate, have been actively developed by many researchers to detect or estimate sugar-protein interactions (Fig. 3a), to assay the activity of the glycosylation-related enzymes or proteins (Fig. 3b),or to discover inhibitors (Fig. 3f) and so on. Interestingly, it was reported that cells of interest were directly detected by the carbohydrate microarray in order to analyze sugar receptors on cell surface (Fig. 3d). To advance the carbohydrate microarray, methods for immobilizing sugars by various organic chemical reactions were actively developed (Fig.4).

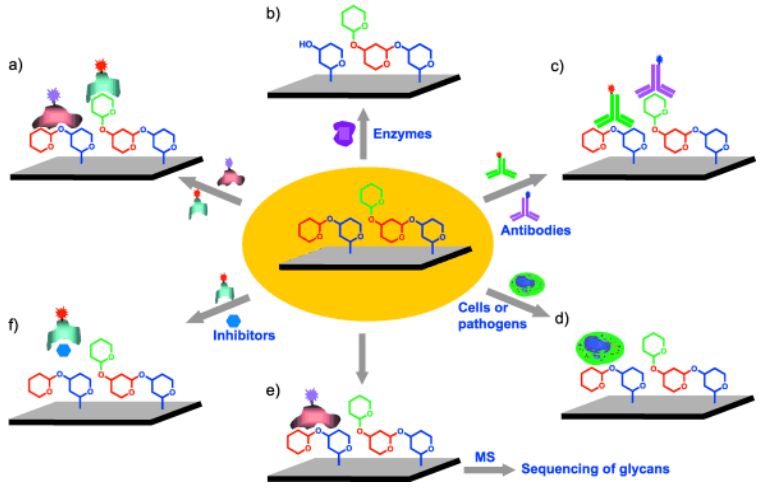


Figure 3. Schematic illustration of carbohydrate microarray [5a]

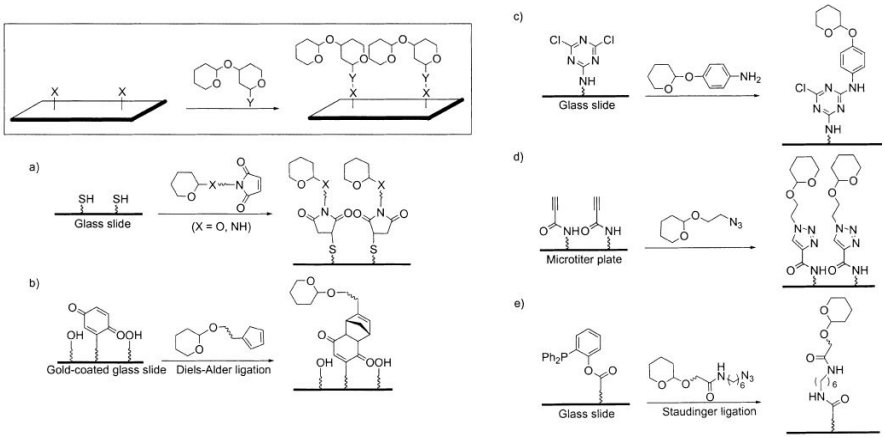


Figure 4. Immobilizing method on carbohydrate microarray [5d]

3-1-4. The concept of lectin microarray

The carbohydrate microarray can not detect sugars and its derivatives, although it can analyze various events involving sugars. In order to detect sugars of biological significance, it is necessary to employ sugar binding proteins (lectin) as molecular tools for the sugar recognition. In this chapter, I developed a unique fluorescent lectin array using a supramolecular hydrogel as an immobilizing matrix.

3-2.Results

3-2-1. Biomolecular fluorescence quenching/recovery method (BFQR) in aqueous solution

With the aim of constructing a lectin chip using a supramolecular hydrogel, we initially comfirmed a lectin-based fluorescence read-out method for a saccharide in aqueous solution.[7] That is a bimolecular fluorescence quenching/recovery (BFQR, Fig.5) method including the FITC (fluoresceinisothiocyanate) labled concanavalin A (FITC-Con A, a mannose and glucose binding lectin) and Mannobiose-appended Dabcyl quencher (**2**, Fig.6). As shown in Fig.7, the fluorescence intensity of FITC-Con A at 516 nm significantly decreased by the addition of the quencher **2** and recovered by Mannotirose (Man-3), the strongest ligand for Con A. Figure 7c shows the fluorescence titration curve of FITC-Con A depending on the concentration of **2** or Man-3. The effective quenching did not occur using Dabcylic acid **6**, a quencher lacking the Mannobiodse unit, and the quenched fluorescence did not recover by, inappropriate sugars such as galactose (Gal), which does not bind Con A. These results indicate that the present fluorescence quenching and recovery can be ascribed to the binding of **2** and its replacement with Man-3, respectively, in the sugar binding pocket of FITC-Con A.

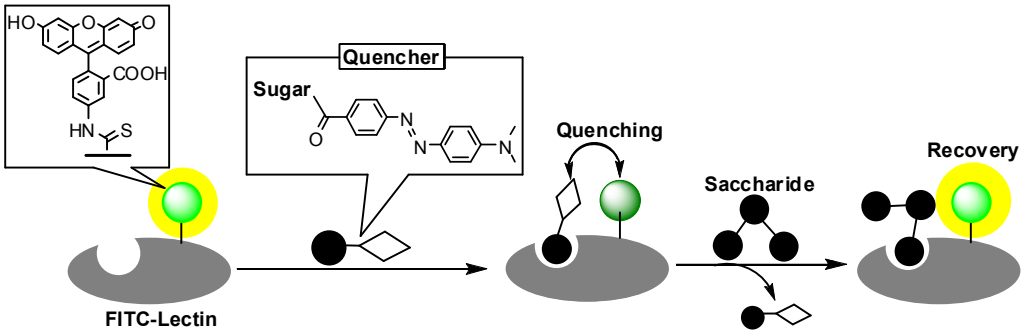


Figure 5. Schematic illustration of BFQR fluorescent detection system

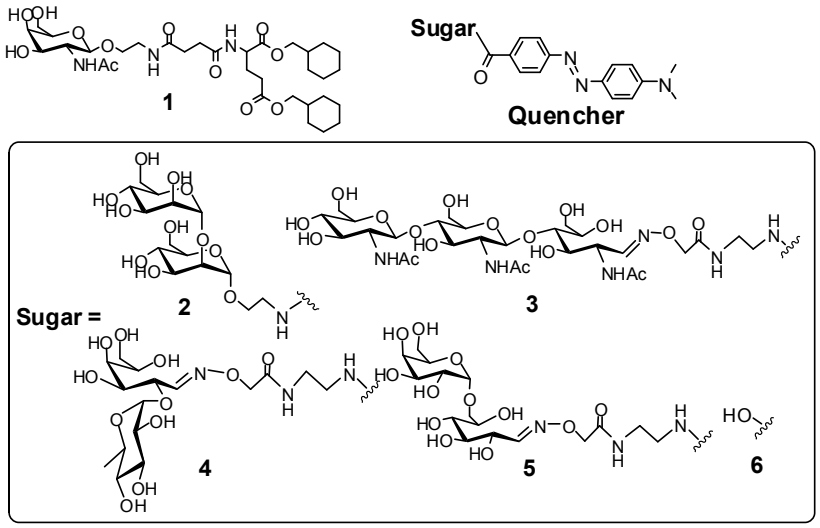


Figure 6. Molecular structure of gelator **1** and quenchers **2-5.**

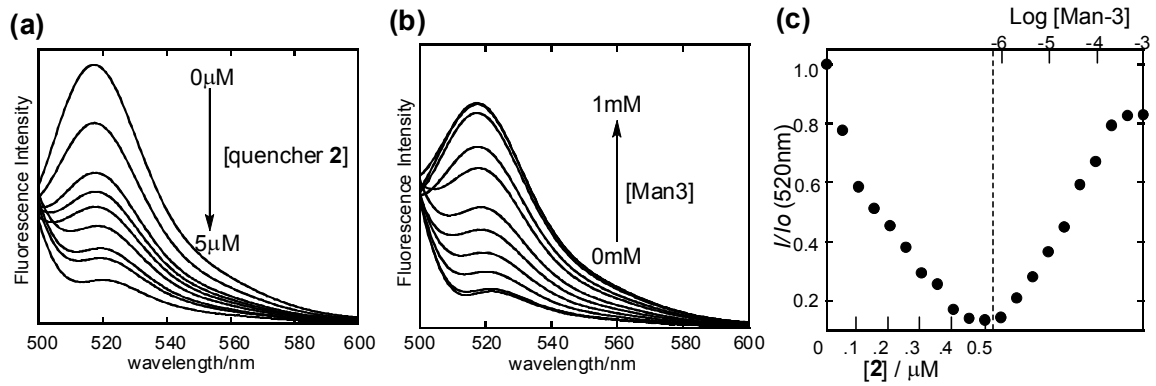


Figure 7. (a) Fluorescence quenching process of FITC-Con A upon addition of **2** (0 – 0.5 μM). (b) Fluorescence recovery of FITC-Con A by the addition of Man-3 (0-1 mM). (c) The titration curve with **2** and then with Man-3. [FITC-Con A] = 0.1 μM, **2** (1 μM), Man-3 (0-100 μM) in 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$ and 0.1 M NaCl), 20 ˚C, λ_{ex} = 488 nm.

<u>3-2-2. Biomolecular fluorescence quenching/recovery method (BFQR) in supramolecular hydrogel.</u>

We applied the present BFQR method to the sugar detection in a hydrogel system comprised of a supramolecular hydrogelator **1** which spontaneously formed a transparent hydrogel without any polymerization reagents/treatment.[8] As recently reported by us, many proteins and enzymes can be immobilized in internal aqueous phases of the hydrogel matrix while retaining the natural activities that are shown in an aqueous solution. After the immobilization of FITC-Con A in the supramolecular hydrogel **1** on the 10 microL scale, the quencher **2** was added to the microgel (Fig.8). Figure 9a shows that the fluorescence peak at 520 nm gradually decreased and leveled off by the increase in **2**. The reduced fluorescence then significantly recovered by the addition of Man-3 (Fig.9b), the behavior of which is perfectly identical with that observed in the aqueous solution. Figure 9c shows the titration curves of various saccharides monitored by the fluorescence recovery process using the Con A chip in the presence of the quencher **2**. It is clear that the saccharides, which have an affinity to Con A, cause the fluorescence recovery in a manner similar to Man-3, whereas Gal and Me-β-Glc, the sugars having no-affinity to Con A, do not induce any significant fluorescence change even at a 100 mM concentration. The sensing selectivity is in the order of Man-3 > Me-α-Man > Me-α-Glc > Glc >> Gal, Me-β-Glc, that is in good agreement with the selectivity of the native Con A in aqueous solution. Although the minimum sensing concentrations are slightly higher than those of the native Con A due the competitive assay of the present BFQR method, the order in sensing agrees well with that evaluated in the aqueous system. The saturation curves also provided the apparent binding constants, by which we can evaluate the net binding constants for each saccahride as follows: 7.0 X 10^4 M^{-1} for Man-3, 1.6 X 10^3 M^{-1} for Me-α-Man, 4.7 X 10^2 M^{-1} for Me-α-Glc, and less than 10 M^{-1} for Gal.[9] These values are almost identical to the values reported in the aqueous system, demonstrating that the original function of the native Con A is retained in the supramolecular hydrogel matrix.

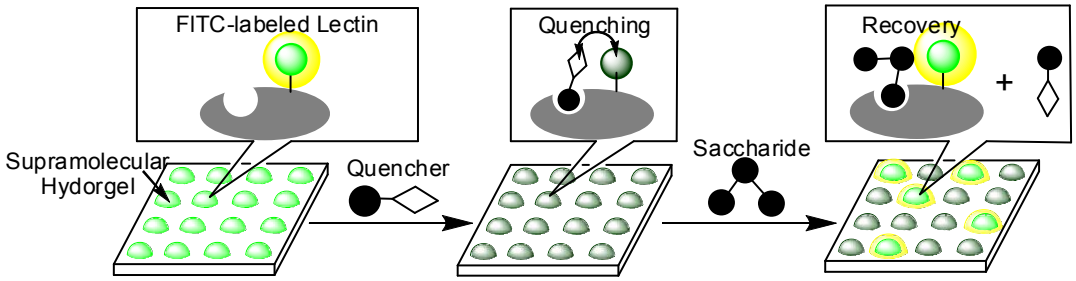


Figure 8. Schematic illustration of BFQR fluorescent detection system on hydrogel chip.

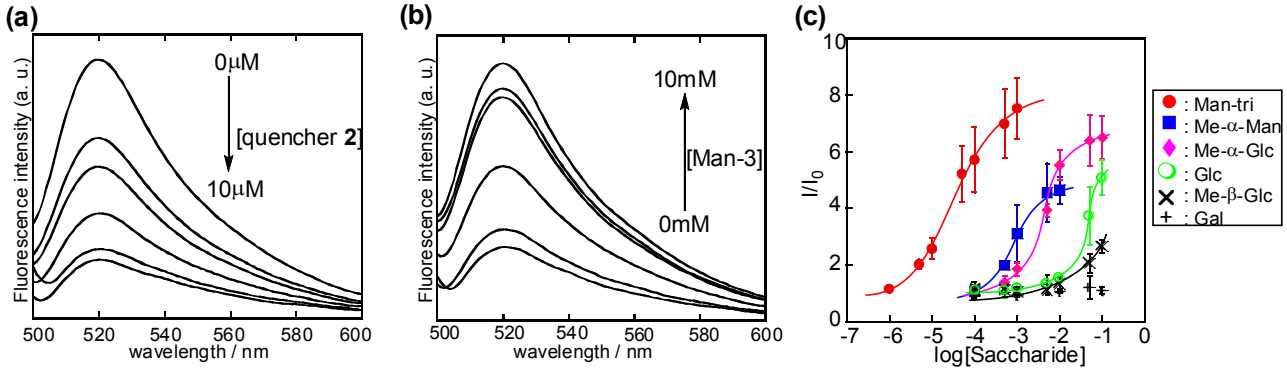


Figure 9. (a) Fluorescence quenching process of FITC-Con A immobilized in the hydrogel matrix by the addition of **2** (0 - 10 μM). (b) Fluorescence recovery of FITC-Con A immobilized in the hydrogel matrix by the addition of Man-3 (0-1 mM). [FITC-Con A] = 1 μM, 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$ and 0.1 M NaCl), in 0.25 wt% of **1**, λ_{ex} = 475nm. (c) Fluorescence titration plots of the relative intensity (I/I_0) of FITC-Con A versus the saccharide concentration (log[saccharide]) in the hydrogel **1** (0.25 wt%): FITC-Con A (0.5 μM), **2** (5 μM) : Man-3 (●), Me-α-Man (■), Me-α-Glc (◆), Glc (○), Me-β-Glc (×), Gal (+). The error bars in graph mean standard deviation of independent 3 measurements.

3-2-3. Fluorescence life time study for BFQR method.

The BFQR events were also confirmed by the fluorescence lifetime measurement.[10] In the absence of the quencher **2**, the decay curve of FITC-Con A in the hydrogel as shown in Fig.10b, gave the fluorescence lifetime of 3.3 ns, the value of which is almost identical with that observed in aqueous solution (3.2 ns in Fig.10a). This suggests that the fluorescein moiety of FITC-Con A embedded in the hydrogel is located in almost the same microenvironment as in the aqueous solution. The lifetime was shortened to 0.4 ns by the addition of the quencher **2** and then subsequently returned to 3.1 ns by the further addition of Man-3, but not by the addition of Gal (0.4 ns). These are reasonably ascribed by the **2**-complexation-induced quenching of the fluorescein of FITC-Con A and the fluorescence recovery by the replacement with Man-3 on the sugar-binding pocket of FITC-Con A, respectively. Such behaviors are again similar to those in aqueous solution (0.4 ns (in the presence of **2**), 3.2 ns (in the presence of **2** and Man-3)). Therefore, it is concluded that the BFQR type of fluorescence read-out method, which has been usually used in an aqueous solution, can successfully operate under the semi-wet supramolecular hydrogel conditions.

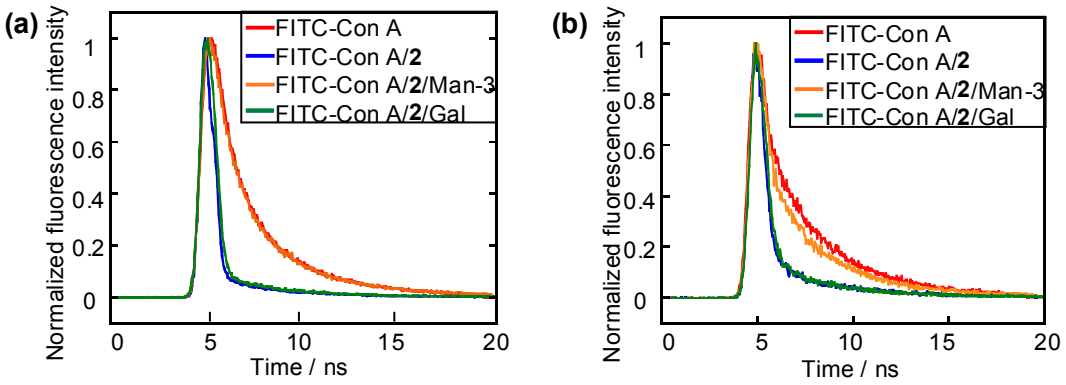


Figure 10. Fluorescent lifetime study of FITC-Con A in aqueous solution (a) and in hydogel **1** matrix (b). Time-dependent fluorescence decay curve of FITC-Con A (0.5 μM) was recorded from 500 to 540 nm after laser excitation at 337 nm in the absence (Red line) or presence of **2** 5μM, (blue line). The decay of FITC-Con A was also measured in the presence of **2** and Man-3 (1 mM, orange line), or in the presence of **2** and Gal (1 mM, green line).

3-2-4.Visualization of glucose and mannose derivatives using semi-wet Con A chip.

Due to the fact that the fluorescence intensity change in the present microgel is clear enough to be distinguished by the naked eye, a semi-wet Con A chip can be used for the rapid estimation of the sugar selectivity of lectin as well as for the convenient detection of a family of saccharides. Figure 11a displays a fluorescence image of the Con A array containing 7 distinct mono-saccharides at different concentrations (0 - 10 mM). According to the binding affinity of the native Con A, a strong green fluorescence appeared at 1mM Me-α-Man, and at 10 mM Man or Me-α-Glc. Glc was also weekly detected at spots containing 10 mM. In a series of oligosaccharides, on the other hand (Fig.11b), the α-linked oligo-mannosides were more sensitively detected in the 1 mM concentration range for Man-bi, the 0.1 mM range for Man-tri, and 0.01 mM for Man-tetra and Man-penta monitored by the fluorescence recovery image. In contrast, the glucose type of oligosaccharides are less sensitively detected (more than 1 mM concentration range) due to the lower affinity to the native Con A. Neither the β-linked saccharides, such as Me-β-Glc and cellobiose, nor the galactose-terminated derivatives, such as Me-α-Gal and lactose, are sensed by this Con A chip. These results imply that the saccharide selectivity of a lectin was rapidly and conveniently evaluated using the present semi-wet lectin chip.

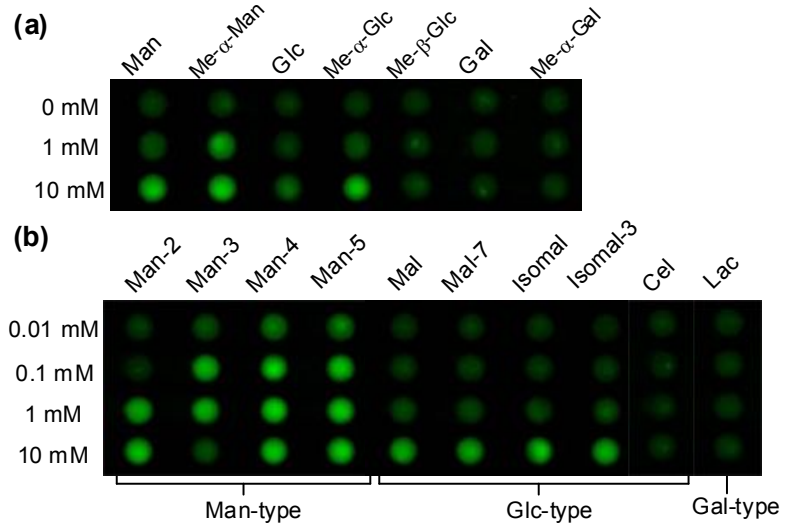


Figure 11. Fluorescence image of Con A chip based on supramolecular hydrogel upon addion (a) mono-saccharides, and (b) oligo-saccharides. [FITC-Con A] = 0.1 μM, [quencher **2**] = 1 μM.

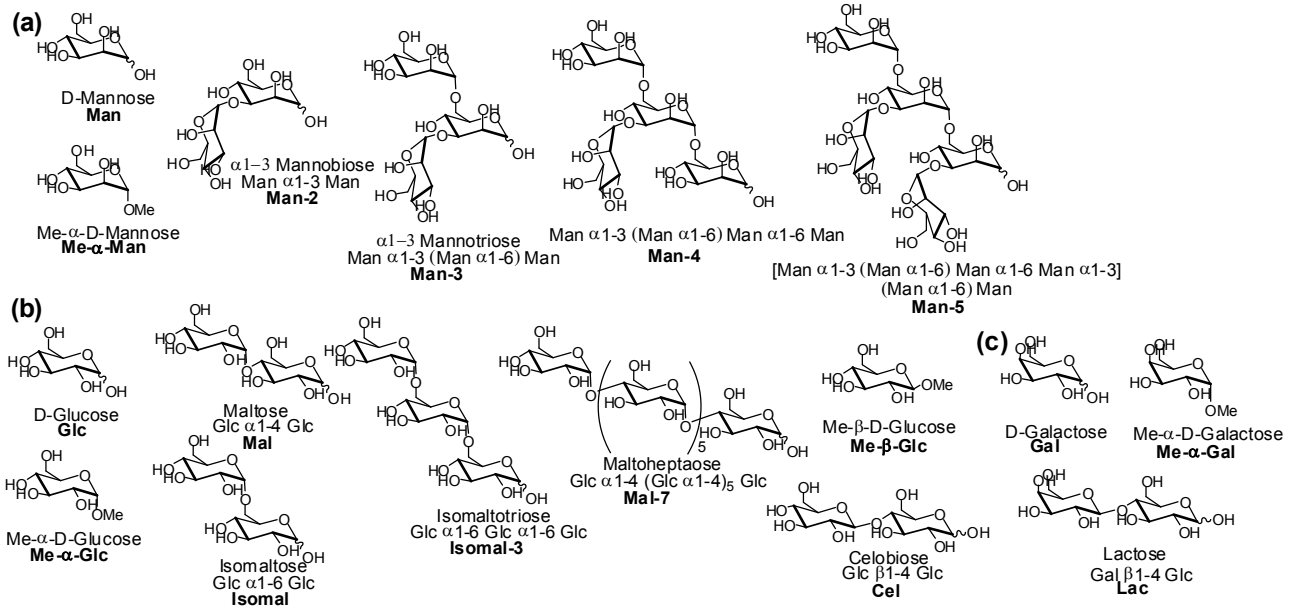


Figure 12. Molecular structures of mannose (a), glucose (b) and galactose (c) derivatives.

3-2-5.Visualization of glucose and mannose derivatives using semi-wet Con A chip.

As proof of the principle experiment to demonstrate the practical utility of the lectin chip, we next conducted the fluoro-colorimetric imaging of glucose (Glc), an important analyte for the diagnosis of diabetes.[11] To produce a clear fluorescent color change, the present semi-wet Con A chip is equipped with a ratiometric mode. According to our previous finding, the present hydrogel matrix possesses many hydrophobic microfibers that can incorporate a hydrophobic fluorescent probe. On the basis of this unique feature, we designed a method to place octadesyl-rhodamin B (OR-B, 13a), an additional fluorophore, in the hydrophobic fiber domain of the supramolecular hydrogel **1**. Since the BFQR event occurred in the aqueous cavity of the hydrogel, the fluorescence of the OR-B localized in the hydrophobic domain should be practically insensitive to the sugar binding of FITC-Con A, so as to become an internal standard in this system. When the Glc

solution was added to the Con A chip containing both the quencher **2** and OR-B, the intensity ratio of FITC over OR-B (FITC/OR-B) changed depending on the Glc concentration as shown in Fig.13c. It is known that the change in the Glc concentration in the range from 5 mM to 20 mM is critical for the diabetes diagnosis. In such a concentration range, it is clear that the ratio value (FITC/OR-B) linearly increased and more interestingly, the fluorescent color change of the ratiometric Con A chip was clearly detected from reddish orange to yellowish green (the second lane of Fig.13b), in contrast to the case of the Con A chip using the simple intensity mode (the first lane of Fig.13b).

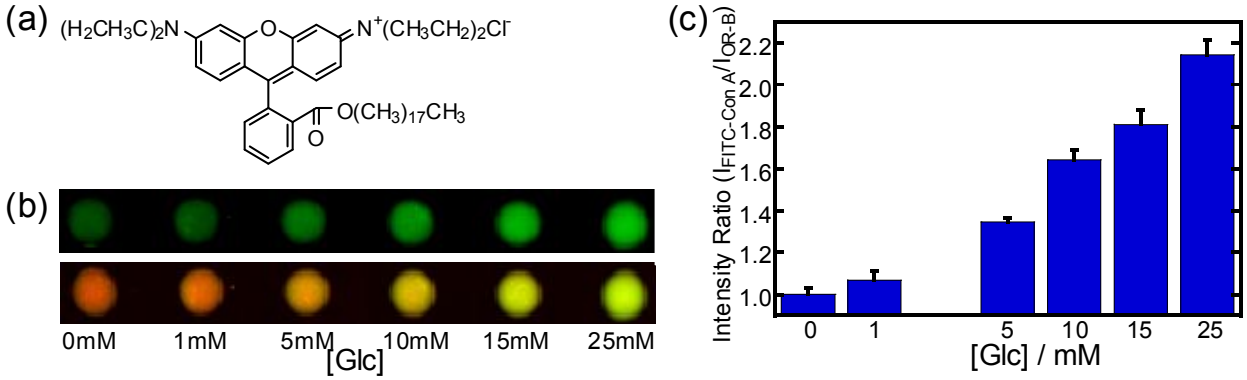


Fig.13 (a) Molecular structure of octadecyl rhodamine B chloride (OR-B). (b) Clear image of glucose using Con A chip containing OR-B. Upper; fluorescent image of FITC-Con A for various concentration of glucose by fluorescein-excitation. Lower; fluorescent merged image by fluorescein (green) excitation and by OR-B (red) excitation. (c) The graph of fluorescent ratio (F-Con A / OR-B) by the addition of various concentration of glucose. [FITC-Con A] = 0.1 μM, [quencher **2**] = 0.5 μM. The bar height and the error bar were estimated by the average and standard deviation of 4 spots on the same plate

3-2-6. Biomolecular fluorescence quenching/recovery method (BFQR) for other lectins

The BFQR method is so general that it can be applied to other fluorescein-tethered lectins showing distinct saccharide selectivities in the presence of a pair of corresponding sugar-appended Dabcyl quenchers as shown in Fig 6. In these 5 lectins (WGA, GSL-II, AAL, UEA-I and GSL-I; the sugar selectivity is shown in Table 1), we confirmed that the quenching and recovery processes. Emission of FITC-labled lectins were decreased by the addition of suitable quenchers, and then increased by the addition of the corresponding saccharides in the presence of quenchers. These results are consistent with in aqueous solution (Fig.14) and in hydrogel matrix (Fig.15).

Table 1 Carbohydrate specificities and suitable quenchers of FITC-labeled Lectins

Lectin	Carbohydrate Specificity	Suitable Quencher
Con A	α-Man,α-Glc	**2**
WGA	β-GlcNAc, α2,3-NeuNAc	**3**
GSL-II	α,β-GlcNAc	
AAL	Fuc	**4**
UEA-I	α1,2-Fuc	
GSL-I	α–Gal,GalNAc	**5**

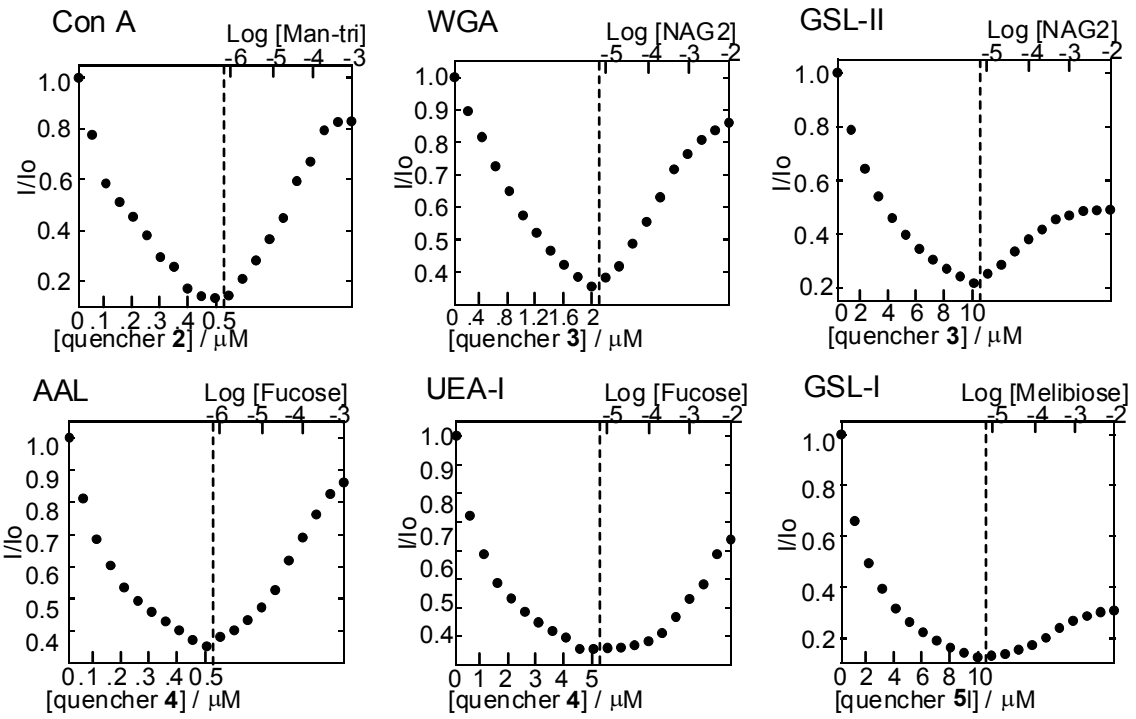


Figure 14. Fluorescent titration curve of Fluorescein-labeled lectins with quenchers and saccharides in aqueous solution. [FITC-lectin] = 0.1 μM in 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$, 0.1 M NaCl), 20 ˚C, λ_{ex} = 488 nm.

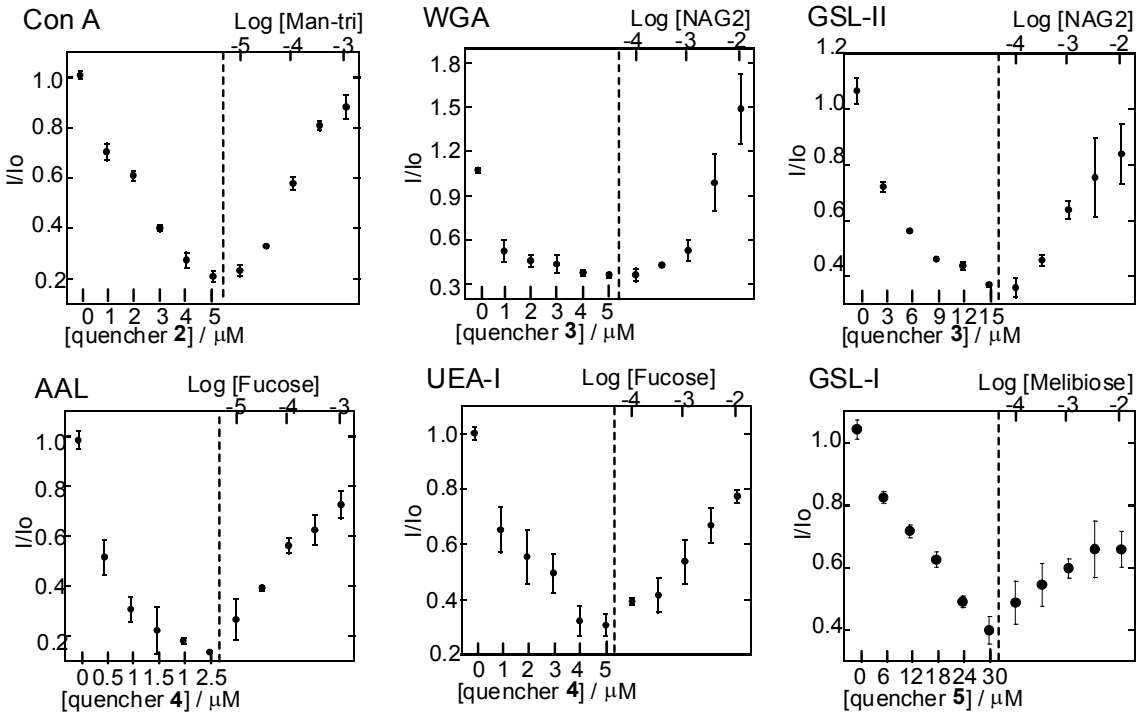


Figure 15. Fluorescent titration curve of Fluorescein-labeled lectins with quenchers and saccharides in hydrogel matrix. [FITC-lectin] = 1 μM in 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$, 0.1 M NaCl), in 0.25 wt% of **1,** 20 ˚C, λ_{ex} = 475 nm. The error bar in graph mean standard deviation of independent 3 times measurement.

3-2-7. Fluorescent detection of simple saccharide using lectin array.

6 kinds of fluorescent lectins were arrayed on a glass plate to produce a lectin array as shown in Fig. 16a. Without sugars, only a weak emission was detected from all the gel spots in the presence of the corresponding quenchers (Fig. 16b). By the addition of Me-α-mannose (Me-α-Man) to these spots, four spots containing Con A strongly emitted a green fluorescence whereas the other spots containing other lectins having no affinity to Man remain dark. In the case of the addition of the *N,N'*-diacetyl chitobiose (NAG-2), the emission increased in the 8 spots containing WGA or GSL-II, both of which display the affinity to the GlcNAc family. Similarly, the Fucose addition induced the recovered emission at spots containing AAL and UEA-I, and Melibiose did it at the spots of GSL-I.

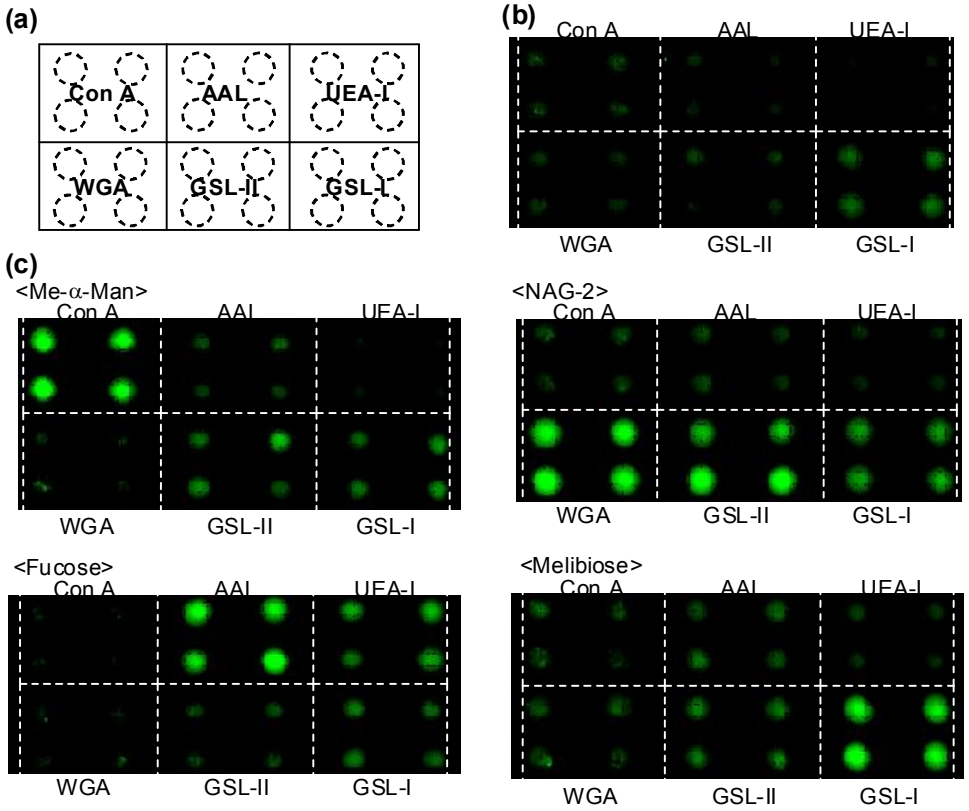


Figure 16. Mono- or oligo-saccharides detection by lectin array. (a) Spatial position of 6 kinds of lectins on hydrogel chip. One lectin was arrayed at 4 spots, so that six lectins were arrayed in 24 spots. (b) Fluorescent image of the lectin array before the addition of analyte. (c) Fluorescent images after the addition of saccharides (Me-α-Man, NAG-2, Fucose, Melibiose (10mM)): [FITC-Con A] = 0.1 μM, [quencher **2**] = 1 μM; [FITC-WGA] = 0.1 μM, [quencher **3**] = 2 μM; [FITC-GSL-II] = 0.1 μM, [quencher **3**] = 30 μM; [FITC-AAL] = 0.1 μM, [quencher **4**] = 1 μM; [FITC-UEA-I] = 0.1 μM, [quencher **4**] = 10 μM; [FITC-GSL-I] = 0.1 μM, [quencher **5**] = 30 μM.

3-2-8. Fluorescent detection and profiling of glycoprotein using lectin array.

Not only the simple saccharides, but also the branched sugar structures tethered to a glycoprotein surface can be roughly characterized by this lectin array. As shown in Fig.17a, the addition of ribonuclease B (Ribo B), which has a high mannoside branch[12a], induced the strong emission at the spots of Con A, and the Fetuin addition intensified the emission at the four spots of WGA due to the tethered sialic acid moiety.[12b] Unlike Ribo B and Fetuin, the fucose-appended Mucin caused the emission intensification at the spots of AAL and UEA-1, [12c] and ovalbumin having the terminal GlcNAc unit intensified the emission at the WGA and GSL-II spots. [12d] In sharp contrast, there are no considerable emission changes observed in the case of the addition of BSA undergoing no saccharides decoration. Therefore, the comparison of the bar graph pattern for the emission recovery by the distinct glycoproteins (Fig. 17b) clearly shows the difference among the types of glycoproteins on the basis of the structural difference of the attached saccharides.

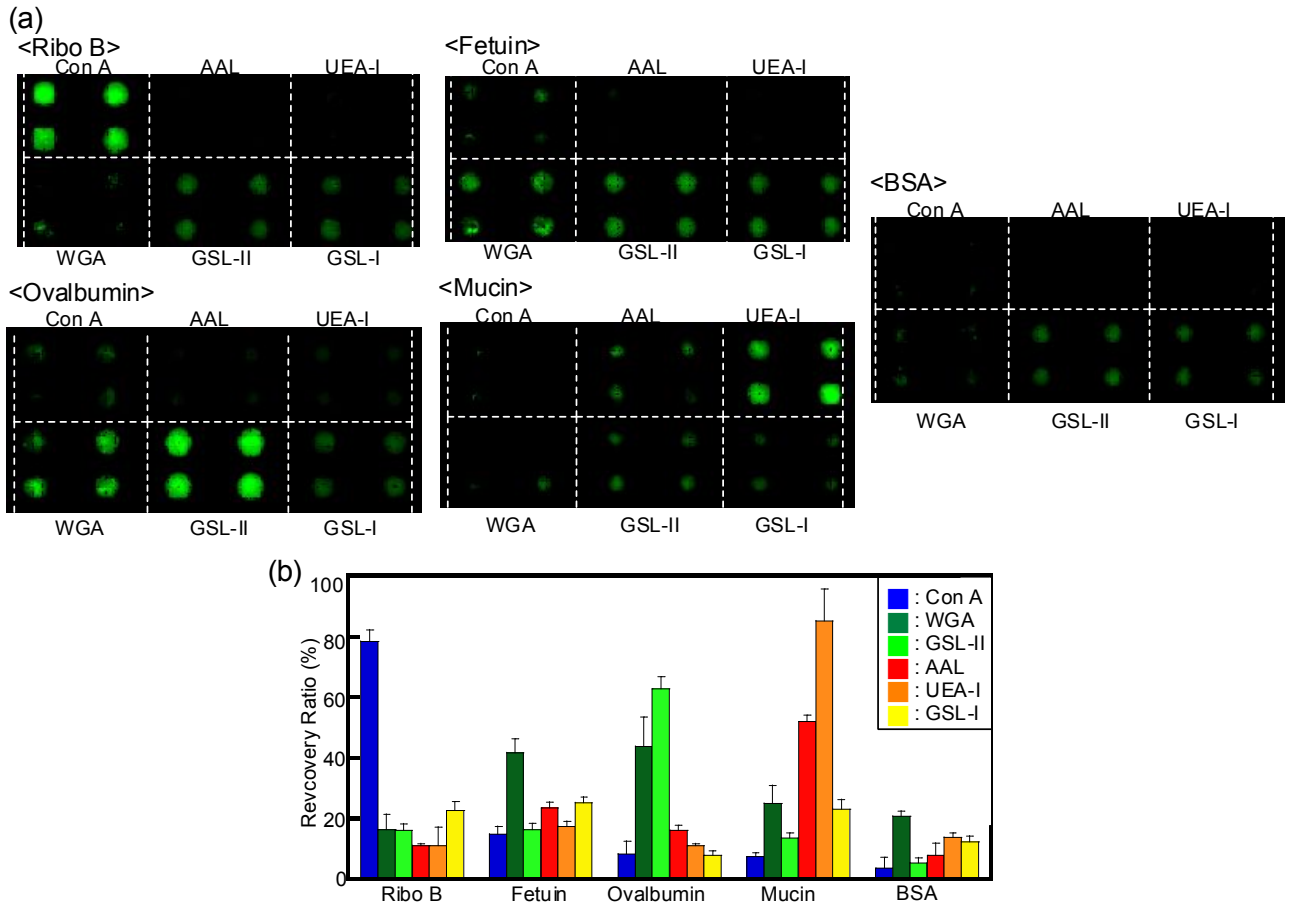


Figure 17. Analysis for glycoproteins (Ribo B: a high-mannose-branched proteins, Fetuin: a sialylated N-type glycan branched proteins, Ovalbumin: a hybrid N-type glycan branched proteins, Mucin; an O-type glycan branched proteins) and no glycosylated protein (BSA) by lectin chip. Fluorescent images (a) and the graph of fluorescent recovery ratio (b) by the addition of proteins (40 μg/μL). Con A (blue), WGA (green), GSL-II (yellow-green), AAL (red), UEA-I (orange), GSL-I (yellow). The bar height and the error bar were estimated by the average and standard deviation of 4 spots on the same plate.

3-2- 9 . Fluorescent detection and profiling of cell lysate

For the more complicated analytes, cell lysates originating from different cell lines were profiled using the present lectin array. In the case of mammalian cell lines, a strong emission was observed at the spots of WGA and AAL (Fig. 18a). The strong response in WGA spots (but not in GSL-II) can be reasonably explained by the existence of sialic acid derivatives on many animal cell surfaces. [13a] The strong response of AAL is consistent with the established profile that there are several kinds of fucosylated sugar chains such as N-linked and Lewis type saccharides in the higher organisms.[13b] In contrast, the AAL spots did not respond to the addition of bacteria cells, but instead, the emission intensification was induced at the GSL-II spots, suggesting that there are many GlcNAc units in these cell lysate (Fig.18b). This is probably attributed to the fact that GlcNAc is the major component of the peptide-glycans located on the bacterial cell wall.

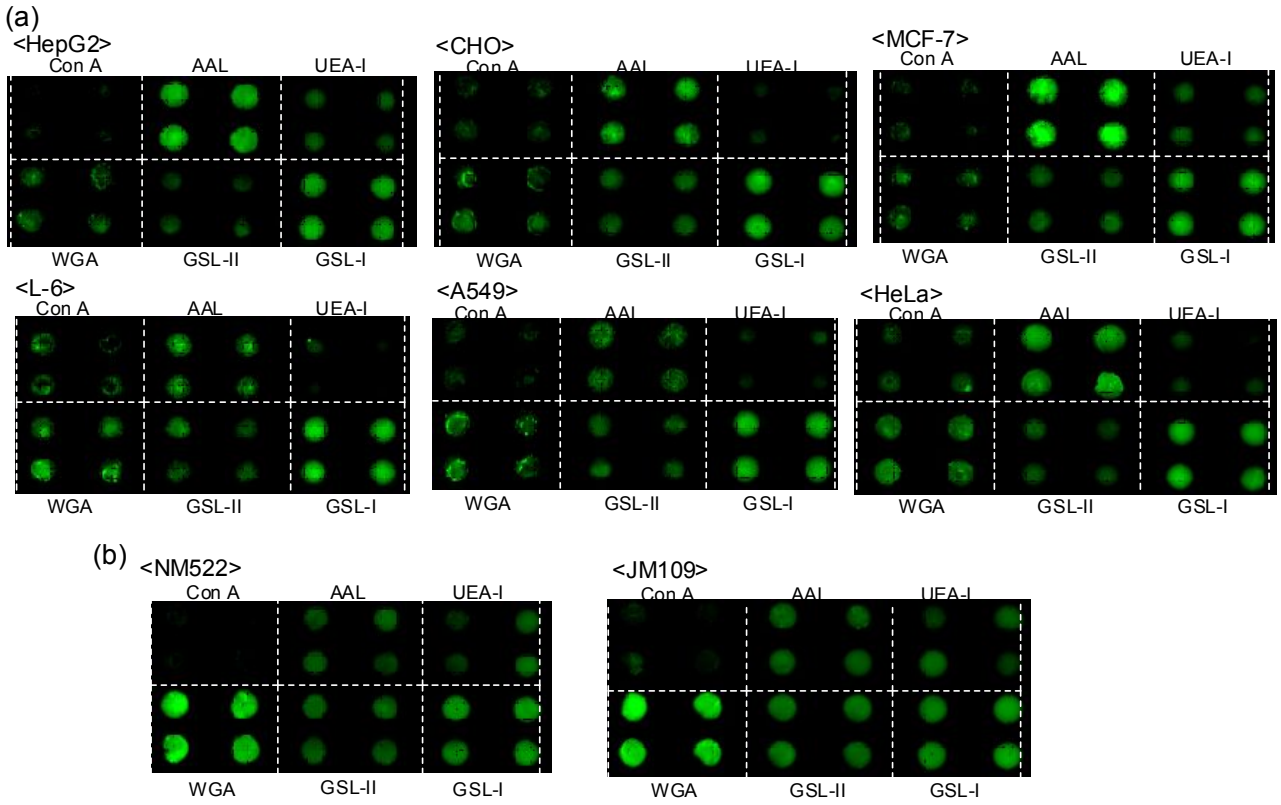


Figure 18. Fluorescent images for lectin chip analyzing mammalian cells (a, MCF-7, L-6, CHO, A549, HeLa, HepG2: 10,000 cells/μL) or bacteria cells (b, NM522, JM109: 10 μg/μL by wet weight).

3-3. Discussion

3-3-1. The advantage of pattern recognition

1-to-1 type discrimination, so called "Rock and Key" mode of molecular recognition, is expected to perfectly recognize a target molecule in an ideal case. However, it is practically difficult for the recognition of structurally complicated and diverse biological glyco-conjugates in many cases (Fig.19a). On the other hand, "Pattern Recognition" enables one to recognize a target molecule based on the response pattern of many receptors (Fig.19b).[14] By using the pattern recognition, it is anticipated to detect and discriminate a complicated glyco-conjugates more precisely.

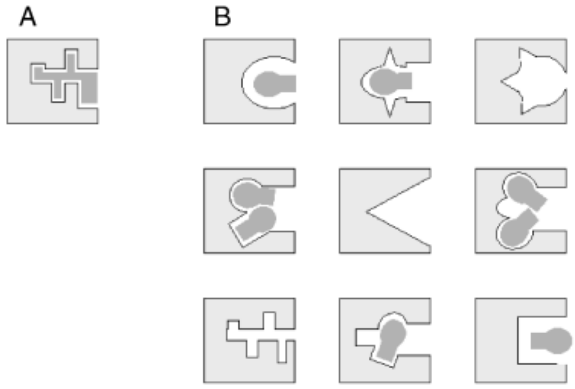


Figure 19. Schematic illustration of host-guest interaction. (A) Rock and key paradigm. (B) Pattern recognition using multi sensors. [14]

Lectins used as sugar recognition tools show high specificities for simple saccharides (mono-, oligo-saccharide). On the other hand, lectins can't specifically interact with the complicated sugar chains in the rock and key mode (Fig.20a). Using this property, a pattern recognition based on the sensing pattern of lectin microarray can be carried out here (Fig.20b). The pattern recognition using lectin microarray enabled not only to discriminate the structurally complicated glyco-conjugates but also to profile various cells. By analyzing the cell lysates, the original patterns of each cell were obtained on the basis of expressed gyco-conjugates in cells, and these patterns help to inform the cell types and the cell conditions (Fig.21).

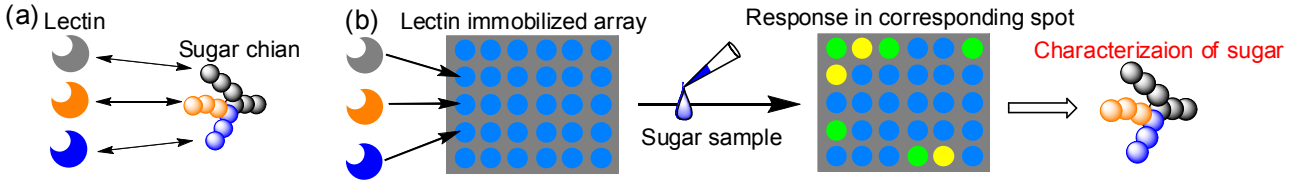


Figure 20 (a) Interaction with lectins and sugar. (b) Schematic illustration of analysis by lectin microarray.

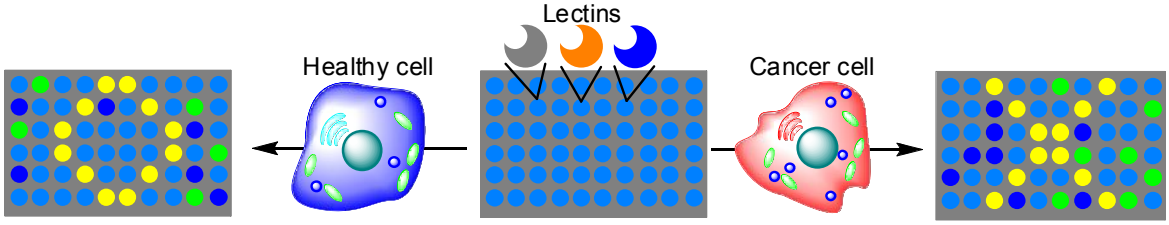


Figure 21 Pattern analysis of cell sample by lectin microarray

In fact, it is clear that the response pattern using the present lectin array is roughly classified into two categories, that is, the cell lines from the mammalian species and the cells from bacteria, by the bar graph in Fig.22a. This is supported by the cluster analysis of these intensity-based data as shown in Fig.22b,c.[15a] Among the mammalian cell lines, a subtle difference in the pattern was further observed, that is, MCF-7 cell showed the predominant response at AAL. It was supported by the previous report that there are richer fucosylated O-linked sugar chains in MCF-7 than in other cells.[16a] This pattern is apparently distinct from the response from the other five cell lines (L-6, HepG2, CHO, A549 and HeLa), and may be due to the major presence of the fucose component in MCF-7. As more details, the cluster analysis showed us that the L-6 is secondarily distinguished from the other four, due to the enhanced presence of α-galactose (GSL-I) as well as fucose (AAL) and sialic acid (WGA). This was consistent with the histological findings that the skeletal muscle cells such as L6 cell are effectively stained by GSL-I.[16b] Thus, it is clear that the response pattern of cell lysates obtained by using this lectin chip is in good agreement with previous general findings and the reports. However, the more subtle differences in other responses are not clearly assigned so far, because the high throughput and systematic analysis for glyco-conjugates has been poorly accomplished. On the other hand, a significant similarity between NM522 and JM109, both of which are bacterial strains, is observed in the cluster analysis of the response pattern. These results demonstrated that the present semi-wet lectin array is useful for pattern analyzing the cell lines on the basis of the characteristic difference in the sugar components of each cell.

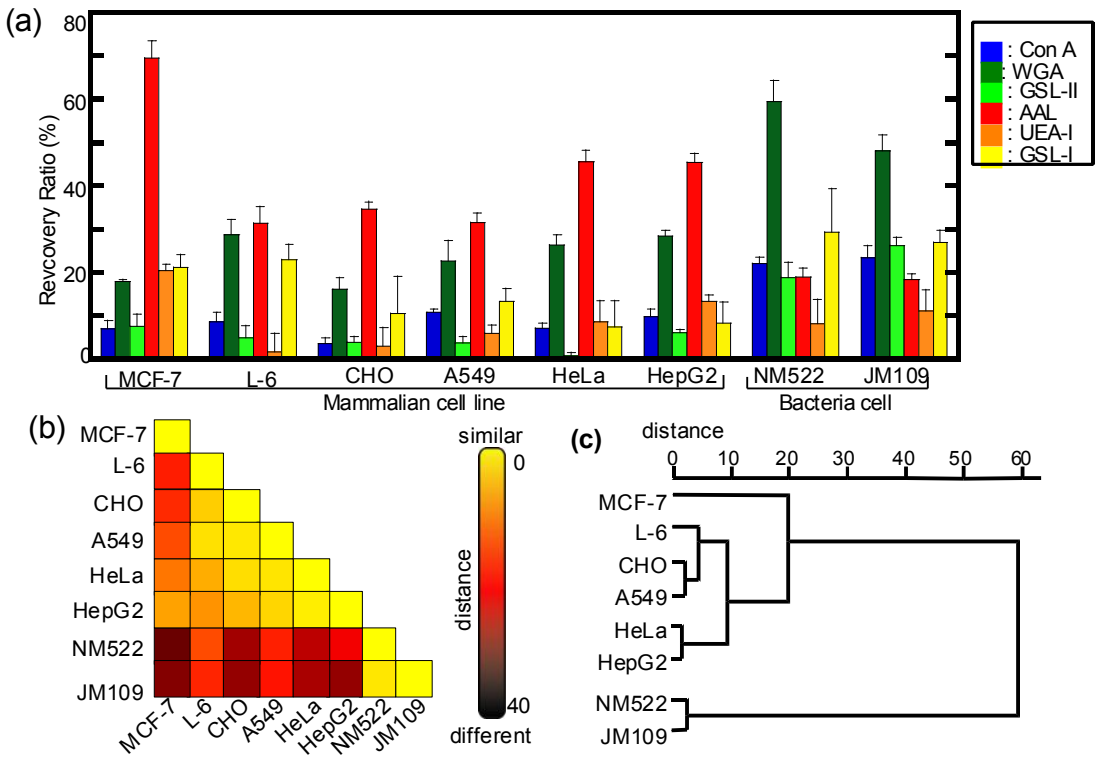


Figure 22. Pattern analysis of carbohydrates in cell lysates. (a) The bar graph of the fluorescent recovery ratio by the addition of mammalian cell lines and bacteria cells. The bar height and the error bar in the graph were estimated by the average and standard deviation of 4 spots on the same plate. (b) The

Euclidean distance matrix between the patterns of the different cell lines obtained by lectin chip was represented by color coding (yellow for the highest similarity and black for the lowest). (c) The dendrogram of the response patterns for eight cell lines generated by the analysis of Euclidean distances. The horizontal axis represents the distances among the normalized lectin chip patterns (left for patterns with the highest similarity and right for patterns with the lowest similarity).

3-3-2. The comparison of the present array with other lectin microarrays

Recently, some examples of the lectin microarrays to detect saccharides were reported (Table 3).[17-21] For example, Mahal and coworkers immobilized lectins on the glass slide covalently, to analyze fluorescent-labeled glycoproteins (Fig.23a).[17] They succeeded in the analysis of the bacteria glycans of bacterial cells into which a fluorescein dye was incorporated (Fig.23b). Hirabyashi and coworkers developed a lectin microarray using the evanescent field fluorescence as a detection method (Fig.24a).[18] By virtue of the evanescent field fluorescence, the fluorescence of fluorescent-labeled glycoproteins which bound to the lectin array can be selectively read out. Since there is no need to wash the array before detection, a highly sensitive assay was realized. Moreover, they successfully analyzed the sugar chain of mammalian cell surface by the incorporation of a fluorescein dye into mammalian cells (Fig.24b).

Table 3. The comparison of the lectin microarrays.

	Analysis object	Immobilization	Spot size	Number of lectins	Detection limit
This study	Mono-,Oligo-Saccharide, Glycoprotein, Lysate (Bacteria Mammalian)	Supramolecular hydrogel	4 mm	6	500 μg/mL
Hirabayashi et al.[17]	Glycoprotein, Lysate (Mammalian), Mammalian Cell	Epoxide	450 μm (220 μm)	45	0.01 μg/mL
Mahal et al.[18]	Glycoprotein, Lysate (Mammalian), Bacteria Cell	NHS-ester	95 μm	58	10 μg/mL
Sprenger et al.[19]	Glycoprotein	Biotin-Avidin	150 μm	4	-
Shin et al.[20]	Mono-Saccharide Mammalian Cell	NHS-ester	300 μm	5	-
Maya et al.[21]	Glycoprotein	Nitrocellulose	450 μm	24	1 μg/mL

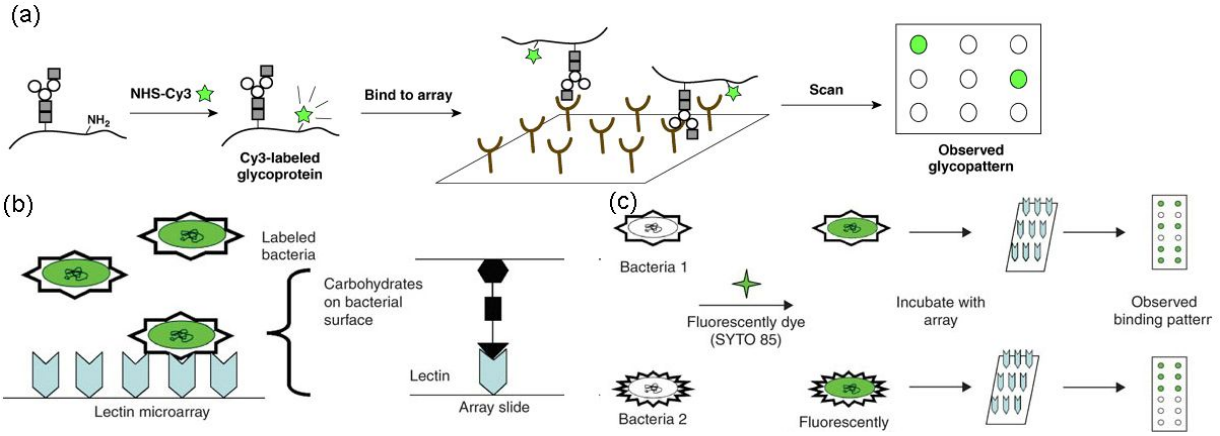


Figure 23. Fluorescent lectin microarray by Mahal et al. (a) Fluorescent detection of glycoprotein.[17a] (b,c) Fluorescent analysis of carbohydrate on bacteria surface.[17b]

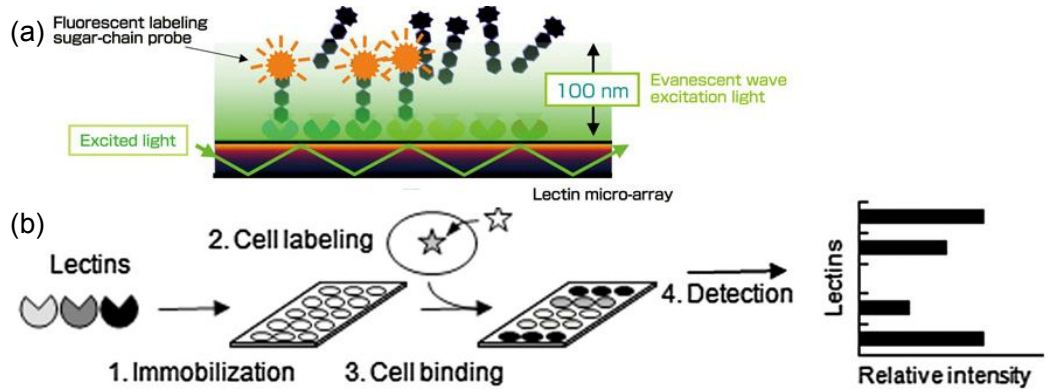


Figure 24. Evanescent field fluorescent lectin microarray by Hirabayashi et al. (a) Schematic illustration of fluorescent detection mechanism.[18a] (b) Fluorescent analysis of carbohydrate on mammalian cell. [18b]

Compared with these examples, it is the greatest advantage of my lectin microarray system that the fluorescent labeling of the samples is not required. This advantage enables not only to omit the labeling step but also to profile the cell lysates which includes various kinds of sugar derivatives. However, there are several disadvantages in this method. For example, the detection limit is suppressed because of usage of the competitive quencher. In fact, the detection limit of my system (500 μg/mL, Fig.25) is much higher than one of the Hirabasyashi's system (0.01 μg/mL, Table 2). But the detection sensitivity can be tuned by the concentration of the quencher. As shown in table 2, we successfully modulated the detection sensitivity of glucose from 5mM to 24mM by changing the concentration of quencher **2**. The other disadvantage is that the synthesis of the corresponding quenchers for each lectin and the subsequent optimization of the lectin/quencher pair are required for extending the arrayed lectins. But this disadvantage may not be so problematic, because the corresponding quenchers were easily synthesized by the reaction of **13** with various saccharides which have a reducing terminal using the aminooxy-method and the pair can be universally used for various analytes once it is optimized.

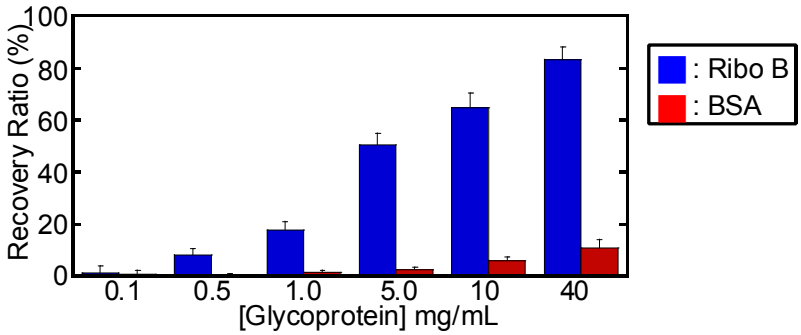


Figure 25. The bar graph of the recovery ratio with various concentration of Ribo B or BSA in Con A chip.

Table 4. The comparison of glucose detectable range by Con A chip to change the concentration of **2**.

quencher **2**	3.0 uM	4.0 uM	5.0 uM	7.5 uM	10 uM
RC_{50} (Glc)	5 mM	12 mM	15 mM	20 mM	24 mM

3-3-3. The immobilization method of lectin on array surface.

As mentioned in **3-1-2**, the immobilization method of proteins is important to construct protein microarray (Fig.26). Nonspecific interactions with the solid surface, such as nitrocellulose and PVDF, or random chemical modifications with the reactive group, such as epoxide and NHS-ester, are simple immobilization techniques of proteins. Through such nonspecific immobilization approaches do not required any modification, these do lack a defined orientation of proteins on the solid support. A site-specific immobilization technique using an affinity tag, such as the biotin moiety and GST, has been developed. This technique provided the attached proteins with the predetermined orientation on the microarray that suppressed the loss of their native conformations and activities. Recently, the chemical tag techniques such as azide were also reported. In addition, the modifications of polyethylene glycol (PEG) and dextran with the solid surface are known to suppress nonspecific adsorption of proteins.

In this study, I have noncovalently fixed fluorescent lectins by using supramolecular hydrogel. Under the semi-wet conditions provided by the supramolecular hydrogel, protein denaturation was effectively suppressed so that the embedded lectin acts as a talented molecular recognition scaffold toward specific saccharides. As shown in Fig.15, it was confirmed that the selectivities of lectins is perfectly retained in the supramolecular hydrogel matrix. However, it was concerned that the sugar head of gelator may interfere with the recognition of lectins. In GSL-I which can bind to GalNAc, the identical with the sugar head moiety of gelator **1**, the binding constant for melibiose in hydrogel matrix (K_a=1.1 X 10^4 M^{-1}) is almost comparable to one of previous report.[22] These result indicated that lectins dose not interact with the gelators because the gelators formed a stable and well-packed hydrogel fiber (Fig.27). Moreover, since many gelators bearing various sorts of saccharide head groups or non-saccharide head groups were developed in our group, it is possible to select a suitable gelator if such a interference was observed.[23]

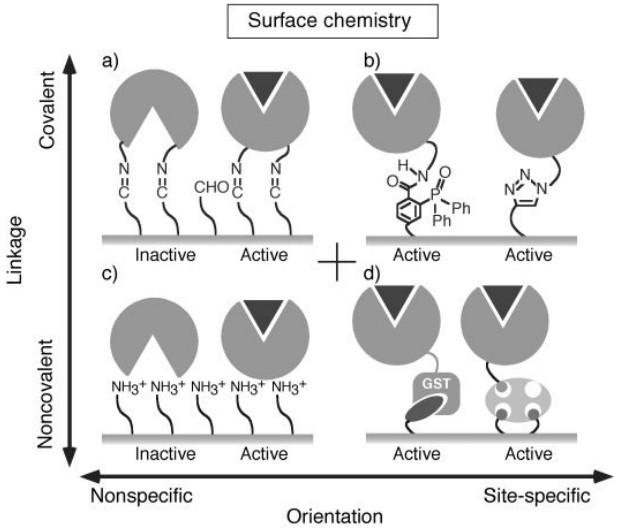


Figure 26. Attachment method of the capture agents [3a]

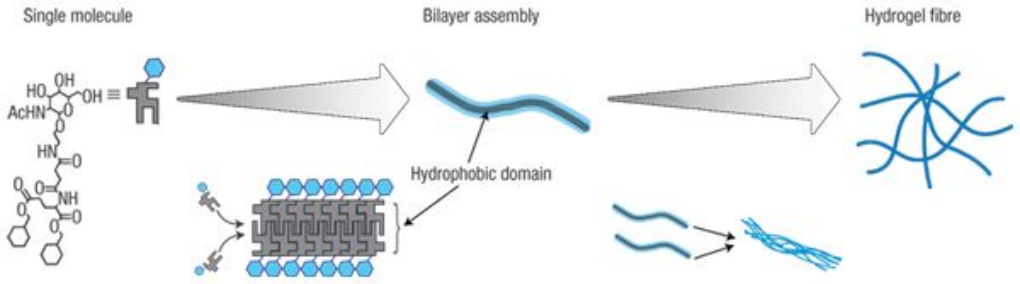


Figure 27. Schematic representation of the hierarchal molecular assembly of gelator **1** to form supramolecular hydrogel.[8a]

3-4. Conclusion

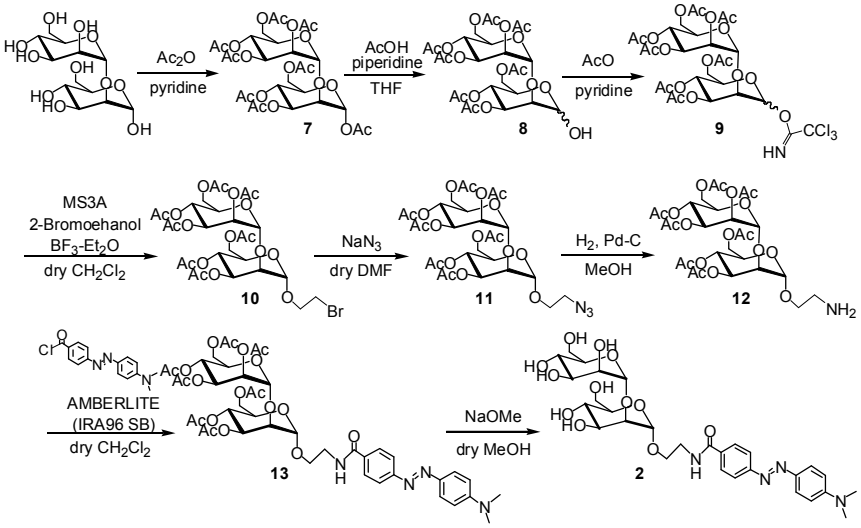
In this chapter, I successfully developed a novel semi-wet lectin microarray that the fluorescent labeling of analytes is not required by the transfer of a BFQR method from a solution system to a supramolecular hydrogel array system. Simple and high-throughput sensing and profiling were carried out using the lectin microarray for diverse glyco-conjugates, that not only included a simple glucose, but also oligosaccharides, and glycoproteins, and furthermore, the pattern recognition and profiling of several types of cell lysates were also accomplished. For the future, more detail analysis for glyco-patterns of various cells is expected by increasing the kinds of lectins immobilized on this lection array.

3-5. Experimental section

General comments for Synthesis.

All chemical reagents were purchased from commercial suppliers (Aldrich, TCI, or Wako) and used without further purification. All air- or moisture-sensitive reactions were performed with distilled solvents (DMF, dichloromethane, or MeOH) under a nitrogen or argon atmosphere.

Synthesis of Quencher 2



Compound (**7**):

α1,2-manbiose 100 mg (292 μmol) was dissolved in 5 mL of pridine 5 and 3mL of acetic

anhydride. The reaction mixture were stirred at for 2 days, evaporated under vacuum, and diluted with ethyl acetate. The organic layer was washed with 5% aqueous citric acid, aqueous $NaHCO_3$ and brine and dry over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **7** 170 mg (251 μmol, 86 %) as a white amorphous solid. ^1H NMR (400MHz, $CDCl_3$), δ /ppm 2.02-2.16 (m, 24H), 4.04-4.26 (m, 7H), 4.95 (d, 1H, J_H = 2.0 Hz), 5.25-5.47 (m, 5H), 6.25 (d, 1H, J_H = 240 Hz).

Compound (**8**):

To the solution of compound **7** 170 mg (250 μmol) in 3mL of THF was added 15 μL of acetic aced (250 μmol) and 250 μL of piperidine (2.5 mmol). The reaction mixtures were stirred at room temperature for 4 hour, evaporated under vacuum, and diluted with ethyl acetate. The organic layer was washed with aqueous $NaHCO_3$ and dry over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **8** 140 mg (220 μmol, 86 %) as a white amorphous solid. ^1H NMR (400MHz, $CDCl_3$), δ /ppm 2.02-2.16 (m, 21H), 3.04 (d, 1H, J_H = 3.6 Hz), 4.09-4.27 (m, 7H), 4.95 (s, 1H), 5.29-5.43 (m, 5H).

Compound (**9**):

To the solution of compound **8** (140 mg, 220 μmol) in 3 mL of CH_2Cl_2 were added 220 μL of trichloro acetonitrile (2.2 mmol) and cesium carbonate 7.2mg (22 μmol). The reaction mixtures were stirred at room temperature for 1 hour, and diluted 50 mL of CH_2Cl_2. The organic layer was washed with $NaHCO_3$.aq and brine, and dry over anhydrous $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **9** 150 mg (192 μmol, 87 %) as a white amorphous solid.

Compound (**10**):

2-Bromoethanol (15 μL, 211 μmol) and molecular sieves 3A were added to a solution of compound **9** (150mg, 192 μmol) in 30 mL of CH_2Cl_2. After stirring for 30 min at -40 ˚C, boron trifluoride-diethyl ether (15 μL, 211 μmo) was added to the above solution. The reaction mixture was then stirred for 1 h at -40 ˚C and diluted with 50 mL of CH_2Cl_2. The organic layer was washed with $NaHCO_3$.aq and brine and dried over anhydrous Na_2SO_4. The crude residue was purified by flash column chromatography on silica gel (elution: ethyl acetate/hexanes 1:1) to give **10** as a white solid (2.17g, 2.92 mmol, 39%). ^1H NMR (400MHz, $CDCl_3$), δ /ppm 2.02-2.16 (m, 21H), 3.52 (t, 2H, J_H = 5.6 Hz), 3.84-3.90 (m, 1H), 3.98-4.26 (m, 8H), 4.94 (d, 1H, J_H = 1.2 Hz), 5.02 (d, 1H J_H = 1.6 Hz), 5.26-5.43 (m, 5H).

Compound (**11**):

To the solution of compound **10** 80 mg (108 μmol) in 3 mL of DMF were added sodium azido 35 mg (538 μmol). The reaction mixtures were stirred at 80 ˚C for 6 hour, evaporated under vacuum, and diluted 50 mL of ethyl acetate. The organic layer was washed with water and dried over $MgSO_4$. The solution was filtered and concentrated under vacuum to afford **11** 75 mg (106 μmol,

98 %). ^{1}H NMR (400MHz, $CDCl_3$), δ /ppm 2.02-2.16 (m, 21H), 3.52 (t, 2H, J_H = 5.6 Hz), 3.64-3.70 (m, 1H), 3.87-3.92 (m, 1H), 3.96-4.27 (m, 7H), 4.94 (d, 1H, J_H = 1.6 Hz), 5.01 (d, 1H J_H = 1.6 Hz), 5.26-5.43 (m, 5H).

Conmpound (**12**):

To a solution of **11** 75 mg (106 μmol) in 3 mL of methanol was added palladium-carbon (10%, 10 mg). Under H_2 atmosphere, the reaction mixture was stirred for 3 h at room temperature. After the palladium-carbon was removed by filtration, the filtrate was evaporated under vacuum to afford **12** as a white solid 65 mg (95.6 mmol, 91%). ^{1}H NMR (400MHz, $CDCl_3$), δ /ppm 2.02-2.16 (m, 21H), 3.33 (m, 2H), 3.85 (m, 1H), 4.01-4.25 (m, 8H), 5.04 (s, 1H), 5.28-5.41 (m, 5H).

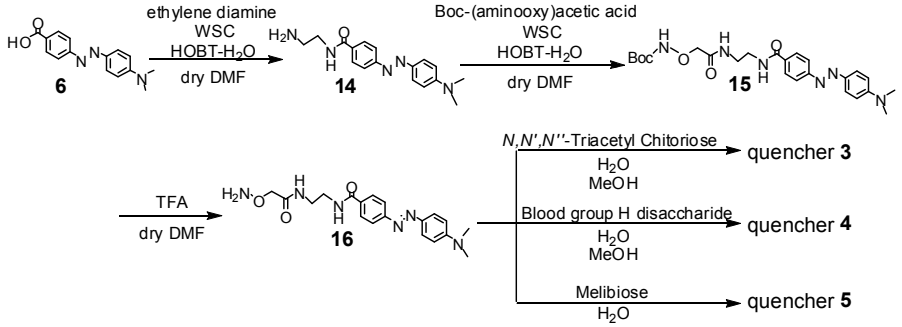
Compound (**13**):

To the solution of **12** 65 mg (95.6 μmol) was in 3 mL of CH_2Cl_2 were added AMBERLITE IRA96SB 500mg and 4-(dimetylamino)-azobenzene-4'-carbonyl chloride 42mg (146 μmol). The solution was stirred at room temperature for 2hr and evaporated under vacuum. The crude residue was purified by flash column chromatography on silica gel (elution: chloroform/methanol = 20/1) to give **13** as a red solid 65 mg (69.8 μmol, 73 %). ^{1}H NMR (400MHz, $CDCl_3$), δ /ppm 2.01-2.15 (m, 21H), 2.80-3.20 (m, 6H), 3.40-4.30 (m, 11H), 4.93 (s, 1H), 5.02 (s, 1H), 5.20-5.45 (m, 5H), 6.75 (d, 1H J_H = 9.2 Hz), 7.75-8.00 (m, 6H). MALDI-Tof MS calcd for [M+H$^+$] ($C_{43}H_{54}N_4O_{19}$): 930.34; found: 931.3. Elemental Anal. Calcd ($C_{43}H_{54}N_4O_{19}$) : C,55.48; H, 5.85; N, 6.02; Found: C, 55.88; H, 5.86; N, 6.12.

Quencher (**2**):

To the solution of **13** 30 mg (32.2 μmol) in 3 mL of methanol were added 1 μL of methanol solution of sodium methoxide (66.2 μmol). The reaction mixtures were stirred at room temperature for 1hr, the reaction was quenched by adding ion exchange resin (Amberlite IRC-50). After filtration, the filtrate was removed under vacuum to afford **2** as a red solid 15 mg (23.6 μmol, 73 %). ^{1}H NMR (400MHz, CD_3OD), δ / ppm; 3.10 (s, 6H, N-$(CH_3)_2$), 3.56 (d, J_H=4.4 Hz, 2H, -CH_2-),3.58-3.98(m, 14H, H-2,3,4,5,6, H'-2,3,4,5,6, -CH_2-), 4.97 (d, J_H=1.6 Hz, 1H, H'-1), 5.16 (d, J_H=1.2 Hz, 1H, H-1), 6.82 (d, J_H=9.2 Hz, 2H, Ar-H), 7.83-7.86 (m, 4H, Ar-H), 7.95 (d, J_H=9.2 Hz, 2H, Ar-H). HRMS (FAB) calcd for [M+Na$^+$] ($C_{29}H_{40}N_4O_{12}Na$): 659.2540; found: 659.2545.

Synthesis of Quencher 3, 4, 5



Compound (**14**):

To the solution of 4-(dimetylamino)-azobenzene-4'-carboxylic acid **6** 495 mg (46.8 mmol) in 15 mL were added WSC 420 mg (2.19 mmol) and HOBt monohydrate 276 mg (1.8mmol). The reaction mixture were stirred at room temperature for 1 hour. The solution of ethylenediamine (151μmol) in 10 mL of DMF was added to the above solution in at 0 ˚C. The reaction mixture were stirred at 0 ˚C for 30 min, evaporated under vacuum, and diluted 100 mL of 5% aqueous citric acid. The aqueous layer was washed with ethyl acetate, and extracted to chloroform. The organic layer was dried over Na_2SO_4. The solution was filtered and concentrated under vacuum to afford **14** 492 mg (40.2 mmol, 86%) as a yellow powder. ^{1}H NMR (400MHz, $CDCl_3$), δ / ppm 2.90 (t, 2H, J_H = 6.2 Hz), 3.13 (s, 6H), 3.51 (t, 2H, J_H = 6.2 Hz), 6.86 (d, 2H, J_H = 9.6 Hz), 7.85-7.88 (m, 4H), 7.96 (d, 2H J_H = 8.8 Hz). MALDI-Tof MS calcd for [M+H$^+$] ($C_{17}H_{21}N_5O_9$): 311.4; found: 313.1.

Compound (**15**):

To the solution of **14** 447 mg (1.43 mmol) 30 mL of DMF were added WSC 408 mg (2.13 mmol), HOBt monohydrate 328 mg (2.14 mmol) and 500 μL of (2.87 mmol). The reaction mixture were stirred at room temperature for 58 h, evaporated to dryness and diluted in ethyl acetate. The organic layer was washed with 5 % aqueous citric acid, $NaHCO_3$.aq and brine and dried over Na_2SO_4. The crude residue was purified by flash column chromatography on silica gel (elution: c chloroform/methanol = 11/1) to (give compound **15** 65 mg (1.04 mmol, 73 %) as a yellow solid. ^{1}H-NMR (400MHz, $CDCl_3$), δ / ppm 1.46 (s, 9H), 3.10 (s, 6H), 3.60-3.66 (m, 4H), 4.36 (s. 2H), 6.76 (d, 2H, J_H = 8.8 Hz), 7.44 (s, 1H), 7.85-7.91 (m, 4H), 7.96 (d, 2H, J_H = 8.4 Hz), 8.03 (s, 1H), 8.48 (s, 1H). MALDI-Tof MS calcd for [M+H$^+$] ($C_{24}H_{32}N_6O_5$): 484.2; found: 484.2.

Compound (**16**):

Compound **15** 49 mg (100 μmol) was dissolved in 5mL of CH_2Cl_2 and 1 mL of TFA. The reaction mixtures were stirred at room temperature for 2 h, and evaporated under vacuum to afford **16** 70.3mg (quantative) as purple solid. ^{1}H NMR (400MHz, $CDCl_3$), δ / ppm 3.08 (s, 6H), 3.40 (d, 2H, J_H = 6.4 Hz), 3.48 (d, 2H, J_H = 6.4 Hz), 4.40 (s. 2H), 6.82 (d, 2H, J_H = 9.2 Hz), 7.73-7.80 (m, 4H), 7.85 (d, 2H, J_H = 8.8 Hz). FAB MS (LS) calcd for [M+H$^+$] ($C_{19}H_{24}N_6O_3$): 384.2; found: 385.2.

Quencher (**3**):

To the solution of compound **16** (42 mg, 68.6 μmol) in 600 μL of methanol were added *N,N',N"*-triacetyl chitriose 10 mg (15.9 μmol). The reaction mixtures were stirred at room temperature for 58 h and the solution was purified by reversed-phase HPLC. The eluent was lyophilized to give quencher **3** 6.3 mg (6.36 μmol, 40 %) as a orange solid. The RP-HPLC was performed using an ODS-A column (YMC, 250 × 10 mm) using a linear gradient of CH_3CN and 10 mM aqueous ammonium acetate from 5 / 95 to 65 / 35 over 60min: flow rate = 3 mL / min, detection at 432 nm. ^{1}H NMR (400MHz, CD_3OD), δ / ppm 1.94-2.00 (m, 9H, -NHAc), 3.11 (s, 6H, N-$(CH_3)_2$), 3.38-3.87 (m, 21H, H-3,4,5,6, H'-2,3,4,5,6, H"-2,3,4,5,6, -CH_2-CH_2-), 4.43 (d, J_H=8.8, 1H, H"-1), 4,47 (s, 2H, -CH_2-), 4.53 (d, J_H=8.8, 1H, H'-1), 4.92 (t, J_H=6.0, 1H, H-2), 6.81 (d, J_H=9.2, 2H, Ar-H), 7.65 (d, J_H=5.6, 1H, H-1), 7.83-7.86 (m, 4H, Ar-H), 7.95 (d, J_H=8.8 Hz, 2H, Ar-H). HRMS (FAB) calcd for [M+H$^+$] ($C_{43}H_{64}N_9O_{18}$): 994.4369; found: 994.4369.

Quencher (**4**):

To the solution of compound **16** 36.1 mg (59.0 μmol) in 500 μL of methanol were added blood group H disaccharide 5 mg (15 μmol). The reaction mixtures were stirred at room temperature for 87 h, and evaporated under vacuum. The crude residue was purified by reversed-phase HPLC. The eluent was lyophilized to give quencher **5** 11.8 mg (quantative) as a orange solid. The RP-HPLC was performed using an ODS-A column (YMC, 250 × 20 mm) using a linear gradient of CH_3CN and 10 mM aqueous ammonium acetate from 25 / 75 to 50 / 50 over 20min: flow rate = 9 mL / min, detection at 432 nm. ^{1}H NMR (400MHz, CD_3OD), δ / ppm 1.16 (s, 3H, CH_3), 3.08 (s, 6H, N-$(CH_3)_2$), 3.47-4.05 (m, 13H, H-3,4,5,6, H'-2,3,4,5, -CH_2-CH_2-), 4.51 (s, 2H, -CH_2-), 4.63 (t, J_H = 7.2 Hz, 1H, H-2), 4.95 (d, J_H = 9.2 Hz, 2H, H'-1), 6.83 (d, J_H = 9.2 Hz, 2H, Ar-H), 7.72 (d, J_H = 7.6 Hz, 1H, H-1), 7.83-7.86 (m, 4H, Ar-H), 7.93 (d, J_H = 8.8 Hz, 2H, Ar-H). HRMS (FAB) calcd for [M+H$^+$] ($C_{31}H_{45}N_6O_{12}$): 693.3095; found: 693.3091.

Quencher (**6**):

The solution of **16** 23.4 mg (37.3 μmol) in 2 mL of water were added to melibiose 16.6 mg (48.6 μmol). The reaction mixtures were stirred at 35 ˚C for 40 h and evaporated under vacuum. The crude residue was purified by HPLC to give **6** 16.9 mg (21.6 μmol, 58 %) as a orange solid. The RP-HPLC was performed using an ODS-A column (YMC, 250 × 20 mm) using a linear gradient of CH_3CN and 10 mM aqueous ammonium acetate from 5 / 95 – 35 / 65 over 40min: flow rate = 9 mL / min, detection at 432 nm. ^{1}H NMR (400MHz, CD_3OD), δ / ppm 3.01 (s, 6H, N-$(CH_3)_2$), 3.38-3.87 (m, 15H, H-3,4,5,6, H-2,3,4,5,6, -CH_2-CH_2-), 4.22 (t, J_H=6.8 Hz, 1H, H-2), 4.41 (s, 2H, -CH_2-), 4.45 (d, J_H=6.8 Hz, 1H, H'-1), 6.75 (d, J_H=9.2 Hz, 2H, Ar-H), 7.54 (d, J_H=6.8 Hz, 1H, H-1), 7.74-7.77 (m, 4H, Ar-H), 7.84 (s, J_H=8.0 Hz, 2H, Ar-H). HRMS (FAB) calcd for [M+H$^+$] ($C_{31}H_{45}N_6O_{13}$): 709.3045; found: 709.3055.

General comments for biochemical assay.

Fluorescein-labeled lectins were purchased by Funakoshi or Seikagaku Corporation and dialyzed by 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$ and 0.1 M NaCl) for one day before usage. Quenchers were dissolved in methanol and diluted by distilled water. The molar absorbance coefficient of F-lectins (ε_{494nm} = 68,000 (pH 7.5)) and quenchers (ε_{428nm} = 30,000 (in MeOH)) were calculated by the coefficient of FITC and dabcyl group, and the concentration was determined by UV-visible spectrometer (Shimazu UV-2550).

Preparation of lectin chip.

The gel array was prepared according to a slight modification of the method reported previously by us. A suspension of 0.25 wt% gelator **1** in 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$ and 0.1 M NaCl) was heated to form a homogenous solution. A 10 μL portion of the hot solution was spotted on glass plate and incubated to complete gel gelation in a sealed box under the high humidity at room temperature for 1 hr. One microliter of the mixed solution of FITC -lectins and quenchers was then dropped onto each resultant hydrogel spot. After 30 min, an analyte solution was dropped onto each spot of lectin chip, and left for 1hr before the measurement.

Fluorescence measurement.

In aqueous solution, F-lectin solution was diluted to 0.1 μM by 50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$ and 0.1 M NaCl) and measured by fluorescence spectrometer (Perkin Elmer LS55). The slit widths for the excitation and emission were set to 15nm and 10nm. The excitation wavelength was 488 nm. In hydrogel matrix, the fluorescent spectral change of F-lectins in hydrogel was traced by multi channel photo detector (MCPD; Otsuka Electronics, excitation at 475 nm) at room.

Fluorescence life time study.

The fluorescence lifetime was measured with a fluorescence limited instrument (Hamamatsu photonics, streak camera, model C4334, optically coupled to a changecoupled-devaice (CCD) array detector). 0.5μM F-Con A in aqueous solution [50 mM HEPES buffer (pH 7.5, containing 1 mM $MnCl_2$, 1 mM $CaCl_2$ and 0.1 M NaCl)] or in hydrogel matrix (0.25 wt%, gelator**1**) as samples and N_2 laser (337nm) were used for excitation.

Fluorescence imaging of lectin chip.

The fluorescence images of the resulting lectin chip were collected by Molecular Imager FX pro (BioRad) under FITC-mode. In the case of saccharide and glycoprotein analysis, the stock solutions dissolved in distilled water were prepared by the measurement of the weight and added 1 uL or 2 μL to lectin chip. Bacteria cells were cultured in LB medium according to general methods, and destroyed by the ultrasonic homogenizer after washing by TES. Mammalian cells were cultured in DMEM (HepG2, L-6, A549), MEM (MCF-7, HeLa) or F12 (CHO) according to general

methods, and destroyed by the ultrasonic homogenizer after washing by PBS. The resultant lysate was 3 μL added to the lectin chip without any modifications. This lectin chip is stably stored for few days and produces the almost constant response pattern for at least 3days. Furthermore, we can convey these chips by car without problem.

The data-analysis of lectin chip.

The fluorescent recovery ratio was calculated by the fluorescence intensity obtained in the gel image of the corresponding lectin chip. The intensity ($I_{quenching}$) at the spots containing the quencher in the absence of analytes is set at 0 % and the intensity ($I_{recovery}$) at the spots in the presence of the large amount of corresponding saccharides (the saturated value of Me-α-Man for Con A, NAG2 for WGA and GSL-II, Fucose for AAL and UEA-I, Melibiose for GSL-I) is set at 100 %. Using these two values, the data was normalized. The recovery ratio of the analysis spots was calculated by comparing the obtained intensity ($I_{analyte}$) with these two values according to lower formula.

$$\text{Ratio (\%)} = (I_{analyte} - I_{quenching}) / (I_{recovery} - I_{quenching}) \text{ X } 100$$

Data treatment using statistical analysis[15]

Euclidean distance was used to determine the similarity between two lectin chips obtained from different cell lines. This is a common measure when considering the distance between two vectors. Before the Euclidean distance analyses were conducted, the lectin chips should be normalized. The normalization formula was as follows:

$$(I_{ori}\text{-}I_{ave})/d=I_{norm}$$

(I_{ori} : a value of a spot in a lectin array, I_{ave} : the average of all value of a lectin array, I_{norm} : a normalized value of a spot in a lectin array)

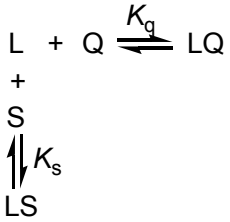
Similarity between the normalized lectin array patterns was measured by Euclidian distance in multidimensional space defined by each lectin array and is represented by color coding (yellow for the highest similarity and black for the lowest) the same as described in the previous section. Additionally, the cluster analysis among the normalized array was conducted. We used Ward's clustering algorithm, and the dendrogram was obtained with the analyses of Euclidean distances using Exel macro program. The horizontal axis represents the distances among normalized lectin array (left for Lectin chips with the highest similarity and right for Lectin chip with the lowest similarity).

Fluorescent merged imaging of Con A chip.

In the preparation of the lectin chip containing OR-B, the OR-B stock solution was added to the heated gelator solution before spotting, and other manipulation was the same as simple lectin chip. The fluorescence images were collected under FITC-mode and TAMRA-mode. Intensity ratio was calculated by the comparison with both fluorescent images.

Calculation of binding constants of Lectin

In BFQR, the competitive binding of quencher (Q) versus saccharide (S) to lectin (L) can be formalistically represented be eq **1**.



(1)

In eq 1, K_q and K_s are the association constants of the lectin-quencher (LQ) and lectin-saccharide (LS) complexes, respectively. If the free and total concentrations of L, Q, S are represented by [L], [Q] and [S], and $[L]_t$, $[Q]_t$ and $[S]_t$, respectively, K_q and K_s can be represented as follows.

$$K_q = [LQ]/[L][Q] = [LQ]/[L]([Q]_t-[LQ]) \quad (2)$$

$$K_s = [LS]/[L][S] = [LS]/[L]([S]_t-[LS]) \quad (3)$$

If $[L]_t << [S]_t$, then $[S]_t$ can be taken to be the measure of [S], and by taking into account the mass balance ($[L]_t = [L]+[LS]+[LQ]$), $[L]_t$ can be represented by eq 4.

$$[L]_t =[LS]+[LS]/K_s[S]_t+K_q[LS][Q]_t/(K_s[S]_t+K_q[LS]) \quad (4)$$

Equation 4 can be simplified to yield the dependence of [LS] on the concentrations and the binding constants of other species (eq 5).

$$[LS] = K_s[S]_t([L]_tK_q-1-K_s[S]_t-K_q[Q]_t+ ((-[L]_tK_q+1+K_s[S]_t+K_q[Q]_t)^2+4K_q[L]_t(K_s[S]_t+1))^{1/2})/2K_q(K_s[S]_t+1) \quad (5)$$

The binding constant of the lectin-saccharide complex determined by monitoring the fluorescent recovery from the quenching state (lectin-quencher complex). Hence, the fluorescence change (△F) as a function of the saccharide ([S]) concentration can be given by eq 6

$$\triangle F/\triangle F_{max} = [LS]/[L]_t \quad (6)$$

Whether $\triangle F_{max}$ is the maximum change in the fluorescence upon complete displacement of quencher form lectin.

3-6. Reference

1) (a) Schena, M.; Shalon, D.; Davis, R. W.; Brown, P. O. *Science,* **1995**, *270*, 467. (b) Lockhart, D. J.; Dong, H.; Byren, M. C.; Follecttie, M. T. Gallo, M. V.; Chee, M. S. Mittmann, M.; Wang, C.; Kobayashi, M.; Horton, H.; Brown, E. L. *Nature Biotechnol.* **1996**, *14*, 1675. (c) J. Inazawa, *Biotechnology Jounarl*, **2007**, *7*, 280-284.

2) (a) http://www.arrayit.com/ (b) http://www.affymetrix.com/technology/manufacturing/index.affx (c) http://www.gene-chips.com/

3) (a) Tomizaki, K.; Usui, K.; Mihara, H. *ChemBioChem*, **2005**, *6*, 782. (b) Hall, D. A.; Ptacek, J.; Snyder, M. *Mech. Ageing. Dev.* **2007**, *128*, 161-167.

4) Rusmini, F.; Zhong, Z.; Feijen, F. *Biomacromolecules* **2007**, *8*, 1775-1789.

5) (a) Shin, I.; Park, S.; Lee, M.-r. *Chem. Eur. J.* **2005**. *11*, 2894. (b) Park, S.; Lee, M.-r.; Pyo, S.-J.; Shin, I. *J. Am. Chem. Soc.* **2004**, *126*, 4812. (c) Park, S.; Shin, I. *Angew. Chem. Int. Ed.* **2002**, *41*, 3180. (d) Burn, M. A.; Disney, M. D.; Seeberger, P. H. *ChemBioChem,* **2006**, *7*. 421. (e) Diseny, M. D.; Seeberger, P. H. *Chem. Biol.* **2004**, *11*, 1701. (f) Adams, E. W.; Ratner, D. M.; Bokesch, H. R.; McMahon, J. B.; O'Keefe, B. R.; Seeberger, P. H. *Chem. Biol.* **2004**, *11*, 875. (g) Blixt, O. et al. *Proc. Nat. Acad. Sci.* **2004**, *101*, 17033. (h) Bryan, M. C.; Lee, L. V.; Wong, C.-H. *Bioorg. Med. Chem. Lett.* **2004,** *14*, 3185. (i) Nimrichter, L.; Gargir, A.; Gorltler, M.; Altstock, R. T.; Shtevi, A.; Weisshaus, O.; Fire, E.; Dotan, N.; Schnaar, R. L. *Glycobiology*, **2004**, *14*, 197. (j) Wang, D.; Liu, S.; Trummer, B. J.; Deng, C.; Wang, A. *Nat. Biotecnol.* **2002**, *20*, 275. (k) Willats, W. G. T.; Rasmussen, S. E.; Kristensen, T.; Mikkelsen, J. D.; Knox, J. P.; *Proteomics*, **2002**, *2*, 1666.

6) (a) Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357. (b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683. (c) Raman, R.; Raguram, S.; Venkataraman, G.; Paulson, J. C.; Sasisekharan, R. *Nat. Meth.* **2005**, *2*, 817.

7) (a) Medintz, I. L.; Clapp, A. R.; Mattoussi, H.; Goldman, E. R.; Fisher, B.; Mauro, J. M. *Nat. Mater.* **2003**, *2*, 630. (b) Medintz, I. L.; Goldman, E. R.; Lassman, M. E.; Mauro, J. M. *Bioconjugate Chem.* **2003**, *14*, 909. (c) Sandros, M. G.; Gao, D.; Benson, D. E. *J. Am. Chem. Soc.* **2005**, *127*, 12198.

8) (a) Kiyonaka, S.; Sada, K.; Yoshimura, I.; Shinkai, S.; Kato, N.; Hamachi, I. *Nat. Mater.* **2004**, *3*, 58. (b) Yoshimura, I.; Miyahara, Y.; Kasagi, N.; Yamane, H.; Ojida, A; Hamachi, I. *J. Am. Chem. Soc.* **2004**, *126*, 12204. (c) Yamaguchi, S.; Yoshimura, I.; Kohira, T.; Tamaru, S.; Hamachi, I. *J. Am. Chem. Soc.* **2005** *127*, 11835. (d) Kiyonaka, S.; Shinkai, S.; Hamachi, I. *Chem. Eur. J.* **2003**, *9*, 976. (e) Kiyonaka, S.; Sugiyasu, K.; Shinkai, S.; Hamachi, I. *J. Am. Chem. Soc.* **2002**, *124*, 10954. (f) Hamachi, I.; Kiyonaka, S.; Shinkai, S. *Chem. Commun.* **2000**, 1281.

9) Benerjee, A. B.; Swanson, M.; Roy, B. C.; Jia, X.; Haldar, M. K.; Mallik, S.; Srivastava, D. K. *J. Am. Chem. Soc.* **2004**, *126*, 10875.

10) (a) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy 2nd Ed.*; Kluwer Academic/Plenum: New York, 1999. (b) Tong, A. K.; Jockusch, S.; Li, Z.; Zhu, H.-R.; Akins, D. L.; Turro, N. J.; Ju, J. *J. Am. Chem. Soc.* **2001**, *123*, 12923.

11) (a) Schwartz, M. W.; Porte, D. Jr. *Science,* **2005**, *307*, 375. (b) Lowell, B. B.; Shulman, G. I.

Science, **20005**, *307*, 384.

12) (a) Williams, R. L.; Greene, S. M.; McPherson, A. *J. Biol. Chem.* **1987**, *262*, 16020. (b) Green, E. D.; Adelt,G.; Baenziger, J. U. *J. Biol. Chem.* **1988**, *263*, 18253. (c) Harvey, D. J.; Wing. D. R.; Kuster, B.; Wilson, I. B. H. *J. Am. Soc. Mass Spectrom.* **2000,** *11*, 564. (d) Karlsson, N. G.; Packer, N. H. *Anal. Biochem.* **2002**, *305,* 173.

13) (a) Angata, T.; Varki, A. *Chem. Rev.* **2002**, *102*, 439. (b) Becker, D. J.; Lowe, J. B. *Glycobiology*, **2003**, 13, 41.

14) (a) Lavigne, J. L.; Anslyn, E. V. *Angew. Chem. Int. Ed.* **2001**, *40*, 3118. (b) Wright, A. T.; Griffin, M. J.; Zhong, Z.; McCleskey, S. C.; Anslyn, E.V.; McDevitt, J. T. *Angew. Chem. Int. Ed.* **2005**, *44*, 6375. (c) Zhang, C.; Suslick, K. S. *J. Am. Chem. Soc.* **2005**, *127*, 11548. (d) Zhou, H.; Baldini, L.; Hong, J.; Wilson, A. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2006**, *128, 2421.* (e) Baldini, L.; Wilson, A. J.; Hong. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2004**, *126*, 5656. (f) Buryak, A.; Severin, K. *J. Am. Chem. Soc.* **2005**, *127*, 3700.

15) (a) Usui, K.; Ojima, T.; Tomizaki, K.; Mihara, H. *Nano Biotech.* **2005**, *1*, 191. (b) Usui, K.; Tomizaki, K.; Ohyama, T.; Nokihara, K.; Mihara, H. *Mol. Biosyst.,* **2006**, *2*, 113. (c) Tkahashi, M.; Nokihara, K.; Mihara, H. *Chem. Biol.* **2003**, *10*, 53.

16) (a) Muller, S.; Hanisch, F.-G. *J. Biol. Chem.* **2002**, *277*, 26103. (b) Christie, K. N.; Thomson, C. *J. Histochem. Cytochem.* **1989**, *37*, 1303.

17) (a) Pilobello, K. T.; Krishnamoorthy, L.; Slawek, D.; Mahal, L. K. *ChemBioChem* **2005**, *6* , 985 (b) Hsu, K.-L.; Pilobello, K. T.; Mahal, L. K. *Nat. Chem. Biol.* **2006**, *2*, 153. (c) Pilobello, K. T.; Slawek, D. E.; Maha, L. K. *Proc. Nat. Acad. Sci.* **200**, *104*, 11534–11539.

18) (a) Kuno, A; Uchiyama, N; Koseki-Kuno, S; Ebe, Y; Takashima, S; Yamada, M; Hirabayashi, J, *Nat. Meth.* **2005**, *2*, 851. (b) Tateno, H.; Uchiyama, N.; Kuno, A.; Togayachi, A.; Sato, T.; Narimatsu, H.; Hirabayashi, J. *Glycobiology*, **2007**, *17*, 1138-1146. (c) Ebe, Y.; Kuno, A.; Uchiyama, N.; Koseki-Kuno, S.; Yamada, M.; Sato, T.; Nishimatsu, H.; Hirabayashi, J.; *J. Biochem.* **2006**. *139*, 323.

19) Angeloni1,S.; Ridet, J. L; Kusy, N.; Gao, H.; Crevoisier, F.; Guinchard, S.; Kochhar, S.; Sigrist, H.; Sprenger, N. *Glycobiology* **2005**, *15*, 31-43.

20) Lee, M.; Park, S.; Shin, I. *Bioorg. Med. Chem. Lett.* **2005**, *16*, 5132-5235.

21) Rosenfeld, R.; Bangio, H.; Gerwig, G. J.; Rosenberg, R.; Aloni, R.; Cohen, Y.; Amor, Y.; Plaschkes, I.; Kamerling, J. P.; Maya, R. B.-Y. *J. Biochem. Biophys. Methods* **2007**, *70*, 415–426.

22) Goldstein, I. J.; Blake, D. A.; Ebisu, S.; Williams, T. J.; Murphy, L. A. *J. Bio. Chem.* **1981**, *256*, 3890.

23) (a) S. Matsumoto et al. submitted. (b) H. Komatsu et al. submitted.

List of Publications

Chapter 1 Luminescent Saccharide Biosensor Using Lanthanide-bound Lectin Labeled with Fluorescein
Yoichiro Koshi, Eiji Nakata and Itaru Hamachi
ChemBioChem **2005**, *6*, 1349 –1352.

Chapter 2 Target-Specific Chemical Acylation of Lectins by Ligand-Tethered DMAP Catalysts
Yoichiro Koshi, Eiji Nakata, Masayoshi Miyagawa, Shinya Tsukiji, Tomohisa Ogawa and Itaru Hamachi,
J. Am. Chem. Soc. **2008**, *130*, 245-251.

Chapter 3 A Fluorescent Lectin Array Using Supramolecular Hydrogel for Simple Detection and Pattern Profiling for Various Glycoconjugates
Yoichiro Koshi, Eiji Nakata, Hiroki Yamane and Itaru Hamachi
J. Am. Chem. Soc. **2006,** *128*, 10413-10422.

Reviews

1. Lectin Functionalization by Post-Photo Affinity Labeling Modification (P-PALM)
   Yoichiro Koshi, Eiji Nakata and Itaru Hamachi
   Trends in Glycoscience and Glycotechnology, **2007**, *19*, 121-131.

2. Development of biosensors based on organic chemistry for proteins
   ～Lectin based sugar sensing probe～
   Yoichiro Koshi and Itaru Hamachi
   Biotechnology Journal (Yodosha), **2007**, *6*, 687-692.

3. Protein Functionalization based on organic chemistry
   －Development of new biosensors from lectin protein—
   Yoichiro Koshi, Eiji Nakata and Itaru Hamachi
   Chemistry Today (Tokyo Kagaku Doujin), **2005**, *407*, 55-60

Other Publications

1. Double-modification of lectin using two distinct chemistries for fluorescent ratiometric sensing and imaging saccharides in test tube or in cell
   Eiji Nakata, Yoichiro Koshi, Erina Koga, Yoshiki Katayama and Itaru Hamachi,
   J. Am. Chem. Soc. **2005**, *127*, 13253-13261.

2. Design of a hybrid biosensor for enhanced phosphopeptide recognition based on a phosphoprotein binding domain coupled with a fluorescent chemosensor.
   Takahiro Anai, Eiji Nakata, Yoichiro Koshi, Akio Ojida and Itaru Hamachi
   J. Am. Chem. Soc. **2007**, *129*, 6232-6239

3. Affinity Labeling-Based Introduction of a reactive handle for Natural Protein Modification
   Haruto Wakabayashi, Masayoshi Miyagawa, Yoichiro Koshi, Yousuke Takaoka, Shinya Tsukiji and Itaru Hamachi
   Angew. Chme. Int. Ed. submitted

List of Presentation

International Symposium

1. Development of Lectin Chip Based on Supramolecular Hydrogel for High-thoughput Glycoside Analysis
   Yoichiro Koshi, Eiji Nakata, Hiroki Yamane, Shan-Lai Zhou and Itaru Hamachi
   The International 21st Century COE symposium of BINDEC Chemistry Network, Senri Hankyu Hotel, Oosaka, Japan, October 2005.

2. New method for lectin labeling using sugar tethered catalyst
   Yoichiro Koshi, Masayoshi Miyagawa, Shinya Tsukiji and Itaru Hamachi
   Annual Conference of the Society for Glycobiology, Boston Park Plaza Hotel, Massachusetts, U.S.A., November 2007.

Domestic Symposium

1. Construction of an enzyme-like lectin using P-PALM
   Yoichiro Koshi, Eiji Nakata, Hiroki Takemoto and Itaru Hamachi
   1st Joint Symposium on 18th Biofunctional chemistry and 7th Biotechnology, Kumamoto University, Kumamoto, Japan, October 2003.

2. Introduction of artificial molecular for lectin by P-PALM
   Youichirou Koshi, Eiji Nakata, Hiroki Takemoto and Itaru Hamachi
   84th Annual meeting of Chemical Society of Japan, Kwansei Gakuin University, Hyougo, Japan, March 2004.

3. Development of new P-PALM method using thioester-ketone chmiestry
   Youichirou Koshi, Eiji Nakata, Hiroki Takemoto and Itaru Hamachi
   Young scientist forum in 19th Symposium on Biofunctional chemistry, Tokyo Institute of Technology, Kanagawa, Japan, October 2004

4. Development of Saccharide Sensor Using LRET on Lectin
   Youichirou Koshi, Eiji Nakata and Itaru Hamachi
   19th Symposium on Biofunctional chemistry, Tokyo University, Tokyo, Japan, October 2004

5. Organic chemistry on protein surface (2):The lectin modification using by thio-ester and ketone chemistry
Yoichiro Koshi, Eiji Nakata, Hiroki Takemoto, Takahiro Anai and Itaru Hamachi
85th Annual meeting of Chemical Society of Japan, Kanagawa University, Kanagawa, Japan, March 2005.

6. Pattern Analysis for Glycoside using by Semiwet Lectin Chip
Yoichiro Koshi, Eiji Nakata, Hiroki Yamane, Zennrai Syu, Itaru Hamachi
20th Symposium on Biofunctional chemistry, Nagoya City University, Aichi, Japan, Septermber 2005.

7. Smart Biomaterials(4):Development of Lectin Chip using Hydogel for the Sugar Profiling
Yoichiro Koshi, Eiji Nakata, Hiroki Yamane and Itaru Hamachi
86th Annual meeting of Chemical Society of Japan, Nihon University, Chiba, Japan, March 2006.

8. Lectin functionalization by organic chemistry -development of glycol-biosensor using P-PALM-
Yoichiro Koshi, Eiji Nakata, Masayoshi Miyagawa and Itaru Hamachi
55th Symposium on Macromolecules, Toyama University, Toyama, Japan, September 2006.

9. A new strategy for protein molecular science(5): Specific chemical protein surface modification by ligand-directed acyl-transfer reaction
Yoichiro Koshi, Masayoshi Miyagawa and Itaru Hamachi
87th Annual meeting of Chemical Society of Japan, Kansai University, Oosaka, Japan, March 2007.

10. A new labeling method for lectins by sugar ligand-tethered DMAP
Yoichiro Koshi, Masayoshi Miyagawa, Shinya Tsukiji and Itaru Hamachi
20th Symposium on Biofunctional chemistry, Tohoku University, Miayagi, Japan, Septermber 2005.

List of Honors

Presentation award for students
86th Annual meeting of Chemical Society of Japan, Nihon University, Chiba, Japan, March 2006

Poster award
Young scientist forum in 19th Symposium on Biofunctional chemistry, Tokyo Institute of Technology, Kanagawa, Japan, October 2004