ChemPubSoc Europe

DOI: 10.1002/cplu.201300130

Synthesis of Cell-Penetrating *S*-Galactosyl–Oligoarginine Peptides as Inducers of Recombinant Protein Expression under the Control of *lac* Operator/Repressor Systems**

Yukina Mizuta, Akinori Takasu,* and Masahiro Higuchi[a]

The synthesis of glycodendrimers and glycopoly(oxazoline)s as inducers of recombinant protein expression has recently been reported; however, these compounds induced the expression of only small amounts of the green fluorescence protein (GFP), which was used as the model recombinant protein, because of their poor ability to penetrate the *Escherichia coli* cell membrane. Therefore, *S*-galactosyl–oligo(Arg) conjugates have now been synthesized to overcome this problem. Following in vivo expression of GFP induced by each of the *S*-galactosyl $(Arg)_n$ constructs ($n=5$, 6, 8) at the T5 promoter in *E. coli* for 18 hours, we visually observed that the cultures fluoresced green light when excited with UV light. The fluorescent intensities for these cultures were greater than that found for a control culture, which indicates that the peptides had induced GFP expression. Quantitative fluorescent measurements also supported the observations that the peptides were better inducers of GFP expression than the galactosyl dendrimers and the poly(oxazoline)s and the natural inducer lactose. Because the level of GFP expression was directly related to the number of arginine moieties in each peptide, we propose that the number of arginine moieties is responsible for how well each peptide passes through the *E. coli* membrane, which affects the expression level. A similar tendency was observed when the T7 promoter was placed upstream from the gene for an artificial extracellular matrix protein and the *S*-Gal–oligo(Arg) peptides were used as inducers. To assess how the distance between two galactosyl moieties as well as how the multivalent effect (cluster effect) in an oligo(Arg) inducer affects the expression level of GFP, we synthesized a conjugate of Lys-$(Arg)_8$ (Lys = lysine) and two *S*-galactosyls, which enhanced the expression of GFP in comparison with that obtained for *S*-Gal-$(Arg)_8$.

Introduction

The RNA polymerase of bacteriophage T5/T7 has a stringent specificity for its own promoters. In 1961, Jacob and Monod suggested the use of the *Escherichia coli* lactose (lac) operon as a model for gene regulation.[1] This model system is still used to study how a structural gene set can be coordinately transcribed or repressed, depending upon the metabolites found in the intercellular environment. The lac repressor, which is the protein product of *lacI* binds to the *lac* operator in the absence of an inducer, for example, the naturally occurring inducer lactose. When the repressor is bound to the operator, transcription of *lacZ*, *lacY*, and *lacA*, which encode β-galactosidase, lac permease, and a transacetylase, respectively, does not occur (Figure 1). The lac repressor promoter contains 360 amino acids and associates to form a homotetramer of 154520 Da.[2] Each monomer contains one saccharide binding site. Notably, isopropyl β-D-thiogalactoside (IPTG)—which is not a substrate for β-galactosidase, but a molecular mimic of allolactose derived from lactose, which may be the "true" inducer—acts as a gratuitous inducer and turns on transcription of the lactose operon through its interaction with the *lac* repressor.[3,4] IPTG permeates *E. coli* without the assistance of the *lac* permease.[5]

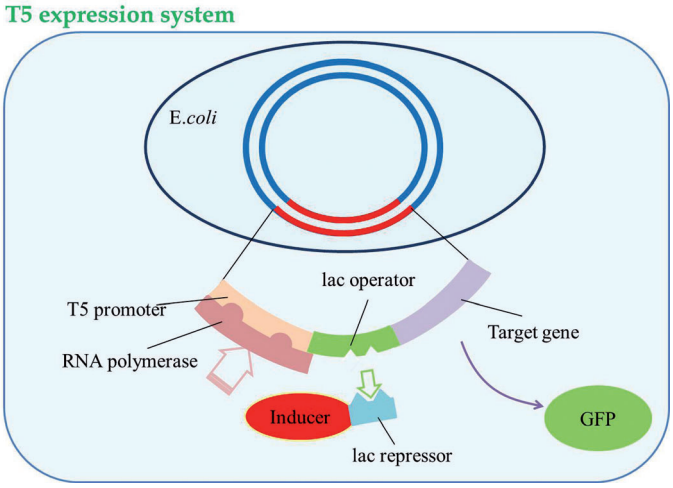


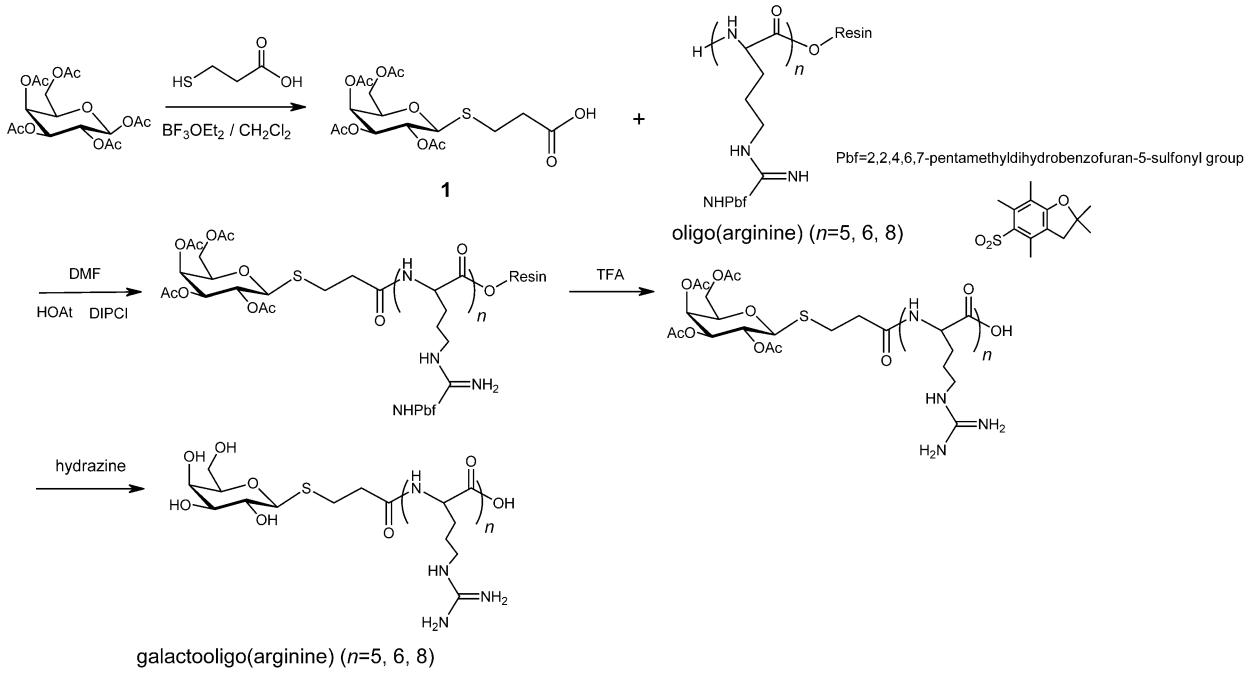
Figure 1. The T5 expression system under the control of a *lac* promoter/operator system.

Much attention is currently being devoted to glycobiology and glycochemistry topics in biomedicine and biochemistry,[6–14] especially as they relate to interactions between pro-

[a] *Y. Mizuta, Prof. A. Takasu, M. Higuchi*
Department of Frontier Materials
Nagoya Institute of Technology
Gokiso-Cho, Showa-Ku
Nagoya 466-8555 (Japan)
Fax: (+81) 52-735-5342
E-mail: takasu.akinori@nitech.ac.jp

[**] *lac* = *lactose*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cplu.201300130.



Scheme 1. Scheme for *S*-Gal–oligo(Arg) peptide synthesis (DIPCI = diisopropylcarbodiimide, HOAt = 1-hydroxybenzotriazole).

teins and carbohydrates. Carbohydrate–protein[6–12] and carbohydrate–carbohydrate[13,14] interactions have been well characterized, and, notably, such interactions are often strengthened by the presence of multiple binding sites. Certain glycopolymers in which multiple saccharide residues are incorporated into their polymer backbones have enhanced binding affinities toward their targeted proteins relative to their monomeric, saccharide-containing building block. This property has been ascribed to multivalent recognition, that is, the cluster effect.[6–12] Notably, certain glycopolymers, in which the saccharide spacing is random, strongly bind their target proteins.[6–12] Aoi et al.[15,16] used a dendrimer skeleton and Matsuura et al.[17,18] reported alternative strategies to prepare periodic glycosylated oligonucleotides (20-mers) as a means of controlling the three-dimensional arrangements of the pendant saccharides. These carbohydrate-containing compounds bound strongly to certain lectins. We have also described the binding affinities of α-helical peptides that contained a pendant saccharide linked to the peptide backbone through an *O*-glycoside linkage[19] (a model for mucine-type glycoproteins and glycopeptides) or through an *N*-glycoside linkage.[20,21] Recent progress involving glycoconjugates for biomedical applications, for example, drug-delivery systems, antibacterial activity, inhibition of viral infection, and the spread of malignant tumors and the human immunodeficiency virus has been remarkable with many excellent results reported.[6]

As far as we know, there have been only a few studies concerning the design of new glycoconjugates that can act as inducers of recombinant protein expression under the control of a *lac* promoter/operator/repressor system.[22,23] Furthermore, no reports have focused on the cluster effect of saccharides and allosteric effects of the repressor proteins until now, as far as we are aware. Recently, we synthesized *S*-galactosyl dendrimers[22] and *S*-galactosyl poly(oxazoline)s,[23] which we expected to act as inducers of recombinant protein synthesis through the cluster effect; however, their limited ability to permeate the *E. coli* membrane precludes their use as inducers. For this study, we designed *S*-galactosyl–oligo(arginine) (*S*-Gal–oligo(Arg)) conjugates (Scheme 1), including one containing the "magic arginine number" eight (R8),[24–26] to use as new types of *lac* operon inducers.

Results and Discussion

We synthesized the *S*-Gal–$(Arg)_n$ conjugates ($n = 5$, 6, or 8) using solid-phase methods (Scheme 1). However, compound **1**, which contains a protected galactoside moiety, was synthesized from penta-*O*-acetyl-β-D-galactopyranoside and 3-mercaptopropionic acid as reported previously.[22,23] The yield was 91%, and its structure was confirmed by ^{1}H NMR spectroscopy. After preparing the three protected arginine peptides through solid-phase fluorenylmethyloxycarbonyl chloride (Fmoc) chemistry, **1** was coupled to each of the resin-bound peptides through their N termini. *O*-Acetyl deprotection of the galactosyl hydroxyl groups was accomplished using hydrazine monohydrate, following the release of the glycopeptides from the resin and the deprotection of the arginine side chains by trifluoroacetic acid. The yields of the peptides were 60–69%. ^{1}H NMR spectroscopy (200 MHz; Figure S1 in the Supporting Information) and MALDI-TOF mass spectrometry (Figure S2) showed the expected structure and that the configuration of the anomeric carbon was β in all cases. The deacetylated products were completely water soluble and their conformations, as indicated by their room-temperature CD spectra, which con-

tained distinct minimums at 197 nm and less pronounced minimums at 222 nm (Figure S3), were primarily aperiodic structures with possibly a small amount (≈5%) of α helices (α helices increase the cell-membrane permeability of poly(Arg) and oligo(Arg) constructs[26]).

Assays that relied on green fluorescent protein (GFP) fluorescence in *E. coli* cells containing pQE9-GFP2/pREP4 were performed to evaluate the abilities of the *S*-Gal–oligo(Arg) peptides to act as inducers of GFP expression. For each assay, when the optical density ($OD_{600\ nm}$) of a culture (5 mL) was 0.6, one of the *S*-Gal–oligo(Arg) peptides or IPTG (each at 1 mM) was added. After an 18 hours induction, the visible fluorescence of each culture was observed (irradiation at 365 nm, Figure 2), quantified by fluorescence spectrometry (excitation at 395 nm and emission at 509 nm), and normalized to the $OD_{600\ nm}$ of each culture ($F/OD_{600\ nm}$; Figure 2).[27,28]

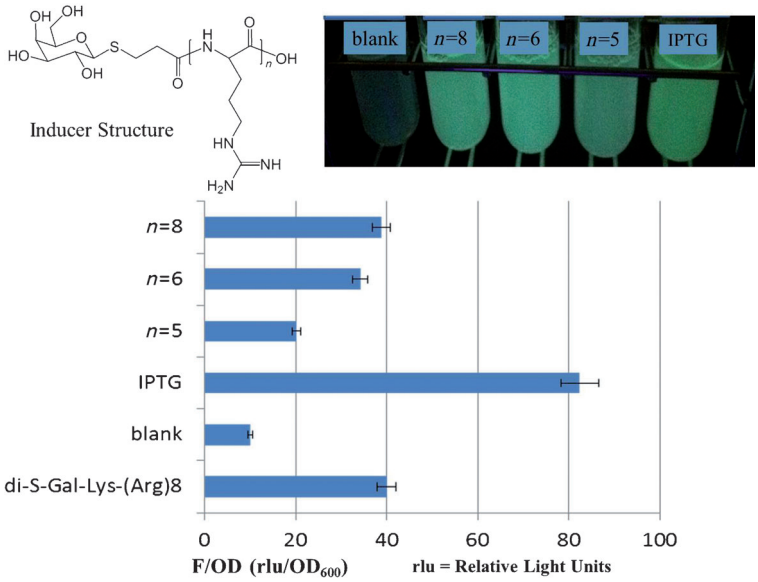


Figure 2. Normalized fluorescent intensities ($F/OD_{600\ nm}$) of *E. coli* cultures after induction with IPTG or an *S*-Gal–oligo(Arg) peptide.

As expected, the normalized fluorescence of the culture was relatively large ($F/OD_{600}=64$) when IPTG was the inducer. When using the *S*-Gal–oligo(Arg) peptides as inducers, fluorescence at 509 nm was detected (F/OD_{600} between 16 and 24), which is smaller than that found for IPTG, although larger than that for lactose ($F/OD_{600}=5$). The fluorescent measurements also showed that the *S*-Gal–oligo(Arg) peptides were better inducers than were the galactosyl dendrimers ($F/OD_{600}=6$–10)[22] and the poly(oxazoline)s ($F/OD_{600}=6$–11).[23] Because GFP expression increased as the number of arginine moieties in the *S*-Gal–oligo(Arg) peptides increased, it seems that an increase in the number of arginine moieties improved their passage through the *E. coli* cell membrane. Induction by deacetylated **1**, prepared by reaction of **1** and hydrazine monohydrate (20:1 molar ratio, hydrazine/galactosyl moiety) was also performed under the same conditions. The F/OD_{600} of 14 was lower than that induced by *S*-Gal–oligo$(Arg)_n$ ($n=8$; $F/OD_{600}=24$), which also supported the theory that the cell-penetrating peptide is a key moiety. However, since oligoarginine peptides are well known to enhance gene transfection efficiency in various cell lines,[24–26] we now had to eliminate the possibility that oligoarginine peptides can enhance protein expression in the *E. coli* cell. The model peptides containing no galactosyl moiety were prepared ($n=8$) and used as the inducer for the expression of GFP under the same conditions. After expression for 18 h, the green fluorescence was not detected ($F/OD_{600}=5$) and we concluded that there is the possible synergistic effect of the galactosyl moiety and the oligoarginine peptides as we expected.

Next, we used a pET expression system that included the gene for an artificial extracellular matrix (ECM) protein (aECM-CS5-ELF-F) that contains a fibronectin cell-binding domain (CS5) and an elastin sequence (ELF).[29,30] To assess the ability of the *S*-Gal–oligo(Arg) peptides to act as inducers, aECM-CS5-ELF-F[29,30] has been expressed in *E. coli* using a T7 expression system. We have also examined how aECM-CS5-ELF-F mutants, containing phenylalanine (Phe) analogues instead of Phe, affect the adhesive behavior of human umbilical vein endothelial cells (HUVEC).[30] The T7 promoter and its downstream target gene reside in a pET plasmid. Residual T7 polymerase activity is inhibited by T7 lysozyme, which is constitutively expressed at low levels, and for which its gene is present in either the pLysS or pLysE plasmid (Lys = lysine; Figure S4).[31,32] The pET system can produce large amounts of protein (up to 100 mg L^{-1} of culture medium). For this study, the *S*-Gal–oligo(Arg) peptides and IPTG was used as inducers of aECM-CS5-ELF-F expression[30] that was controlled by a T7 promoter upstream from a *lac* operator. Whole cell lysates and purified aECM-CS5-ELF-F were subjected to sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE). Substantial amounts of a protein with a calculated molecular weight of 42 600 that corresponds to aECM-CS5-ELF-F were seen, which indicates that the *lac* operator was activated by the *S*-Gal–oligo(Arg) peptides in the T7 expression system. The expressed protein concentrations were greater than that of the 10 mg L^{-1} culture and the yields are summarized in Figure 3.

The expression levels of aECM-CS5-ELF-F were smaller (20–39 mg L^{-1}) when the *S*-Gal–oligo(Arg) peptides were used as inducers, than when IPTG served as the inducer (82 mg L^{-1}), but greater than that found when lactose was used as the inducer (10 mg L^{-1}). Because aECM-CS5-ELF-F expression increased as the number of arginine moieties in the *S*-Gal–oligo(Arg) peptides increased, an increase in the number of arginine moieties apparently improved their passage through the *E. coli* cell membrane as had been shown for the GFP/T5 expression system.

The 600 MHz ^{1}H NMR spectra of the proteins purified from the T7 expression system using *S*-Gal$(Arg)_8$ or IPTG as the inducer are the same, which indicates that aECM-CS5-ELF-F[30,33] had been expressed in both systems (Figure 4, lower panel). Lower critical solution temperature (LCST) values were also

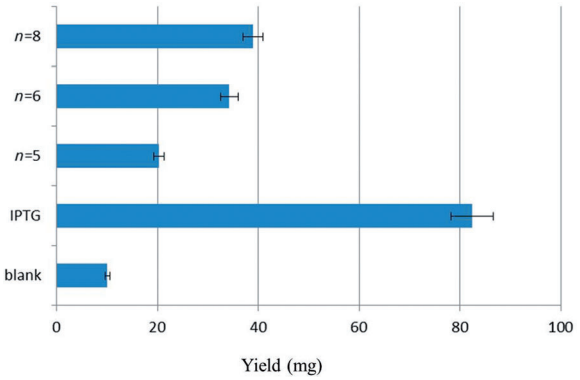


Figure 3. Yield of aECM-CS5-ELF-F expressed in 1 L of an *E. coli* culture after induction with IPTG or an *S*-Gal–oligo(Arg) peptide from a T7 expression system under *lac* repressor control.

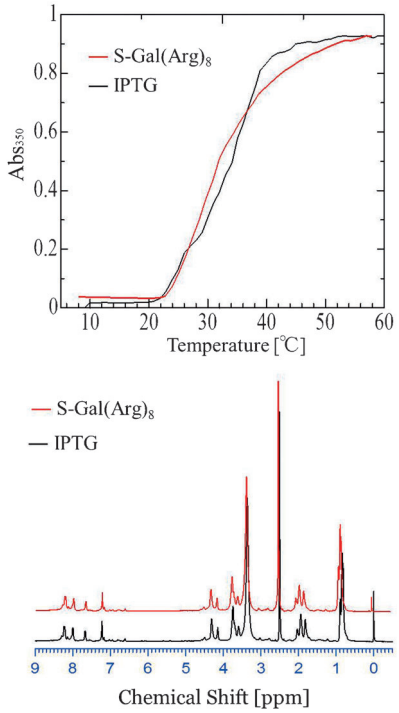


Figure 4. Above: LCST measurements (protein concentration, 1 mg mL^{-1} in H_2O). Below: 600 MHz ^{1}H NMR spectra of proteins expressed after addition of IPTG (black) or *S*-Gal$(Arg)_8$ (red) (solvent, [D_6]DMSO).

measured using 1 mg of aqueous expressed protein per milliliter. The temperature was increased at a rate of 1 °C min^{-1} while monitoring the percent transmission at 350 nm. A property that distinguishes elastin-like polypeptides from other types of polypeptides is that they are miscible at all concentrations in water below their LCST value.[29,30,33–38] For the protein preparations both LCST values were 34 °C (Figure 4, upper panel).

Finally, to assess how the distance between and the presence of two galactosyl moieties in an oligo(Arg) inducer as well as the multivalent effect (cluster effect) affect the expression levels of GFP, we synthesized a conjugate of Lys$(Arg)_8$ and **1** (Scheme 2). This inducer showed the expression of GFP ($F/OD_{600} = 40$; see also Figure 2, bottom), which is higher than that obtained for *S*-Gal–oligo$(Arg)_8$ ($F/OD_{600} = 24$). The X-ray structure of the *lac* repressor[3] shows that the distance between two sugar-binding sites is approximately 2.4 nm. The intersugar distance (between the C4 carbon atoms in the two galactosyl moieties) in a Corey–Pauling–Koltun (CPK) model of the conjugate would be between 0.8 and 2.2 nm depending on the conformation. Therefore, to better estimate the distance between the two galactosyl moieties, an energy minimization using a semiempirical molecular orbital method[39] was performed for a derivative of Lys$(Arg)_3$ that has **1** covalently attached to the N terminal and the Nε of the lysine, Gal-S-$(CH_2)_2$CONLys[NCO$(CH_2)_2$-S-Gal]$(Arg)_3$-COOH. After energy minimization, the intersugar distance, as defined above, was 1.0 nm (Figure S5), which is shorter than the 2.4 nm separation found for the *lac* repressor binding sites,[3] thereby indicating that the enhanced expression level induced by the *S*-diGal-$(Arg)_8$ peptide can be ascribed, in part, to the cluster (multivalent) effect. It seems that the abilities of the peptides synthesized for this study to induce protein expression mainly depend on their ability to penetrate *E. coli*.

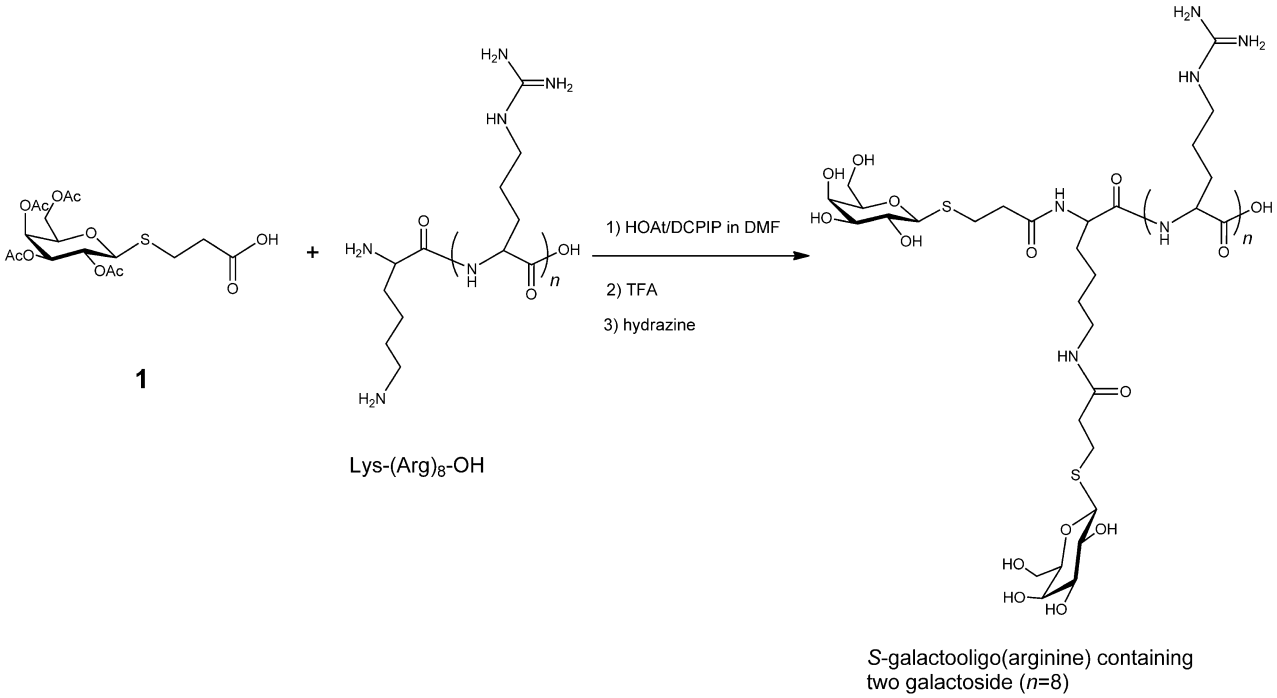
Conclusion

For the study reported herein, we prepared a set of *S*-Gal–oligo(Arg) peptides and a di-*S*-Gal-Lys-$(Arg)_8$ peptide using solid-phase methods that represent a new type of inducer for protein expression under the control of a *lac* operator/promoter/repressor system. Excellent galactosyl functionalities (coupling yield > 99%) were obtained. Their structures were verified by NMR spectroscopy and MALDI-TOF mass spectrometry. The abilities of these compounds to act as inducers were verified by the fluorescence associated with GFP expression in *E. coli* by using a T5 promoter. A similar tendency was observed when the T7 promoter was placed upstream from the gene for an artificial extracellular matrix protein. These fundamental results provide new possibilities for a recombinant protein expression system that is better than IPTG, although they are not yet as effective as IPTG.

Experimental Section

Materials

Pentaacetyl β-D-galactose was purchased from Sigma–Aldrich Corp. 3-Mercaptopropionic acid was purchased from Wako Corp. Fmoc/2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf)-protected Arg-Clear acid resin and Fmoc-Arg(Pbf) were purchased from the Peptide Institute, Inc. 1-Hydroxybenzotriazole (HOAt) and *N*,*N*′-diisopropylcarbodiimide (DIPCI) were purchased from Watanabe Chemical Industries. Piperidine, trifluoroacetic acid, hydrazine monohydrate, boron trifluoride diethyl ether complex, THF, and all other reagents were analytical grade and acquired from Peptide Institute Inc., Nacalai Tesque, Wako, and Sigma-Aldrich; these compounds were used without further purification.



Scheme 2. Strategy for the synthesis of Gal-S-$(CH_2)_2$CONLys[NCO$(CH_2)_2$-S-Gal](Arg)$_8$-COOH.

Measurements

^{1}H and ^{13}C NMR spectra were recorded at 27 °C using a Bruker DPX200 spectrometer (200 MHz for ^{1}H and 50 MHz for ^{13}C) or a Bruker DPX-600 spectrometer (600 MHz for ^{1}H). Chemical shifts were referenced to the tetramethylsilane signal ($\delta = 0$ ppm). MALDI-TOF mass spectra were recorded using a JMS-S3000 mass spectrometer (JEOL), with dithranol as the matrix agent. CF_3COONa was used to generate monosodium-cationized ions of the oligo-(Arg) derivatives ($[M+Na]^+$). CD spectra were measured using 0.02 mg mL^{-1} aqueous samples of the *S*-Gal–oligo(Arg) peptides, a 1 mm quartz cell, and a Jasco J820 K spectropolarimeter. LCST values were measured for 1 mg mL^{-1} aqueous samples of the recombinant, artificial ECM protein aECM-CS5-ELF-F that contains a fibronectin cell-binding domain (CS5) and an elastin sequence (ELF).[29,30] The temperature was increased at a rate of 1 °C min^{-1} while measuring the percent transmission of a sample at 350 nm with a V-550 spectrometer.

Semiempirical molecular orbital calculation

A model of Gal-S-$(CH_2)_3$CONLys[NCO$(CH_2)_3$-S-Gal](Arg)$_3$-COOH was subjected to a semiempirical molecular orbital calculation by the AM1 molecular orbital (MO) method[27] in MOPAC 2000 to investigate the preferred distance between the two galactosyl moieties.[40] As a starting conformation, the backbone of the (Arg)$_3$ chain was modeled as a planar zigzag structure.

Preparation of 3-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosylthio)propionic acid (1)

Compound **1** was prepared (91% yield) as described (Scheme 1).[22] ^{1}H NMR ($CDCl_3$): $\delta = 1.99$, 2.06, 2.16, 2.17 (4 s, 12 H; $COCH_3$), 2.77 (t, $J = 6.7$ Hz, 2 H; SCH_2CH_2COO), 2.96 (t, $J = 6.0$ Hz, 2 H; SCH_2CH_2COO), 3.91 (t, $J = 6.1$ Hz, 1 H; H-5), 3.99–4.28 (m, 2 H; H-6), 4.55 (d, $J = 9.9$ Hz, 1 H; H-1), 5.05 (dd, $J = 3.3$, 9.8 Hz, 1 H; H-3), 5.23 (t, $J = 9.9$ Hz, 1 H; H-2), 5.43 ppm (d, $J = 3.1$ Hz, 1 H; H-4).

Preparation of *S*-Gal–oligo(Arg) peptides

Pbf-protected (Arg)$_n$ peptides ($n = 5$, 6, 8) were first prepared by solid-phase synthesis using an EYELA Solid Organic Synthesizer CCS-150M, Fmoc-(Pbf)Arg-resin (0.50 g, 0.0176 mol), and Fmoc chemistry. Next, **1** (0.230 g, 0.528 mmol, 3 equiv) was coupled to the N terminal of each peptide while it was still bound to the resin by Fmoc chemistry. Removal of the acetylated *S*-Gal–oligo(Arg) preparations from the resin and deprotection of the arginine side chains were performed by adding ice-cold trifluoroacetic acid (28.5 mL) and H_2O (1.5 mL) to each ice-cold peptide solution and then each mixture was stirred at room temperature for 1.5 h. Each mixture was then filtered, and each filtrate was subjected to reduced pressure by rotary evaporation to remove any remaining trifluoroacetic acid. Next, each filtrate was poured into cold diethyl ether (200 mL) to precipitate the product, which was, in all cases, a white powder. Each ether/product mixture was centrifuged to isolate the product, which was then dried. To remove the acetyl groups, the acetyl-protected *S*-Gal–oligo(Arg) peptides were individually dissolved in THF containing hydrazine monohydrate (20:1 molar ratio, hydrazine/galactosyl moiety). After the solutions had been left standing at 0 °C for 16 h, acetone was added into each solution to quench the reaction. The solvent was completely removed under vacuum, and the residue was dissolved in H_2O. Each preparation was dialyzed against water (2 L) in seamless cellulose tubing (Spectra/Por Biotech Cellulose Ester, cut-off molecular weight of 1000) for 3 days, with the dialysate replaced every 12 h. For each preparation, a white powder was obtained upon lyophilization. The yields were 60–69%.

Expression of GFP in *E. coli* by a T5 expression system under the control of a *lac* promoter/operator/repressor system

E. coli K10 cells were transformed with a pQE9-GFP2/pREP4 vector containing the gene for GFP for which expression was under the control of a *lac* promoter/operator/repressor system.[22,23] The cells were cultured at 37 °C in M9 medium (5 mL) supplemented with 0.2% (w/v) glucose, 35 mg L^{-1} thiamine, 0.1 mM $MgSO_4$, 0.1 mM $CaCl_2$, the 20 common amino acids (4 mg mL^{-1} each), 25 mg L^{-1} kanamycin, and 20 mg L^{-1} chloramphenicol to an $OD_{600\,nm}$ of 0.6, after which the culture was divided into aliquots for the expression experiments. GFP expression was assessed at 37 °C by adding one of the *S*-Gal–oligo(Arg) peptides, IPTG, or lactose into a culture (each at final concentration of 1 mM) and incubating the cultures for 4 h, after which the fluorescence of each was visualized by UV-light irradiation (395 nm). A culture to which no peptide or inducer had been added served as the negative control. To quantify expression levels, the fluorescence of each culture was measured at 509 nm (excitation at 395 nm) using a Hitachi F-2700 spectrometer. The fluorescence (*F*) of each culture, expressed as relative light units, was normalized to the number of cells ($F/OD_{600\,nm}$).[27,28] At least three independent experiments were carried out to check the reproducibility.

Expression of aECM-CS5-ELF-F in *E. coli* by a T7 expression system

The pET28-CS5-ELF-PheRS* plasmid,[29,30] in which a linker sequence encoding a T7 tag, a hexahistidine tag, and an enterokinase cleavage site was cloned into a pET28 plasmid between its *Nco* I and *Xho* I sites, was transformed into the phenylalanine-auxotrophic *E. coli* BL21(DE3) strain AF (*HsdS gal* (nclts857 *ind*1 *S*am7 *nin*5 *lac*UV5-T7 gene 1) *pheA*), which was constructed in the Tirrell laboratory. This *E. coli* strain AF-IQ[pET28-CS5-ELF-PheRS*] served as the expression system.[29,30] Initially, samples (300 mL) of M9 minimal medium supplemented with 0.4% (w/v) glucose, 35 mg L^{-1} thiamine, 0.1 mM $MgSO_4$, 0.1 mM $CaCl_2$, the 20 common amino acids (40 mg L^{-1} each), 25 mg mL^{-1} kanamycin, and 20 mg L^{-1} chloramphenicol were inoculated with 5 mL of the same medium containing the expression strain that had been cultured at 37 °C overnight. Next, 1 L preparations of the same medium were each inoculated in culture (300 mL). When the $OD_{600\,nm}$ of a culture was between 0.7 and 1.0, a medium shift was performed. To do so, each culture was centrifuged for 7 minutes at 20 100*g*, 4 °C, and the supernatant was removed. Cells were resuspended in the supplemented M9 medium described above that was deficient in Phe. Protein expression was then induced by addition of an *S*-Gal–oligo(Arg) peptide or IPTG (final concentration, 1 mM). After 10 min, Phe was added at concentrations between 50 and 250 mg L^{-1} as previously described.[30] Cells were cultured for an additional 4 h, and protein expression was monitored by SDS PAGE using a normalized OD_{600} of 0.5 per sample. The aECM-CS5-ELF-F purification scheme, which takes advantage of its inverse temperature transition, has been reported.[33–38] Briefly, wet cell masses were each dispersed in TEN buffer (10 mM tris(hydroxymethyl)aminomethane (Tris)-HCl/1 mM ethylenediaminetetraacetate (EDTA), pH 8.0) at a concentration of 1 g mL^{-1}, frozen, and thawed at 4 °C with 1.0 μg mL^{-1} deoxyribonuclease (DNase), 10 ng mL^{-1} ribonuclease (RNase), and 50 ng mL^{-1} phenylmethylsulfonyl fluoride (PMSF) added. aECM-CS5-ELF-F samples, which were found in the pellets of the whole-cell lysates after centrifugation (20 100*g*, 60 min, 37 °C), were resuspended in 4 M urea. The solutions were centrifuged (20 100*g*, 60 min, 4 °C) to remove nonprotein debris and then dialyzed against water (3 days, 4 °C). Precipitates were removed by centrifugation (20 100*g*, 60 min, 2 °C), and the clarified retentates were lyophilized. The purity of the proteins and the uniformity of their molecular masses were confirmed by SDS-PAGE and ^{1}H NMR spectroscopy.

Synthesis of *S*-$(Gal)_2$-Lys-$(Arg)_n$ peptide ($n = 8$) containing two galactoside residues

Solid-phase chemistry, as described above was used to prepare Lys$(Fmoc)_2$-$(Arg)_8$(Pbf) starting from Fmoc-(Pbf)Arg-resin (0.50 g, 0.0176 mol). Next, after removal of the Fmoc groups, **1** (0.460 g, 1.056 mmol) was coupled to the N terminal of the peptide and the lysine side-chain amino group while the peptide was still bound to the resin using the procedure described above. The acetylated di-*S*-Gal-peptide product was removed from the resin while on ice by addition of an ice-cold mixture of trifluoroacetic acid (28.5 mL) and H_2O (1.5 mL), after which the mixture was stirred at room temperature for 1.5 h with simultaneous deprotection of the arginine side chains. Next, the mixture was filtered, and the filtrate was subjected to reduced pressure to remove any remaining trifluoroacetic acid. Then the filtrate was poured into ice-cold diethyl ether (200 mL) and the mixture was centrifuged to isolate the precipitate, which was then dried to give the acetylated peptide as a white powder. Next, to deprotect the galactosyl hydroxyl groups, the product was dissolved in THF containing hydrazine monohydrate (20:1 molar ratio, hydrazine/galactosyl moiety) and was left standing at 0 °C for 16 h, after which acetone was added to the solution, and the mixture was stirred for 1 h to quench the reaction. The solvent was evaporated under vacuum, and the residue was dissolved in H_2O. The solution was dialyzed as described above using seamless cellulose tubing (Spectra/Por Biotech Cellulose ester, cut-off molecular weight of 1000) for 3 days. A white powder was obtained by lyophilization with a yield of 54%.

Acknowledgements

We are grateful to Professor David A. Tirrell (California Institute of Technology) and Dr. Inchan Kwon (Virginia University) for the gift of the E. coli K10 GFP expression system, as well as helpful suggestions and discussion. This study was funded by the Ministry of Education, Culture, Sports, Science and Technology of Japan (Grant-in-Aid for Development Scientific Research, no. 24550132) and Japan Science and Technology (JST) "Research for Promoting Technological Seeds".

Keywords: bacteria · biosynthesis · fluorescence · glycopeptides · protein expression

[1] F. Jacob, J. Monod, *J. Mol. Biol.* **1961**, *3*, 318–356.
[2] W. Gilbert, B. Müller-Hill, *Proc. Natl. Acad. Sci. USA* **1966**, *56*, 1891–1898.
[3] M. Lewis, G. Chang, N. C. Horton, M. A. Kercher, H. C. Pace, M. A. Schumacher, R. G. Brennan, P. Lu, *Science* **1996**, *271*, 1247–1254.
[4] K. Matthews, *Science* **1996**, *271*, 1245–1246.
[5] A. Khlebnikov, J. D. Keasling, *Biotechnol. Prog.* **2002**, *18*, 672–674.
[6] Y. C. Lee, R. T. Lee, *Neoglycoconjugates: Preparation and Applications*, Academic Press Inc., San Diego, **1994**.
[7] A. Kobayashi, T. Akaike, K. Kobayashi, H. Sumitomo, *Makromol. Chem. Rapid Commun.* **1986**, *7*, 645–650.
[8] R. Roy, F. C. Troopper, *J. Chem. Soc. Chem. Commun.* **1988**, 1058–1060.
[9] D. D. Manning, X. Hu, P. Beck, L. L. Kiessling, *J. Am. Chem. Soc.* **1997**, *119*, 3161–3162.

[10] S.-I. Nishimura, T. Furuike, K. Matsuoka, K. Maruyama, K. Nagata, N. Nishi, S. Tokura, *Macromolecules* **1994**, *27*, 4876 – 4880.
[11] G. Wulff, J. Schmid, T. Venhoff, *Macromol. Chem. Phys.* **1996**, *197*, 259 – 274.
[12] K. Yamada, M. Minoda, T. Miyamoto, *Macromolecules* **1999**, *32*, 3553 – 3558.
[13] N. Kojima, S. Hakomori, *J. Biol. Chem.* **1991**, *266*, 17552 – 17558.
[14] K. Matsuura, R. Oda, H. Kitakouji, M. Kiso, K. Kitajima, K. Kobayashi, *Biomacromolecules* **2004**, *5*, 937 – 941.
[15] K. Aoi, K. Ito, M. Okada, *Macromolecules* **1995**, *28*, 5391 – 5393.
[16] K. Aoi, K. Tsutsumiuchi, A. Yamamoto, M. Okada, *Tetrahedron* **1997**, *53*, 15415 – 15427.
[17] T. Hasegawa, S. Kondoh, K. Matsuura, K. Kobayashi, *Macromolecules* **1999**, *32*, 6595 – 6603.
[18] K. Matsuura, M. Hibino, Y. Yamada, K. Kobayashi, *J. Am. Chem. Soc.* **2001**, *123*, 357 – 358.
[19] A. Takasu, T. Houjyou, Y. Inai, T. Hirabayashi, *Biomacromolecules* **2002**, *3*, 775 – 782.
[20] S. Horikoshi, A. Takasu, T. Houjyou, Y. Inai, T. Hirabayashi, *Polym. Prepr. Am. Chem. Soc. Div. Polym. Chem.* **2002**, *43*, 1110 – 1111.
[21] A. Takasu, S. Horikoshi, T. Hirabayashi, *Biomacromolecules* **2005**, *6*, 2334 – 2342.
[22] A. Takasu, T. Makino, T. Hirabayashi, *J. Polym. Sci. Part A* **2009**, *47*, 310 – 314.
[23] A. Takasu, H. Kojima, *J. Polym. Sci. Part A* **2010**, *48*, 5953 – 5960.
[24] S. Futaki, T. Suzuki, W. Ohashi, T. Yagami, S. Tanaka, K. Ueda, Y. Sugiura, *J. Biol. Chem.* **2001**, *276*, 5836 – 5840.
[25] T. Suzuki, S. Futaki, M. Niwa, S. Tanaka, K. Ueda, Y. Sugiura, *J. Biol. Chem.* **2002**, *277*, 2437 – 2443.
[26] Y. Takechi, H. Yoshii, M. Tanaka, T. Kawakami, S. Aimoto, H. Saito, *Langmuir* **2011**, *27*, 7099 – 7107.
[27] H. Lebeau, S. Lindow, *J. Bacteriol.* **2001**, *183*, 6752 – 6762.
[28] K.-S. Ko, J. Kruse, N. L. Pohl, *Org. Lett.* **2003**, *5*, 1781 – 1783.
[29] I. S. Carrico, S. A. Maskarinec, S. C. Heilshorn, M. L. Mock, J. C. Liu, P. J. Nowatzki, C. Franck, G. Ravichandran, D. A. Tirrell, *J. Am. Chem. Soc.* **2007**, *129*, 4874 – 4875.
[30] A. Takasu, S. Kondo, A. Ito, Y. Furukawa, M. Higuchi, T. Kinoshita, I. Kwon, *Biomacromolecules* **2011**, *12*, 3444 – 3452.
[31] A. H. Rosenberg, B. N. Lade, D. Chui, S. Lin, J. J. Dunn, F. W. Studier, *Gene* **1987**, *56*, 125 – 135.
[32] F. W. Studier, A. H. Rosenberg, J. J. , Dunn, J. W. Dubendorff, *Meth. Enzymol.* **1990**, *185*, 60 – 89.
[33] E. R. Welsh, D. A. Tirrell, *Biomacromolecules* **2000**, *1*, 23 – 30.
[34] A. Panitch, T. Yamaoka, M. J. Fournier, T. L. Mason, D. A. Tirrell, *Macromolecules* **1999**, *32*, 1701 – 1703.
[35] S. C. Heilshorn, K. A. DiZio, E. R. Welsh, D. A. Tirrell, *Biomaterials* **2003**, *24*, 4245 – 4252.
[36] P. J. Nowatzki, D. A. Tirrell, *Biomaterials* **2004**, *25*, 1261 – 1267.
[37] J. C. Liu, S. C. Heilshorn, D. A. Tirrell, *Biomacromolecules* **2004**, *5*, 497 – 504.
[38] J. C. Liu, D. A. Tirrell, *Biomacromolecules* **2008**, *9*, 2984 – 2988.
[39] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, *J. Am. Chem. Soc.* **1985**, *107*, 3902 – 3909.
[40] MOPAC 2000, version 1.0, Fujitsu Ltd., Tokyo, Japan, **1999**.

Received: April 12, 2013
Published online on May 23, 2013