A note on expression and purification of recombinant galectins

Nozomu Nishi

I graduated from the Department of Chemistry, Faculty of Science, Osaka University in 1975 and then obtained a doctorate degree (Ph.D.) from the Graduate School of Science, Osaka University in 1980. I started my postdoctoral career in the Department of Biochemistry, University of Southern California as a research scholar. After that, I became an assistant professor in the Faculty of Medicine, Kagawa Medical University and then associate professor in the Life Science Research Center, Kagawa University. Currently, I am a visiting researcher at Kagawa University. I was involved in research on photosynthetic bacteria, prostate/growth factors, and galectins as a protein biochemist.

Summary

This note describes unpublished data regarding the expression (in *E. coli*) and purification of human galectins. I hope that the data and know-how provided below will be useful for researchers (especially for those having limited experience in biochemistry) who are not acquainted with the galectin family and have a need to prepare recombinant galectins.

Introduction *p2*

1. **Basic methods for expression and purification of galectins and the yield of recombinant galectins.** *p2*
2. **Points to be noted for each galectin** *p7*
3. **Supplementary explanation** *p13*

Glycoforum https://www.glycoforum.gr.jp/

A note on expression and purification of recombinant galectins
https://www.glycoforum.gr.jp/article/23A15.html

DOI：https://doi.org/10.32285/glycoforum.23A15

This work is licensed under a Creative Commons Attribution 4.0 International License

Introduction

Table 1 summarizes galectin clones (about 60 human clones and about 20 mouse/rat clones) deposited in the Riken BioResource Research Center by the author. The clone information can be obtained at the homepage of the center: **Riken BioResource Research Center → Gene Engineering Division > Resources for > Search and Browse > Depositors List → Nishi, Nozomu → individual clone.** Detailed description of the clones is given in "Points to be noted for each galectin", if needed. Table 1 contains expression vectors based on pTrcHisB (an *E. coli* expression vector with a tag sequence with a molecular weight of 4,000 - 5,000). There is almost no benefit to use the pTrcHisB-based vectors for the expression of galectins. Please refer to the right most column of Table 1 for the abbreviations of galectins (ex. galectin-1, G1). The fundamental molecular properties of galectins are summarized in Table 2. The methionine residue corresponding to the initiation codon was assigned to residue 1 in this note, if not otherwise specified.

Proto-type galectins consist of a single CRD (carbohydrate recognition domain). On the other hand, a chimera-type galectin (G3) consists of an N-terminal non-CRD domain (collagen-like domain) and a CRD, and tandem-repeat-type galectins (G4, G8, G9, and G12) consist of an N-terminal CRD, a linker peptide, and a C-terminal CRD. Because it is difficult to strictly determine the boundary between the non-CRD domain/linker peptide and the CRD(s), the region attributed to the CRD described below includes some ambiguity. In addition, the term "CRD" does not mean the minimum unit required for sugar-binding activity but, e.g., the whole molecule of proto-type galectins.

1. Basic methods for expression and purification of galectins and the yield of recombinant galectins.

The Data in Table 3, the yields of recombinant galectins, were newly obtained for this note by using the basic methods shown in Fig. 1 (pET vector) and Fig. 2 (pGEX vector). Three different conditions (37°C, 2 h; 30°C, 3 h; 20°C, 16 h) were used for the expression of G1, G3, G4, G7, G8M, and G9S. The expression of other galectins (CSG1, G2, G9Null, G10, G12, G13, and G14/PPL13) was performed under a single condition (20°C, 16 h). Each experiment was carried out twice as shown in Table 3. The GST-free form is designated as the purified protein in the case of GST-fusion proteins (Fig. 2). The protocol shown in Fig. 1 is essentially the same as that posted on the GlycoPOD (GlycoScience Protocol Online Database; Sugar binding proteins: Expression and purification of recombinant human galectin-9), with some modifications. Although this note deals with only *E. coli* expression system, several types of host cells including bacterial cells (*E. coli*), yeast cells (*S. cerevisiae and P. pastoris*), baculovirus-infected insect cells (expres SF+), mammalian cells (COS-1/7, HEK293, and CHO), and moss cells (*Physcomitrella patens*) were evaluated as to the expression of G9, one of the most difficult-to-express galectins. Among them, *E. coli* was found to be the best expression host. In addition, codon optimization did not improve the expression of G9 in *E coli*.

Currently, I am carrying out the expression of galectins at 20°C overnight (14 - 18 h). The yields (mg, mg/L culture) of all the galectins, the expression of which was performed under the three conditions described above, were the highest under 20°C/16 h conditions, as shown in Table 3. However, there is only a little difference in the yields when it is expressed per wet *E. coli* weight (mg/*E. coli* [g]) among the three conditions in the cases of G1 and G7. On the other hand, the expression of G3 and G9S is highly sensitive to the expression temperature. It is indispensable to use the low temperature conditions (20°C/16 h) to obtain sufficient amounts of G3 and G9S (also G9Null and ssG9, data not shown). The data shown in Table 3 were obtained through small scale experiments. In general, the yield (mg/L culture) increases on scaling

up of the culture volume.

SDS-PAGE analysis of samples at different stages of expression and purification (see "Points to be noted for each galectin") revealed that the centrifugal supernatant of *E. coli* extract treated batchwise with lactose-agarose (non-adsorbed fraction) and the wash fraction of the lactose-agarose column contained significant amounts of recombinant proteins in some cases. An attempt to recover the recombinant proteins from these fractions resulted in the isolation of considerable amounts of the proteins. The results of recovery experiments for G1 and G8M are shown in Table 3 as red letters*. Since it is cumbersome and time-consuming, there is little merit to carry out the recovery operation when the yield is satisfactory. It may be possible to reduce the recombinant proteins lost in the wash fraction by decreasing the wash buffer volume without lowering the purity of the product. In the case of glutathione-Sepharose, leakage of recombinant proteins (GST-fusion proteins) into the wash fraction was not evident.

* The non-adsorbed fraction is treated batchwise with lactose-agarose in the same manner as the *E. coli* extract. The recombinant proteins in the wash fraction are isolated by using a column packed with 1 mL of lactose-agarose gel. Other conditions are the same as in Fig. 1 (Step 18/19 - Step 24) .

Comments on the expression and purification steps (Figs. 1 and 2)

・Step 1

According to the manual for a GST Gene Fusion System (GST Gene Fusion System Handbook 18-1157-58 AD 11/2014, GE Healthcare Life Sciences; hereinafter referred to as "GST manual"), it is recommended that 2xYT (+ ampicillin) be used as the culture medium from this step.

・Step 3

The main culture is performed using a 1,000-mL flask containing 220 mL of the expression medium (+ overnight culture). In general, an increase in the culture volume (ex. 400 mL of the medium in a 1,000-mL flask) does not affect the yield (mg/L culture) of recombinant proteins. Note that, because two-ply aluminum foil is used to cover the flasks in the experiment, it is possible that the yield changes when a culture stopper with lower air permeability is used.

GST manual: Dilute the overnight culture 1:100 with the expression medium.

・Step 4

Two mL of the culture medium is periodically collected from the flask placed in an incubator shaker to check the absorbance*. The process, thus, is not completely sterile. In the case that the culture is carried out at 30°C or 20°C after the addition of IPTG (isopropyl β-D-1-thiogalactopyranoside), the temperature setting of the incubator shaker is changed when $A_{600\ nm}$ reaches around 0.5. IPTG is added at the point when $A_{600\ nm}$ reaches 0.60 - 0.65 in most experiments. The value ($A_{600\ nm}$) is somewhat ambiguous, because it fluctuates significantly when a conventional spectrophotometer is used.

* Actually it is not the absorption but mainly due to scattering of the measuring light.

GST manual: Add IPTG when $A_{600\ nm}$ reaches 0.5 - 2 at a final concentration of 0.1 - 1 mM.

・Step 8

E. coli resuspension buffer, 10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM PMSF (phenylmethanesulfonyl fluoride, synonym: benzylsulfonyl fluoride)*, used in the experiments is not optimized for the purification of recombinant galectins. The volumes of resuspension buffer and culture medium are kept in a constant ratio of 0.18 (0.2 including Triton X-100) when the culture volume is changed. PMSF is unstable in water. Since decomposition of PMSF proceeds even if DMSO (dimethyl sulfoxide) is used as a solvent (due to contaminating water), it is essential to prepare the PMSF solution at the time of use.

* The buffer is supplemented with 1 mM DTT (dithiothreitol) when used for G1/mG1 (refer to the section for G1). Although DTT may affect the purity and yield of recombinant proteins in addition to the lectin activity of G1/mG1, we have not examined this point.

GST manual: Use 5 mL of PBS per 100 mL of culture to resuspend the *E. coli* pellet.

・Step 9

A rosette cooling cell (Fig. 3A) placed in an ice-water bath is used to facilitate efficient cooling of the *E. coli* suspension during sonication. Although it is not clear to what extent the increase in the sample temperature induced by sonication affects the yield and/or activity of recombinant galectins, it may be necessary to decrease one sonication time and increase the number of processings when a sonication container with poor cooling efficiency is used.

・Step 10

Triton X-100 (final concentration, 1%[v/v]) is added to facilitate solubilization of recombinant galectins (same condition as in the GST manual). Because the condition is customarily used like the *E. coli* resuspension buffer, a thorough examination has not been made as to purification of recombinant galectins.

・Step 14

Generally 1.5 mL of affinity gel is used for 400 - 500 mL of culture. Instead, 1 mL of the gel per 200 mL of culture was used in the experiments, since too small an amount of the gel may result in a significant error in the yield due to the loss of the gel in the steps hereafter. The binding capacity of lactose-agarose is 5.0 mg/mL gel (= 51 nmol/mL gel) of peanut agglutinin (molecular weight = 98,000) according to the product specification (Honen Corporation). The results in Table 3 show that about 600 nmol (300 nmol, if a dimer is considered as a single molecule) and about 200 nmol of G1 and G8M, respectively, are purified using 1 mL of lactose-agarose gel.

Taking into account that significant amounts of G1 and G8M are found in the non-adsorbed and wash fractions, as described above, the possibility cannot be ruled out that the amount of the gel is insufficient for highly expressed galectins like G1, G3, G7, and G8M. The GST manual recommends that 0.2 mL of glutathione-Sepharose be used for 400 mL of culture (column method; assumed yield, 1 mg of GST-tagged protein per 400 mL of culture). I feel that the amount of glutathione-Sepharose gel recommended in the GST manual is too small. The binding capacity of glutathione-Sepharose is > 5.0 mg GST/mL gel and 25 mg horse liver GST/mL gel according to the product specification (GE Healthcare) and GST manual, respectively.

The batch adsorption method is used in this note, because there were some cases in which the column method was practically unavailable due to the high viscosity of *E. coli* extracts in the purification of recombinant proteins. Note that I have not experienced such a viscous *E. coli* extract in the case of galectins.

・Step 15

Refer to Fig. 3B.

・Step 18

An Econo-Column chromatography column (1.0 x 10 cm, Bio-Rad) is used. An extension tube of about 10 cm long is attached to the column outlet during gel packing and washing to increase the flow rate (Fig. 3C). The extension tube is removed during elution.

・Step 19

The washing buffer contains 0.03% CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate), an amphoteric (twitter ionic) surfactant. I prefer to use CHAPS because of its various advantages for handling proteins [1], even though it is relatively expensive. Washing buffers without a detergent or with a non-ionic detergent*, like Tween 20, may be used without lowering the purity of the product recombinant galectins.

* In general, non-ionic detergents have some disadvantages when compared with ionic detergents: They are difficult to remove by dialysis due to lower critical micellar concentrations and a higher mass of micelles, show higher cytotoxicity, and frequently exhibit high absorbance at 280 nm. Non-ionic detergents must be washed out before elution if they are included in the washing buffer.

・Step 20 (Fig. 1)

The elution buffer, TBS, 0.2 M lactose, is routinely used for lactose-agarose affinity chromatography. G9 shows relatively high affinity for lactose among the members of the galectin family. Purified G9 adsorbed to a lactose-agarose column is eluted with lactose at concentrations of 10 - 15 mM when gradient elution is performed. Therefore it may be possible to lower the lactose concentration in the elution buffer to 50 - 100 mM. It is also possible to use galactose (500 mM) instead of lactose as an eluent, though it is more expensive than lactose. G9 is eluted with galactose at concentrations of 200 - 250 mM under the above mentioned conditions. On X-ray crystallographic analysis, lactose molecules bound to a part of galectin molecules are detectable in some cases even if extensive dialysis is performed to remove lactose in the sample [2]. Although I have not examined it, it may be possible to obtain sugar-free galectin crystals by using galactose as an eluent when lactose affinity chromatography is used for purification.

The affinity gels (lactose-agarose and glutathione-Sepharose) can be used repeatedly after regeneration. The used gels are stored at 4°C in the presence of a preservative like sodium azide, and regenerated as follows: The used gel (~ 20 mL) packed in an Econo-Column chromatography column (2.5 x 10 cm) is washed in sequence with 2 bed volumes of PBS, 2 bed volumes of 6 M guanidine hydrochloride, and 5 bed volumes of PBS, and then suspended in PBS, 0.05% NaN_3 to make a 50% (v/v) slurry. The regeneration time is not limited. In the case that thrombin is added to the GST-fusion proteins adsorbed to glutathione-Sepharose gel (Fig. 2, Step 21 and thereafter), the gel is washed with 50 mM Tris-HCl (pH 8.0), 10 mM glutathione (1 mL x 4, 5-min interval between washings) after recovery of GST-free galectins to remove the GST moiety adsorbed to the gel.

・Step 21 (Fig. 1)

The dialysis conditions (volume of external solution, etc.) are routinely used. A reduced volume of the external solution and shorter exchange cycles with an increased exchange number may result in more efficient removal of lactose, etc., I do not have supporting data though.

・Step 24 (Fig. 1)

The stability of each galectin is described in "Points to be noted for each galectin", if needed. The tandem-repeat type galectins (except for Null-type mutants) are generally unstable compared to the proto-type and chimera-type galectins. The linker region of the tandem-repeat type galectins is highly susceptibility to contaminating proteases in the purified preparations. Care must be taken to store purified galectins in a freezer, to avoid possible irreversible insolubilization.

・Step 20 (Fig. 2)

GST-fusion proteins adsorbed to glutathione-Sepharose gel are treated with thrombin in the experiments (Fig. 2, Step 21 and thereafter). In the cases that galectins are used as GST-fusion proteins or thrombin treatment is carried out in solution, GST-fusion proteins are eluted with 50 mM Tris-HCl (pH 8.0), 10 mM glutathione at this step. Refer to the GST manual for the precautions when eluting GST-fusion proteins.

・Step 21, 22 (Fig. 2)

The thrombin solution is mixed with the glutathione-Sepharose gel in a column with a vortex mixer. Although efficient mixing is not possible, the column covered with parafilm is shaken in an incubator shaker during overnight incubation. In this note, overnight processing was performed uniformly, but the reaction may be completed within several hours. When degradation of recombinant

galectins by contaminating *E. coli*-derived proteases is evident, it is necessary to examine the treatment conditions such as shortening the time (increasing thrombin if necessary).

・Step 23 ~ 27 (Fig. 2)

The solutions (eluate-1 and eluate-2), which contain GST-free galectins, show relatively high absorbance at 280 nm. However, the absorbance greatly decreases on dialysis and centrifugation. The case of G4 (20°C, 16 h) is shown below as an example. Eluate-1 and eluate-2 (1st and 2nd) are mixed, dialyzed, and then centrifuged. The resulting supernatant is sterilized by filtration. The sample name "supernatant" means the sterilized supernatant.

Sample	Volume (mL)	$A_{280\ nm}$
eluate-1	1.0	1.346
eluate-2 (1st)	1.0	1.241
eluate-2 (2nd)	1.0	0.348
supernatant	2.9	0.204

The protein (GST-free G4) concentration estimated by SDS-PAGE of each sample reveals that the decrease in $A_{280\ nm}$ cannot be explained by loss of the protein during the treatments (data not shown). It is most probable that eluate-1 and eluate-2 contain nucleic acids, which are removable by dialysis, because the UV absorbance peaks of eluate-1/-2 and supernatant are at 260 nm and 280 nm, respectively. The situation is the same for other GST-fusion proteins.

Each purified galectin preparation contains thrombin. It is possible to remove thrombin by using a benzamidine-Sepharose or HiTrap benzamidine FF(HS) column (GE Healthcare Life Sciences), if needed (refer to the GST manual). Note that care must be taken when using benzamidine-immobilized carriers. Since benzamidine has a plus charge at around the neutral pH (acid dissociation constant: 11.90 ± 0.40), acidic proteins like G2 (GST-free G2, pI = 5.92) in addition to thrombin are adsorbed if eluate-1/-2 (PBS solution) is directly applied to the benzamidine immobilized carriers (Fig. 4D, AFT; see below for explanation of the figure). The problem can be avoided by increasing the NaCl concentration of the sample to 0.5 M (Fig. 4D, BFT). For example, add 0.08 mL of 5 M NaCl to 1 mL of PBS solution.

Explanation for Fig. 4D

The experiment shown in Fig. 4D is done at 2X scale compared to that shown in Fig. 2. A mixture of eluate-1 and eluate-2 (6 mL) is divided into two equal volumes. One part (A) is directly applied to a HiTrap benzamidine FF(HS) column (1 mL) equilibrated with PBS. The other part (B) is supplemented with 5 M NaCl to make the final concentration 0.5 M, and then applied to another column equilibrated with 20 mM Na-Pi (sodium phosphate) (pH 7.5), 0.5 M NaCl. The non-adsorbed fraction (flow-through fraction) is saved. The adsorbed substances are eluted with 20 mM Na-Pi (pH 7.5), 0.5 M NaCl and then with 0.1 M Glycine-HCl (pH 2.7) (each 2 mL) in the case of A. While they are eluted with 0.1 M Glycine-HCl (pH 2.7) (2 mL) in the case of B. The amounts of samples analyzed by SDS-PAGE are adjusted to make it possible to directly compare the intensity of the stained bands. Note that two types of molecular weight marker kits (Unstained Protein Standard, Broad Range [10-200 kDa], P7704 and P7717, New England BioLabs) are used in Fig. 4 - 7 & 9. In Fig. 4D, P7717 is used.

G2: eluate-1 + eluate-2
AFT, BFT: non-adsorbed fraction
AE1: elution fraction with 20 mM Na-Pi (pH 7.5), 0.5 M NaCl
AE2, BE2: elution fraction with 0.1 M Glycine-HCl (pH 2.7)

GST manual: 0.05 M Tris-HCl, 0.5 M NaCl (pH 7.4) is recommended as equilibration and sample binding buffers for benzamidine-immobilized carriers

2. Points to be noted for each galectin

The expression and purification profiles of galectins on SDS-PAGE are shown in Fig. 4 (A - C), Fig. 5 (A - C), Fig. 6 (A - C), and Fig. 7 (A, C, D, and E). The samples are obtained from the experiments shown in Table 3 (one of the two experiments). Explanation of the samples subjected to SDS-PAGE is as follows.

H0: *E. coli* cells just before the addition of IPTG (Step 4)

H (37H/30H/20H; the same hereinafter): *E. coli* cell homogenate treated with Triton X-100* (Step 11)

S: supernatant obtained by centrifugation of H (Step 12)

P: precipitate obtained by centrifugation of H (Step 12)

LFT: supernatant obtained by centrifugation of S treated with lactose-agarose in a batchwise manner (Step 16)

GFT: supernatant obtained by centrifugation of S treated with glutathione-Sepharose in a batchwise manner (Step 16)

LE: final preparation (purified G1, etc.) (Step 23, Fig. 1)

GE: final preparation (purified G2, etc.) (Step 27, Fig. 2)

Amount of the sample subjected to SDS-PAGE:
H0 - LFT/GFT, 0.1 mg wet pellet equivalent**/10 μL (10 μL/lane)
LE/GE, differs with galectin species (described in legends to figures)

* Although the reason is not clear, the expressed protein band is not detectable in some cases when the *E. coli* cell homogenate before treatment with Triton X-100 (Step 8, 9) is used for SDS-PAGE.

** In the cases of H - LFT/GFT, the sample corresponding to or derived from 0.1 mg of wet *E. coli* pellet (0.1 mg wet pellet equivalent) obtained at Step 7 is analyzed. For example, the wet weight of the *E. coli* pellet is 0.70 g in the case of G1 (the first experiment performed at 37°C, see Table 3). Since the pellet is suspended in 40 mL of the extraction buffer (including Triton X-100, Step 10/11), the concentration is 17.5 μg wet *E. coli* pellet/μL (17.5 μg wet pellet equivalent/μL) when the pellet volume is ignored. The conditions to prepare 100 μL of sample solution subjected to SDS-PAGE are shown below.

Sample (Step 11), 57.1 μL +
SDS sample buffer (4X)*, 25 μL + H_2O, 17.9 μL

* 0.25 M Tris-HCl (pH 6.8), 8%(w/v) SDS, 20%(v/v) mercaptoethanol, 25%(w/v) glycerol

To prepare the H0 sample for electrophoresis, the relationship between *E. coli* wet weight and $A_{600\ nm}$ of the *E. coli* culture is calculated from $A_{600\ nm}$ of the overnight culture (Step 6). In the case of G1 described above, the *E. coli* wet weight obtained from 1 mL of the overnight culture with an absorbance of 1.0 ($A_{600\ nm}$ = 1.0) is determined to be 2.38 mg. This relationship is applied to the culture just before the addition of IPTG (H0 sample). The value significantly fluctuates among different *E. coli* clones. The average and standard deviation are as follows: 1.93 ± 0.55 mg (n = 50). Because the absorbance of the culture obtained just before the addition of IPTG is 0.63, the wet weight of the *E. coli* pellet obtained by centrifugation of 1 mL of the culture is calculated to be 1.50 mg*. The sample solution for SDS-PAGE (0.1 mg wet pellet equivalent/10 μL) can be obtained by treating the pellet (1.50 mg) with 150 μL of SDS sample buffer (1X) when the pellet volume is ignored. It is difficult to strictly determine $A_{600\ nm}$ of the *E. coli* culture using a conventional spectrophotometer, as described in the comment to Step 4. In addition, the linearity between *E. coli* cell density and the absorbance is lost as the density increases. The absorbance of the overnight culture (Step 6) is determined after dilution to make the value lower than 0.7.

* The *E. coli* cell pellet (H0 sample) is stored frozen until the relationship between *E. coli* wet weight and $A_{600\ nm}$ of the overnight culture is determined.

•G1

Previously, galectins were termed S-type lectins, as opposed to C-type lectins, due to their presumed

requirement of reduced cysteine residues for the lectin activity. The name, S-type lectin, is still used, though the term is inappropriate. It is well known that G1 loses lectin activity upon oxidation of cysteine residues (both intramolecular and intermolecular disulfide bond formation), which induces three-dimensional structural alteration of the molecule. The lectin activity of G1 does not depend on reduced cysteine residues [3]. DTT is added to the *E. coli* resuspension buffer at a concentration of 1 mM (Step 8, Figs. 1 and 2) but not to the dialysis buffer in the case of G1/mG1. The lectin activity of G1 determined by hemagglutination assay decreases to about 1/4 after 10 days of storage at 4°C under sterile conditions [4]. The activity decreases to about 1/4 after 20 days of storage even if the final preparation is supplemented with 1 mM DTT. The hemagglutination assay is a simple and sensitive assay and widely used to assess lectin activity. In general, galectins exert their action on cultured cells at concentrations higher than those required for hemagglutination. It is possible, thus, galectin preparations with detectable hemagglutination activity practically lack biological activity toward cultured cells (refer to the sections for G7 and G8).

CSG1 is a cysteine-less mutant of G1 with all Cys residues substituted with Ser residues [4]. CSG1 can be stored without loss of activity in the absence of a reducing agent. Because the mutations do not alter the oligosaccharide specificity or three-dimensional structure, CSG1 can be used as a stable substitute for G1. G1 can obtain neurotrophic activity through oxidation. It has been reported that G1 acquires axonal regeneration promoting activity via macrophages with concomitant loss of lectin activity on the formation of three intramolecular disulfide bonds [5]. Therefore, CSG1 is not a perfect substitute for G1, in that there is not the possibility of oxidation-induced acquisition of neurotrophic activity.

・G2

It is practically impossible to purify G2 by lactose-agarose affinity chromatography, though G2 has detectable affinity for lactose [6,7]. Note that, because the affinity of G2 for lactose was determined using pyridylamidated or p-nitrophenylated saccharide derivatives in previous studies, the affinity for lactose immobilized to agarose may differ from the reported value. It is most probable that the minor material with a molecular weight of about 26,000 in the purified G2 preparation (Fig. 4C) is contaminating GST moiety removed from GST-G2 on thrombin digestion (see Fig. 7B).

・G3

Purified G3 preparations (3 μg/10 μL [10 μL/lane]) obtained at three different temperatures are electrophoresed in the right lanes of Fig. 5A. The protein concentrations of the preparations are 47 μg/mL (37°C), 323 μg/mL (30°C), and 2.40 mg/mL (20°C). Therefore it is not possible to prepare the sample for electrophoresis (3 μg/10 μL) using SDS sample buffer (4X) except for the purified G3 preparation obtained at 20°C. The samples obtained at 37°C and 30°C are concentrated with SCR (StrataClean Resin) before the SDS treatment to achieve the final concentration of 3 μg/10 μL. Refer to "Supplementary explanation" for the concentration of dilute protein solutions with SCR.

G3 is a sole member of the chimera-type galectin subfamily consisting of an N-terminal non-CRD domain (collagen-like domain) and a C-terminal CRD. G3CRD corresponds to the C-terminal CRD*. The yield of G3CRD (pET-G3CRD, 20°C/overnight) is about 3 mg/L culture (less than 10% of that of G3) under conditions similar to those in Table 3.

* An extra Met residue corresponding to the cloning site (NdeI site) is attached to the N-terminal of the product of pET-G3CRD.

・G4

It is practically impossible to purify G4 by lactose-agarose affinity chromatography like G2.

The purified G4 preparations obtained at all the

temperatures are concentrated with SCR before the SDS treatment (Fig. 5B, 37GE*/30GE*/20GE*). Although the intensity of the stained bands should be the same for the three preparations, they apparently differ from each other. It is likely that the protein concentrations of the samples, especially those obtained at 37°C and 30°C, calculated based on the absorbance at 280 nm are overestimated due to contaminating non-protein substances, possibly nucleic acids. Two minor bands with molecular weights of about 76,000 (upper band) and 63,000 (lower band) in the purified G4 preparation are DnaK and GroEL, respectively (Fig. 5B). The identities of the bands are confirmed by N-terminal amino acid sequence analysis*. It is possible to remove DnaK to some extent by washing the glutathione-Sepharose column (Fig. 2, Step 19) with an ATP solution (refer to the section for G10).

* N-terminal amino acid sequence of the upper band: GKIIGIDLGT
N-terminal amino acid sequence of the lower band: AAKDVKFGNDARVKM

The structures of G4Null, G4NT, and G4CT are as follows:

G4Null: N-terminal CRD (Met^{1} - Gln^{153}) + two parts of the linker peptide (Pro^{154}-Leu^{155} + Thr^{187} - Pro^{191}) + C-terminal CRD (Val^{192} - Ile^{323})

G4NT: N-terminal CRD (Met^{1} - Gln^{153}) + a part of the linker peptide (Pro^{154} - Tyr^{166})

G4CT: a part of the linker peptide (Thr^{181} - Pro^{191}) + C-terminal CRD (Val^{192} - Ile^{323})

The yields of G4Null, G4NT, and G4CT (GST-free proteins) are about 1.6 mg/L culture, 10 mg/L culture, and 20 mg/L culture, respectively, under conditions similar to those in Table 3.

・G7

G7R54H is a mutant with a substitution of Arg^{54} to His. Arg^{54} is one of the important amino acid residues which constitute the carbohydrate-binding site of G7. The carbohydrate-binding activity of G7 is almost completely abolished by the mutation. However, it is possible that the mutant retains detectable hemagglutination activity. The situation is the same for other galectins with analogous mutations (refer to the section for G8).

・G8

There are at least two G8 isoforms (splicing isoforms: G8M and G8L), which only differ in the linker peptide region. The lengths of the linker peptides are 28 and 70 amino acid residues for G8M and G8L, respectively. The yield of G8L (pET-G8L) is comparable to that of G8M. G8MR69H (G8LR69H) and G8MR233H (G8LR275H) are mutant forms carrying Arg-to-His mutations at the N- and C-terminal CRDs, respectively, of G8M (G8L). The mutations are analogous to that of G7 described above. G8MR69,233H (G8LR69,275H) has the mutation at both CRDs. The two CRDs of G8 exhibit quite different affinities for lactose: The N-terminal CRD has relatively high affinity, while the C-terminal CRD has practically no affinity for lactose [6,7]. Therefore, lactose-affinity chromatography is not applicable to the G8 mutants with an inactivated N-terminal CRD even if the C-terminal CRD is in the native form. I have reported that G8MR69H exhibits a lower (compared to the wild-type G8) but significant affinity for asialofetuin-agarose [8]. G8MR69H, however, cannot be retained on an asialofetuin-agarose column unless it is washed minimally.

As mentioned in the section for G7, it is possible that a mutant galectin with a substitution(s) of His for Arg at the carbohydrate biding site(s) may exhibit hemagglutination activity. Fig. 8 shows the hemagglutination activity of G8MR69,233H in comparison with that of G8M (refer to "Supplementary explanation" for the hemagglutination assay). G8M exhibits a relatively low activity among the members of the galectin family: The minimum concentration required for hemagglutination under the conditions used is 0.2 - 0.4 μM (Fig. 8A). G8MR69,233H exhibits the activity at 6.4 μM or higher, suggesting that the double

mutant retains about 3 - 6% of the wild-type activity. The activity of G8M at 6.4 μM or lower is inhibited by 50 mM lactose. Although hemagglutination by high concentrations (12.8 and 25.6 μM) of G8M is hardly inhibited, a weak inhibitory effect is observed at 12.8 μM when the incubation time is extended to 2 h (Fig. 8B). On the other hand, G8MR69,233H-induced hemagglutination is not inhibited by 50 mM lactose regardless of the incubation time (Fig. 8C). The results apparently show that G8MR69,233H has hemagglutination activity. However, it is not clear whether this is based on the residual lectin activity.

G8LR197A is a mutant with a substitution of Arg^{197} to Ala at the linker peptide. There is a sequence recognized by thrombin (-IAPRT-) in the linker peptide of G8L but not that of G8M. The mutation eliminates the sensitivity of G8L to thrombin [9]. Therefore, thrombin treatment of GST-G8L without this mutation (the products of pGEX-G8(L), pGEX-G8(L)R69H, pGEX-G8(L)R275H, and pGEX-G8(L) R69,275H) results in not only excision of the GST moiety but also dissociation of the two CRDs.

The linker peptide of the tandem-repeat-type galectins is a so-called "disordered region" (refer to "Supplementary explanation"), which does not form a fixed three-dimensional structure and is highly susceptible to a variety of proteases. In addition, it is not possible to completely remove proteases of *E. coli* origin from recombinant protein preparations by a single affinity chromatography. G8Null is a mutant form of G8 lacking most of the linker peptide and highly resistant to proteolytic degradation/inactivation [10]. The yield of G8Null (pET-G8Null) is about 40 mg/L culture. Note that a restriction enzyme site (NdeI site: CATATG), which is used to construct the expression vector, exists at the connecting point of two cDNAs coding for the N- and C-terminal CRDs. Accordingly, an extra two amino acid residues (HM) corresponding to the restriction enzyme site exist in G8Null. The structures of G8Null, G8NT, G8CT, G8NCRD, and G8CCRD are as follows:

G8Null: N-terminal CRD* (Met^1 - Ser^{155}) + residues corresponding to the NdeI site (HM) + C-terminal CRD (Arg^{184} - Trp^{317})
G8NT: N-terminal CRD* (Met^1 - Ser^{155}) + a part of the linker peptide (Asp^{156} - Thr^{168})
G8CT: a part of the linker peptide (Ser^{171} - Leu^{183}) + C-terminal CRD (Arg^{184} - Trp^{317})
G8NCRD: N-terminal CRD* (Met^1 - Phe^{153})
G8CCRD: C- terminal CRD (Arg^{184}- Trp^{317})

* It would be more appropriate to assign (Met^1 - Ser^{155}) as the N-terminal CRD of G8.

・G9

There are at least three G9 isoforms (splicing isoforms: G9S, G9M, and G9L), which only differ in the linker peptide region. The lengths of the linker peptides are 33, 45, and 72 amino acid residues for G9S, G9M, and G9L, respectively. Although (Thr^{166} - Thr^{311}) (the amino acid residue numbers are those of G9S) was assumed to be the C-terminal CRD of G9 in a previous paper [10], it would be more appropriate to assign (Pro^{182} - Thr^{311}) as the CRD considering the three-dimensional structure. The yield of G9 markedly differs among the isoforms. It is difficult to obtain recombinant G9L that is sufficient for general biochemical experiments. The yield of G9M is about 30% of that of G9S, and the linker peptide of G9M is significantly more susceptible to proteolysis than that of G9S (data not shown). The intensity of the stained main band of 30LE* is apparently lower than that of 20LE* (Fig. 6B), similar to the case of G4. It is possible that the protein concentration of 30LE is overestimated. In addition, the formation of SDS-resistant oligomers/polymers (the bands near top of the gel) is observed in several galectin preparations including G9 (refer to the section for G13, G14/PPL13), and this may result in the differences in the stained band intensity.

G9MR65H, G9MR239H, and G9MR65,239H are mutant forms of G9M carrying Arg-to-His mutations analogous to those of G8.

G9Null is a mutant form of G9 lacking a part of the linker peptide and highly resistant to proteolytic degradation [10]. Sixteen amino acid residues of the linker peptide exist in G9Null as described above, and an extra two amino acid residues (HM) are also inserted like in the case of G8Null. The yield of G9Null is about two times higher than that of G9S. G9 can exist at relatively high concentrations shortly after elution from lactose-agarose. However, the solution starts to turn cloudy within a few minutes after elution, and most G9 (including G9Null) becomes insoluble after dialysis of the eluate against PBS and other buffers of physiological pH. The maximum concentration of G9Null, which is more soluble than G9S, in PBS is about 400 μg/mL. For that reason, it is possible to raise the yield of G9 by diluting the elution fractions of high G9 concentrations immediately after elution with those of low G9 concentrations, which are normally discarded. However, if the concentration of G9Null is adjusted to 400 μg/mL before dialysis, the concentration of the final preparation becomes significantly lower than this value because insolubilization of G9Null during dialysis occurs even at that concentration. The yield data for G9S and G9Null shown in Table 3 were obtained by using this technique. It is possible to prevent the insolubilization of G9 in the eluate by lowering the pH of both the elution buffer and dialysis buffer [11]. Refer to "Supplementary explanation" for purification of G9 under acidic conditions. G9NullR65D/G9NullR211D are mutants analogous to G9MR65H/G9MR239H. Because it is not possible to obtain recombinant proteins carrying an Arg to His mutation (G9NullR65H/G9NullR211H), G9NullR65D/G9NullR211D are used instead.

Although part of the problem associated with wild-type G9, high sensitivity to proteolytic inactivation, has been solved by the development of G9Null, the structural modification has only slightly improved the solubility at physiological pH, another issue with wild-type G9. ssG9 (highly stable and soluble form of G9) has been developed to solve the solubility problem [12]. ssG9 having a 10-amino acid deletion and a single amino acid substitution at the remaining linker peptide region of G9Null can stably exist at concentrations higher than 2 mg/mL under physiological pH. The modification has no adverse effect on the protease sensitivity or biological activity of G9Null. Six amino acid residues, HPPYPM (wild-type sequence: HPAYPM), exist between the N- and C-terminal CRD of ssG9, and no extra amino acid residue exists. The cDNA for wild-type G9 contains a BamHI site. The sequence has been modified (GGATCC → GTATCC) in ssG9 cDNA. The structures of G9Null, ssG9, G9NT, G9CT, G9NCRD, and G9CCRD are as follows:

G9Null: N-terminal CRD (Met^{1} - Gln^{148}) + residues corresponding to the NdeI site (HM) + a part of the linker peptide + (Thr^{166} - Met^{181}) + C-terminal CRD (Pro^{182} - Thr^{311})

ssG9: N-terminal CRD (Met^{1} - Gln^{148}) + a part of the linker peptide with a modification (HPPYPM) + C-Terminal CRD (Pro^{182} - Thr^{311})

G9NT: N-terminal CRD (Met^{1} - Gln^{148}) + a part of the linker peptide (Pro^{149} - Thr^{167})

G9CT*: a part of the linker peptide (Ile^{160}- Met^{193}) + C-terminal CRD (Pro^{194} - Thr^{323})

G9NCRD: N-terminal CRD (Met^{1} - Gln^{148})

G9CCRD: a part of the linker peptide + (Thr^{166} - Met^{181}) + C-Terminal CRD (Pro^{182} - Thr^{311})

* The amino acid residue numbers of G9M are used for G9CT, and those of G9S are used for other clones.

・G10

It is practically impossible to purify G10 by lactose-agarose affinity chromatography.

The purified G10 preparation is concentrated with methanol before the SDS treatment (Fig. 7A, 20GE*). Refer to "Supplementary explanation" for concentration of dilute protein solutions with methanol.

The purified G10 preparation contains *E. coli* chaperones like G4. The procedure to remove one of the chaperones, DnaK, described in the GST manual is applied to G10 to check the effectiveness of the method (Fig. 7B). In the experiment shown in Fig. 7B, incubation of GST-G10 with ATP is carried out while it is adsorbed to glutathione-Sepharose, although the GST-manual recommends incubation of the *E. coli* extract with ATP before affinity chromatography. The GST manual also mentions the use of an ATP-immobilized column. In addition, GST-G10 is treated with thrombin after elution from the glutathione-Sepharose column, which is different from the conditions given in Fig. 2. DnaK is released by the incubation with ATP (Fig. 7B ATP·W). The content of DnaK in the purified G10 preparation treated with ATP (the right most lane in Fig. 7B) is apparently lower than that in the untreated preparation. The treatment also has some effect on the content of GroEL.

Explanation for Fig. 7B

The experiment shown in Fig. 7B is performed on 2X scale compared to that shown in Fig. 2 until Step 17. The gel is packed into two columns (1 mL gel/column), and then washed as described in Step 19. GST-G10 is eluted from one column (column A) with 50 mM Tris-HCl (pH 8.0), 10 mM glutathione. One mL of ATP solution, 50 mM Tris-HCl (pH 7.5), 2 mM ATP, 10 mM $MgCl_2$, is added to the other column (column B; the effluent is discarded) and then incubated for 10 min at 37°C. The procedure is repeated three times. The second and third effluents are saved as the ATP-wash fraction. After washing the column with 3 mL of TBS, 0.03% CHAPS, GST-G10 is eluted with 50 mM Tris-HCl (pH 8.0), 10 mM glutathione. The eluate from the column is supplemented with 30 µL of thrombin solution (1 u/µL), and then incubated for 16 h at 25°C, followed by dialysis and centrifugation under the conditions given in Fig. 2 (Step 25 and Step 26). The resulting supernatant is applied to a glutathione-Sepharose column (0.5 mL gel). The non-adsorbed fraction is collected and designated as the purified G10 preparation. Thirty µL of each sample was mixed with 10 µL of SDS sample buffer (4X) and then subjected to heat treatment. The heat-treated samples were analyzed by SDS-PAGE (10 µL/lane).

GE: eluate from glutathione-Sepharose with 50 mM Tris-HCl (pH 8.0), 10 mM glutathione
GE·Th: GE treated with thrombin
GE·Th·d: supernatant fraction obtained by centrifugation of dialyzed GE·Th
G10: purified G10
ATP·W: ATP-wash fraction (column B)

·G12

It is practically impossible to obtain recombinant G12, LNG12, G12NT and G12CT in a soluble form using the expression vectors shown in Table 1*. It is most probable that the main band of the purified fraction (Fig. 7C, 20GE) is GroEL.

G12 (the cDNA insert of pGEX-G12) corresponds to the clone reported by Yang et al. [13]. Hotta et al. [14] have reported the presence of an alternative initiation codon in the 5' upstream region of that described by Yang et al. LNG12 corresponds to the latter clone with an N-terminal extension of 22 amino acid residues. Western blot analysis and/or *in vitro* translation in the presence of radio-labeled methionine are used to detect expressed G12 and LNG12, and no experiment using purified recombinant G12/LNG12 preparation has been conducted. Note that the radio-labeled LNG12 shows low affinity for lactose-agarose [14]. The structures of G12NT and G12CT are as follows:

G12NT: N-terminal CRD (Met^{1} - Leu^{161}) + a part of the linker peptide (Asn^{162} - Ser^{170})
G12CT: a part of the linker peptide (Pro^{174} - Val^{188}) + C-terminal CRD (Pro^{189} - Ser^{314})

Because the three-dimensional structure of G12 has not been reported except for the results of "*in silico* analysis", the assignment of CRD and linker regions is tentative.

* Maller et al. [15] have reported the expression and purification of soluble mouse G12 using the pET-28a vector and *E. coli* host strain Rosseta 2(DE3)pLysS, a BL21 derivative designed to facilitate the expression of eukaryotic proteins. The expression of tRNAs for seven rare codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) is enhanced in the strain. Although the mouse G12 expressed by the method has a His-tag, it is not possible to purify the protein with either Ni-NTA carrier or lactose-agarose. The protein is purified by ion-exchange chromatography on Q-Sepharose and CM-Sepharose columns (yield, 5 mg/L culture).

・G13, G14/PPL13

It is practically impossible to purify G13 and G14/PPL13 by lactose-agarose affinity chromatography.

The purified G14/PPL13 preparation is concentrated with methanol before the SDS treatment (Fig. 7E, 20GE*). Refer to “Supplementary explanation” for concentration with methanol. The values for "Protein concentration (mg/mL, A280 nm = 1.0)" in Table 2 and “Purified protein yield” in Table 3 for G14/PPL13 are not reliable because of the presence of a large amount of impurities (provably GroEL) (Fig. 7E).

The samples, 20GE§ and 20GE*,§, shown in Fig. 7D are not heat-treated before electrophoresis. It is evident that heat treatment of purified G13 in SDS sample buffer results in the formation of SDS-resistant oligomers/polymers (Fig. 7D, 20GE and 20GE*, the bands near top of the gel). A similar phenomenon is observed in the case of G9S etc., but it is not necessarily related to the heat treatment.

3. Supplementary explanation

Concentration of proteins for SDS-PAGE

It is usually needed to concentrate and dialyze protein preparations in biochemical research. This is problematic and time-consuming when it is indispensable to preserve their biological activity. On the other hand, it is possible to use more conventional methods for the concentration of dilute protein solutions for SDS-PAGE (Coomassie staining, Western blotting, lectin blotting, etc.). Three concentration methods, including those used for the concentration of purified galectin preparations described above, are described below, though they have no specific relation to the field of galectin research. The three methods comprise the methanol precipitation method, the SCR adsorption method, and the TCA-DOC (trichloroacetic acid-deoxycholic acid) precipitation method (hereinafter referred to as the methanol method, SCR method, and TCA-DOC method). The results with the three methods applied to rat tissue extracts and purified galectin preparations are shown in Fig. 9A and B.

Explanation for Fig. 9

Rat tissue extracts are concentrated by the three methods (Fig. 9A), and purified galectin preparations (and BSA, bovine serum albumin) are concentrated by the methanol method and SCR method (Fig. 9B). The control sample (C) contains 7.5 μg protein/lane and 1 μg protein/lane for Figs. 9A and 9B, respectively. MeOH, TCA-DOC, and SCR indicate the samples concentrated with the methanol method, TAC-DOC method, and SCR method, respectively. The methanol method is best suited for the concentration of rat tissue extracts under the conditions given below. It is evident that some protein bands are not observed at all or observed to only a limited extent with the SCR method (Fig. 9A, arrowheads). Almost the same results are obtained with the galectin preparations: There is only a small difference between the methanol method and SCR method in the recovery of galectins with relatively high molecular weights, i.e., the chimera-type and tandem-repeat-type galectins, while the recovery of the proto-type galectins and BSA is low when the SCR method is used. Although the SCR method is an excellent protein concentration method, care must be taken as to the applicability of the method to the proteins of interest. The effect of priming (see below, “SCR method”) on the concentration efficiency of the SCR method is uncertain, because comparative experiments have not been performed. The G4

preparation used for the concentration experiments contains degradation products formed during storage (Fig. 9B). All the G13 samples for SDS-PAGE are not subjected to heat treatment (refer to the section for G13, G14/PPL13).

・Methanol method

This is one of the protein precipitation methods like ammonium sulfate precipitation. In general, the precipitation method is not applicable to dilute protein solutions. The precipitation method is usually used for samples with protein concentrations of higher than 1 mg/mL. The method is intentionally applied to the low concentration protein solutions (rat tissue extracts, 100 μg/mL; galectin solutions, 20 μg/mL) in the experiments shown in Fig. 9. Four times the volume of methanol is added to each sample (final methanol concentration, about 80% [v/v]).

Rat tissue extract (100 μg/mL in TBS), 300 μL in a 2-mL tube
↓ + 1.2 mL of methanol, followed by mixing
↓ 30 min on ice
↓ cent., 14,000 rpm (15,000 xg at R_{av}), 20 min
Ppt.
↓ residual methanol was removed by an additional brief centrifugation and careful aspiration
↓ dissolved in 40 μL of SDS sample buffer (1X), and then subjected to heat treatment
SDS-PAGE

Galectin solution (20 μg/mL in PBS), 200 μL in a 1.5-mL tube
↓ + 0.8 mL of methanol, followed by mixing
↓ 30 min on ice
↓ cent., 14,000 rpm, 20 min
Ppt.
↓ residual methanol was removed by an additional brief centrifugation and careful aspiration
↓ dissolved in 40 μL of SDS sample buffer (1X), and then subjected to heat treatment
SDS-PAGE

・SCR method

SCR is hydroxylated silica particles originally developed to remove proteins (restriction enzymes, etc.) from DNA solutions in place of the phenol-chloroform extraction method. SCR can also be used for the concentration of dilute protein solutions exploiting the non-specific nature of its protein adsorption. The SCR method, in contrast to the precipitation method, is applicable to dilute protein solutions and takes only a short time. In the protocol below, 10 μL of SCR slurry is added to 300 μL of rat tissue extract (100 μg protein/mL) or 200 μL of purified galectin solution (20 μg protein/mL). The amount of sample (concentration and/or volume) can be increased as long as it is within the adsorption capacity of the SCR used. Like the methanol method, the SCR method can be applied for the preparation of samples mainly for SDS-PAGE and mass spectrometry analyses, in which preservation of biological activity is not needed. It seems difficult to recover proteins adsorbed to SCR in a biologically active form. Since detergents inhibit the interaction between SCR and proteins, care must be taken when applying this method to detergent-containing samples.

Otto et al. have reported a general protocol for the enrichment and storage of proteins using SCR [16]. In the protocol, SCR is subjected to “priming” (incubation in hydrochloric acid at 100°C for 6 h) before use. Commercially available SCR without "priming" is used in the experiment below. Note that SCR is a relatively expensive reagent.

Rat tissue extract (100 μg/mL in TBS), 300 μL in a 1.5-mL tube
Galectin solution (20 μg/mL in PBS), 200 μL in a 1.5-mL tube
↓ + 10 μL of SCR slurry (50%[v/v] in H_2O)
↓ mixed for 1 min
↓ cent., 30 sec (table top centrifuge)
Ppt.
↓ residual supernatant was removed by an additional brief centrifugation and careful aspiration

↓ suspended in 40 μL of SDS sample buffer (1X), and then subjected to heat treatment
↓ cooled on ice
↓ cent., 30 sec (table top centrifuge)
Sup.
↓
SDS-PAGE

・TCA-DOC method

This method has been developed to determine the protein concentrations of samples containing interfering substances by Lowry's method [17]. Although this method is one of the precipitation methods, the use of DOC as a carrier makes it applicable to dilute protein solutions. An additional washing step with methanol is added to the original protocol because TCA is used as a precipitant.

Rat tissue extract (100 μg/mL in TBS), 300 μL
↓ + 300 μL of TBS
↓ + 5 μL of 2% sodium DOC*, followed by mixing
↓ 15 min at room temperature
↓ + 150 μL of 30% TCA, followed by mixing
↓ 15 min on ice
↓ cent., 14,000 rpm (15,000 xg at R_{av}), 20 min
Ppt.
↓ residual supernatant was removed by an additional brief centrifugation and careful aspiration
↓ + 1 mL of methanol, followed by mixing
↓ cent., 14,000 rpm, 5 min
Ppt.
↓ residual supernatant was removed by an additional brief centrifugation and careful aspiration
↓ dissolved in 40 μL of SDS sample buffer (1X), and then subjected to heat treatment
SDS-PAGE

* The pH of the 2% sodium DOC solution is adjusted to 8.0 - 8.5 with NaOH to facilitate solubilization of sodium DOC. The solution may become cloudy during storage but can be used after through mixing.

Hemagglutination assay

The following method was used in this note.

96-well plate, V-bottom (U-bottom)
↓ + 50 μL/well of sample (serially diluted with PBS, 10 mg/mL BSA, 0.05% NaN_3)
↓ + 50 μL/well of 2% (v/v) suspension of trypsinized, glutaraldehyde-fixed rabbit erythrocytes (diluted with PBS, 10 mg/mL BSA, 0.05% NaN_3), and then mixed by pipetting several times using a multichannel pipette set at 50 μL
↓ stand for 1 h (2 h) at room temperature
The extent of hemagglutination was observed visually.

Since hemagglutination is visually evaluated by an appearance of mat formation, it may be difficult to determine the minimum effective concentration depending on the nature of the lectin. In addition, the shape of the plate and the concentration of the erythrocyte suspension may affect the results. Fig. 10 shows the results of hemagglutination assay of G8M, G8MR69,233H, G9Null, and CSG1 using two types of plates. In the case of G9Null, the shape of the plate (V-bottom or U-bottom) does not affect visual judgment of agglutination, while it is difficult in the case of G8MR69,233H when a U-bottom plate is used. Also, the aggregation state changes gradually with increasing concentrations of G8M and CSG1 (especially CSG1) regardless of plate type, making it difficult to unequivocally determine the minimum effective concentration.

The disordered region of tandem-repeat-type galectins

The predicted disordered region of galectins (G1 - G13), which has been obtained by using the Web service of POODLE (Prediction Of Order and Disorder by machine LEarning), is shown in Fig. 11. Although this service is not currently available, it seems possible to perform the analysis on a local

PC (Protein Structure Prediction Active Workflow: https://togo.medals.jp/active.html). The N-terminal non-lectin domain of G3 and the linker peptide region of tandem-repeat-type galectins exhibit high disorder probability. There is no typical disordered region in G8Null and G9Null, which lack most part of linker peptide. In the case of LNG12/G12, there is a region with a slightly high disorder probability. However, this region does not correspond to the expected linker peptide. This may have be related to the unusual intracellular localization of the protein: G12 is localized to the inner surface of lipid droplets of 3T3-L1 cells [15]. The reason for the presence of regions with relatively high disorder probability in the proto-type galectins, G7 and G13, is unclear.

Purification of G9 under acidic conditions

The yield of G9Null purified under the modified protocol (Step 20 and thereafter of Fig. 1, see below) is as follows (200-mL culture; 20°C, 16 h):

Acidic conditions*: 0.708 mg/mL** x 3.85 mL
Original conditions*: 0.385 mg/mL x 5.00 mL

* The experiments are performed simultaneously for direct comparison.
** The value is calculated on the assumption that the absorption coefficient at 280 nm of G9Null is the same at pH 6.0 as it is at 7.5.

The eluted fractions with high G9Null concentrations are diluted with those with low concentrations in the latter case (original conditions; refer to G9 section) but not in the former case (acidic conditions). This results in differences in the sample volume. A significant amount of precipitate (insolubilized G9Null) is obtained at Step 22 (Fig. 1) under the original conditions. An effort to recover G9Null in a soluble form from the precipitate with an acidic buffer (20 mM Na-Pi (pH 6.0), 0.15 M NaCl, 0.2 M lactose) is not successful (data not shown).

20) Wash the gel with 2 gel-bed volumes of 0.15 M NaCl.
↓
21) Elute the recombinant protein from the gel with 20 mM Na-Pi (pH 6.0), 0.15 M NaCl, 0.2 M lactose.
↓
22) Dialyze the eluate against 20 mM Na-Pi (pH 6.0), 0.15 M NaCl*.
*1st, 4 - 5 h (500 mL); 2nd, overnight (1,000 mL); 3rd, 4 - 5 h (500 mL)
↓
23) Transfer the dialyzed solution to a centrifuge tube, and then spin down the insoluble material at 25,000 xg for 20 min.
↓
24) Sterilize the supernatant with a sterile filter (0.2 μm).
↓
25) Store the sterilized solution at 4 °C.

Acknowledgements

I would like to thank Drs. Aiko Itoh (Kagawa University), Takanori Nakamura (Kagawa University), and Jun Hirabayashi (Advanced Industrial Science and Technology) for their valuable cooperation and support for the preparation and publication of this manuscript.

Abbreviations

BSA: bovine serum albumin
CHAPS: 3-[(3-Cholamidopropyl) dimethylammonio]propanesulfonate
CRD: carbohydrate recognition domain
DMSO: dimethylsulfoxide
DOC: deoxycholic acid
DTT: dithiothreitol
FAC: frontal affinity chromatography
G1 - G14: galectin-1 - galectin-14
GST: glutathione S-transferase
IPTG: isopropyl β-D-1-thiogalactopyranoside
Na-Pi: sodium phosphate
PMSF: phenylmethanesulfonyl fluoride (synonym: benzylsulfonyl fluoride, etc.)
SCR: StrataClean Resin
TBS: Tris-buffered saline
TCA: trichloroacetic acid

References

1) Matuo, Y., Nishi, N., Muguruma, Y., Yoshitake, Y., Masuda, Y., Nishikawa, K., and Wada, F. (1988) Stabilization of fibroblast growth factors by a non-cytotoxic zwitterionic detergent, 3-[(3-cholamidopropyl)dimethylammonio]-1-propane sulfonate (CHAPS). *In Vitro Cell Dev Biol* **24**, 477-480
2) Nonaka, Y., Ogawa, T., Yoshida, H., Shoji, H., Nishi, N., Kamitori, S., and Nakamura, T. (2015) Crystal structure of a Xenopus laevis skin proto-type galectin, close to but distinct from galectin-1. *Glycobiology* **25**, 792-803
3) Hirabayashi, J., and Kasai, K. (1991) Effect of amino acid substitution by sited-directed mutagenesis on the carbohydrate recognition and stability of human 14-kDa beta-galactoside-binding lectin. *J Biol Chem* **266**, 23648-23653
4) Nishi, N., Abe, A., Iwaki, J., Yoshida, H., Itoh, A., Shoji, H., Kamitori, S., Hirabayashi, J., and Nakamura, T. (2008) Functional and structural bases of a cysteine-less mutant as a long-lasting substitute for galectin-1. *Glycobiology* **18**, 1065-1073
5) Horie, H., Kadoya, T., Hikawa, N., Sango, K., Inoue, H., Takeshita, K., Asawa, R., Hiroi, T., Sato, M., Yoshioka, T., and Ishikawa, Y. (2004) Oxidized galectin-1 stimulates macrophages to promote axonal regeneration in peripheral nerves after axotomy. *J Neurosci* **24**, 1873-1880
6) Hirabayashi, J., Hashidate, T., Arata, Y., Nishi, N., Nakamura, T., Hirashima, M., Urashima, T., Oka, T., Futai, M., Muller, W. E., Yagi, F., and Kasai, K. (2002) Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. *Biochim Biophys Acta* **1572**, 232-254
7) Iwaki, J., Tateno, H., Nishi, N., Minamisawa, T., Nakamura-Tsuruta, S., Itakura, Y., Kominami, J., Urashima, T., Nakamura, T., and Hirabayashi, J. (2011) The Galbeta-(syn)-gauche configuration is required for galectin-recognition disaccharides. *Biochim Biophys Acta* **1810**, 643-651
8) Nishi, N., Shoji, H., Seki, M., Itoh, A., Miyanaka, H., Yuube, K., Hirashima, M., and Nakamura, T. (2003) Galectin-8 modulates neutrophil function via interaction with integrin alphaM. *Glycobiology* **13**, 755-763
9) Nishi, N., Itoh, A., Shoji, H., Miyanaka, H., and Nakamura, T. (2006) Galectin-8 and galectin-9 are novel substrates for thrombin. *Glycobiology* **16**, 15C-20C
10) Nishi, N., Itoh, A., Fujiyama, A., Yoshida, N., Araya, S., Hirashima, M., Shoji, H., and Nakamura, T. (2005) Development of highly stable galectins: truncation of the linker peptide confers protease-resistance on tandem-repeat type galectins. *FEBS Lett* **579**, 2058-2064
11) Nonaka, Y., Ogawa, T., Oomizu, S., Nakakita, S., Nishi, N., Kamitori, S., Hirashima, M., and Nakamura, T. (2013) Self-association of the galectin-9 C-terminal domain via the opposite surface of the sugar-binding site. *J Biochem* **153**, 463-471
12) Itoh, A., Fukata, Y., Miyanaka, H., Nonaka, Y., Ogawa, T., Nakamura, T., and Nishi, N. (2013) Optimization of the inter-domain structure of galectin-9 for recombinant production. *Glycobiology* **23**, 920-925
13) Yang, R. Y., Hsu, D. K., Yu, L., Ni, J., and Liu, F. T. (2001) Cell cycle regulation by galectin-12, a new member of the galectin superfamily. *J Biol Chem* **276**, 20252-20260
14) Hotta, K., Funahashi, T., Matsukawa, Y., Takahashi, M., Nishizawa, H., Kishida, K., Matsuda, M., Kuriyama, H., Kihara, S., Nakamura, T., Tochino, Y., Bodkin, N. L., Hansen, B. C., and Matsuzawa, Y. (2001) Galectin-12, an Adipose-expressed Galectin-like Molecule Possessing Apoptosis-inducing Activity. *J Biol Chem* **276**, 34089-34097
15) Maller, S. M., Cagnoni, A. J., Bannoud, N., Sigaut, L., Perez Saez, J. M., Pietrasanta, L. I., Yang, R. Y., Liu, F. T., Croci, D. O., Di Lella, S., Sundblad, V., Rabinovich, G. A., and Marino, K. V. (2020) An adipose tissue galectin controls endothelial cell function via preferential recognition of 3-fucosylated glycans. *FASEB J* **34**, 735-753
16) Otto, A., Maass, S., Bonn, F., Buttner, K., and Becher, D. (2017) An Easy and Fast Protocol for Affinity Bead-Based Protein Enrichment and Storage of Proteome Samples. *Methods Enzymol* **585**, 1-13
17) Bensadoun, A., and Weinstein, D. (1976) Assay of proteins in the presence of interfering materials. *Anal Biochem* **70**, 241-250

Table 1. Galectin clones deposited in the RIKEN Bioresource Center DNA Bank

Catalog #[1)]	Name of clone	Vector (cloning site)	Expressed recombinant protein
Galectin-1			
RDB04217	pGEX-G1-1a	pGEX-4T-2 (BamHI/XhoI)	Human galectin-1 (G1), GST fusion protein
RDB08392	pET-G1-1a	pET-11a (NdeI/BamHI)	G1, tag-free protein
RDB08448	pGEX-CSG1-2a	pGEX-4T-2 (BamHI/XhoI)	G1 in which all Cys residues are substituted with Ser residues (G1C3,17,43,61,89,131S: CSG1), GST fusion protein
RDB08449	pET-CSG1-1a	pET-11a (NdeI/BamHI)	CSG1, tag-free protein
RDB08433	pGEX-mG1-1a	pGEX-4T-2 (BamHI/XhoI)	Mouse galectin-1 (mG1), GST fusion protein
RDB08434	pET-mG1-1a	pET-11a (NdeI/BamHI)	mG1, tag-free protein
Galectin-2			
RDB04218	pGEX-G2-1a	pGEX-4T-2 (BamHI/EcoRI)	G2, GST fusion protein
RDB08435	pGEX-mG2-1a	pGEX-4T-2 (BamHI/EcoRI)	mG2, GST fusion protein
Galectin-3			
RDB04219	pGEX-G3-1a	pGEX-4T-2 (EcoRI/XhoI)	G3, GST fusion protein
RDB08393	pET-G3-1a	pET-11a (NdeI/BamHI)	G3, tag-free protein
RDB08394	pGEX-G3CRD-1a	pGEX-4T-2 (BamHI/EcoRI)	G3 (C-terminal lectin domain/carbohydrate recognition domain, G3CRD) , GST fusion protein
RDB08395	pET-G3CRD-1a	pET-11a (NdeI/BamHI)	G3CRD, tag-free protein
RDB08436	pGEX-mG3-3a	pGEX-4T-2 (BamHI/EcoRI)	mG3, GST fusion protein
Galectin-4			
RDB08396	pGEX-G4-1a	pGEX-4T-2 (EcoRI/XhoI)	G4, GST fusion protein
RDB04220	pTrc-G4-1a	pTrc-HisB (XhoI/EcoRI)	G4, $(His)_6$-containing-tag protein
RDB18238	pGEX-G4Null	pGEX-4T-2 (EcoRI/XhoI)	G4 with a truncated linker peptide (G4Null), GST fusion protein
RDB18237	pGEX-G4NT	pGEX-4T-2 (EcoRI/XhoI)	G4 (N-terminal CRD with a part of the linker peptide, G4NT), GST fusion protein
RDB08398	pGEX-G4CT-1a	pGEX-4T-2 (EcoRI/XhoI)	G4 (C-terminal CRD with a part of the linker peptide, G4CT), GST fusion protein
RDB08437	pGEX-mG4-1a	pGEX-4T-2 (BamHI/EcoRI)	mG4, GST fusion protein
Galectin-7			
RDB04221	pGEX-G7-5a	pGEX-4T-2 (BamHI/XhoI)	G7, GST fusion protein
RDB08399	pET-G7-1a	pET-11a (NdeI/BamHI)	G7, tag-free protein
RDB08400	pGEX-G7R54H-12a	pGEX-4T-2 (BamHI/XhoI)	G7 with an R54H mutation[2)], GST fusion protein
RDB08438	pGEX-mG7-7a	pGEX-4T-2 (BamHI/XhoI)	mG7, GST fusion protein
RDB08439	pET-mG7-1a	pET-11a (NdeI/BamHI)	mG7, tag-free protein
Galectin-8			
RDB04222	pGEX-G8(M)-2a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8 (the predominant isoform of G8 in most cell and tissue types, G8(M)/G8M), GST fusion protein
RDB08404	pET-G8(M)-1a	pET-11a (NdeI/BamHI)[4)]	G8M, tag-free protein
RDB08405	pGEX-G8(L)-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8 with a long linker peptide (G8(L)/G8L), GST fusion protein
RDB08410	pET-G8(L)-12a	pET-11a (NdeI/BamHI)[4)]	G8L, tag-free protein
RDB08401	pGEX-G8(M)R69H-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8M with an R69H mutation[2)], GST fusion protein
RDB08402	pGEX-G8(M)R233H-2a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8M with an R233H mutation[2)], GST fusion protein
RDB08403	pGEX-G8(M)R69,233H-1a	pGEX-4T-2 (BamHI/XhoI)[2)]	G8M with R69,233H mutations[2)], GST fusion protein
RDB08454	pGEX-G8Null-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8 with a truncated linker peptide (G8Null), GST fusion protein
RDB08455	pET-G8Null-5a	pET-11a (NdeI/BamHI)	G8Null, tag-free protein
RDB08411	pGEX-G8NT-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8 (N-terminal CRD with a part of the linker peptide, G8NT), GST fusion protein
RDB08412	pGEX-G8CT-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8 (C-terminal CRD with a part of the linker peptide, G8CT), GST fusion protein
RDB08406	pGEX-G8(L)R69H-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8L with an R69H mutation[2)], GST fusion protein
RDB08407	pGEX-G8(L)R275H-3a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8L with an R275H mutation[2)], GST fusion protein
RDB08408	pGEX-G8(L)R69,275H-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8L with R69,275H mutations[2)], GST fusion protein
RDB08409	pGEX-G8(L)R197A-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8L with an R197A mutation[5)], GST fusion protein
RDB08413	pGEX-G8NCRD-1a	pGEX-4T-2 (BamHI/XhoI)[3)]	G8 (N-terminal CRD, G8NCRD), GST fusion protein

Table 1. Galectin clones deposited in the RIKEN Bioresource Center DNA Bank

RDB08414	pGEX-G8CCRD-1a	pGEX-4T-2 (BamHI/XhoI)[3]	G8 (C-terminal CRD, G8CCRD), GST fusion protein
RDB08440	pGEX-mG8(M)-2a	pGEX-4T-2 (BamHI/XhoI)	mG8 (the predominant isoform of mG8 in most cell and tissue types, mG8(M)/mG8M), GST fusion protein
RDB08441	pET-mG8(M)-1a	pET-11a (NdeI/BamHI)	mG8M, tag-free protein
RDB08456	pET-mG8Null-2a	pET-11a (NdeI/BamHI)	mG8 with a truncated linker peptide (mG8Null), tag-free protein
RDB08445	pGEX-RaG8(M)-1a	pGEX-4T-2 (BamHI/EcoRI)	Rat G8 (the predominant isoform of rat G8 in most cell and tissue types, RaG8(M)/RaG8M), GST fusion protein
RDB08446	pET-RaG8(M)-1a	pET-11a (NdeI/BamHI)	RaG8M, tag-free protein
Galectin-9			
RDB08415	pGEX-G9(S)-1a	pGEX-4T-2 (EcoRI/XhoI)	G9 with a short linker peptide (G9(S)/G9S), GST fusion protein
RDB08416	pET-G9(S)-3a	pET-11a (NdeI/BamHI)[4]	G9S, tag-free protein
RDB04223	pGEX-G9(M)-1a	pGEX-4T-2 (EcoRI/XhoI)	G9 (the predominant isoform of G9 in most cell and tissue types, G9(M)/G9M), GST fusion protein
RDB08420	pET-G9(M)-4a	pET-11a (NdeI/BamHI)[4]	G9M, tag-free protein
RDB08421	pGEX-G9(L)-2a	pGEX-4T-2 (EcoRI/XhoI)	G9 with a long linker peptide (G9(L)/G9L), GST fusion protein
RDB08422	pET-G9(L)-5a	pET-11a (NdeI/BamHI)[4]	G9L, tag-free protein
RDB08417	pGEX-G9(M)R65H-2a	pGEX-4T-2 (EcoRI/XhoI)	G9M with an R65H mutation[2], GST fusion protein
RDB08418	pGEX-G9(M)R239H-1a	pGEX-4T-2 (EcoRI/XhoI)	G9M with an R239H mutation[2], GST fusion protein
RDB08419	pGEX-G9(M)R65,239H-1a	pGEX-4T-2 (EcoRI/XhoI)	G9M with R65,239H mutations[2], GST fusion protein
RDB08451	pET-G9Null-11a	pET-11a (NdeI/BamHI)	G9 with a truncated linker peptide (G9Null), tag-free protein
RDB08452	pET-G9NullR65D-2a	pET-11a (NdeI/BamHI)	G9Null with an R65D mutation[2], tag-free protein
RDB08453	pET-G9NullR212D-3a	pET-11a (NdeI/BamHI)[4]	G9Null with an R212D mutation[2], tag-free protein
RDB15282	pET-ssG9	pET-11a (NdeI/BamHI)	G9Null with a 10-amino acid deletion and a single amino acid mutation at the remaining linker peptide (ssG9), tag-free protein
RDB08423	pGEX-G9NT-1a	pGEX-4T-2 (EcoRI/XhoI)	G9 (N-terminal CRD with a part of the linker peptide, G9NT), GST fusion protein
RDB08424	pGEX-G9CT-1a	pGEX-4T-2 (EcoRI/XhoI)	G9 (C-terminal CRD with a part of the linker peptide, G9CT), GST fusion protein
RDB08425	pGEX-G9NCRD-1a	pGEX-4T-2 (EcoRI/XhoI)	G9 (N-terminal CRD, G9NCRD), GST fusion protein
RDB08426	pGEX-G9CCRD-1a	pGEX-4T-2 (EcoRI/XhoI)	G9 (C-terminal CRD, G9CCRD), GST fusion protein
RDB08442	pGEX-mG9(M)-2a	pGEX-4T-2 (BamHI/EcoRI)	mG9 (the predominant isoform of mG9 in most cell and tissue types (mG9(M)/mG9M), GST fusion protein
RDB08443	pET-mG9(M)-1a	pET-11a (NheI/BamHI)	mG9M, tag-free protein
RDB08444	pET-mG9(L)-3a	pET-11a (NheI/BamHI)	mG9 with a long linker peptide (mG9(L)/mG9L), tag-free protein
RDB08457	pET-mG9Null-5a	pET-11a (NdeI/BamHI)	mG9 with a truncated linker peptide (mG9Null), tag-free protein
RDB08458	pET-mG9NullR64H-1a	pET-11a (NheI/BamHI)[6]	mG9Null with an R64H mutation[2], tag-free protein
RDB08459	pET-mG9NullR211H-2a	pET-11a (NheI/BamHI)[6]	mG9Null with an R211H mutation[2], tag-free protein
RDB08447	pTrc-RaG9(M)-2a	pTrc-HisB (XhoI/PstI)	RaG9 (the predominant isoform of RaG9 in most cell and tissue types (RaG9(M)/RaG9M), $(His)_6$-containing-tag protein
Galectin-10			
RDB08427	pGEX-G10-8a	pGEX-4T-2 (BamHI/EcoRI)	G10, GST fusion protein
RDB04224	pTrc-G10-4a	pTrc-HisB (BglII/EcoRI)	G10, $(His)_6$-containing-tag protein
Galectin-12			
RDB08428	pGEX-G12-1a	pGEX-4T-2 (BamHI/EcoRI)	G12, GST fusion protein
RDB04225	pGEX-LNG12-3	pGEX-4T-2 (BamHI/EcoRI)	G12 with an N-terminally extended NCRD (LNG12), GST fusion protein
RDB08429	pGEX-G12NT-1b	pGEX-4T-2 (BamHI/EcoRI)	G12 (N-terminal CRD with a part of the linker peptide, G12NT), GST fusion protein
RDB08430	pGEX-G12CT-1b	pGEX-4T-2 (BamHI/EcoRI)	G12 (C-terminal CRD with a part of the linker peptide, G12CT), GST fusion protein
Galectin-13			
RDB08431	pGEX-G13-1a	pGEX-4T-2 (BamHI/EcoRI)	G13, GST fusion protein
Galectin-14/PPL13			
RDB08432	pGEX-PPL13-1a	pGEX-4T-2 (BamHI/EcoRI)	G14/PPL13, GST fusion protein

1) Catalog number in RIKEN BioResouce Center DNA bank.
2) The mutations abolish lectin activity of respective CRD almost completely.
3) The insert sequence was amplified using forward and reverse primers tagged with extra 5' BglII and SalI sequences, respectively. The amplified cDNA was digested with BglII and SalI, and the resulting cDNA fragment was inserted into the BamHI-XhoI sites of pGEX-4T-2.
4) The insert sequence was amplified using forward and reverse primers tagged with extra 5' NdeI and BglII sequences, respectively. The amplified cDNA was digested with NdeI and BglII, and the resulting cDNA fragment was inserted into the NdeI-BamHI sites of pET-11a.
5) The mutation abolishes thrombin sensitivity of G8L.
6) The insert sequence was amplified using forward and reverse primers tagged with extra 5' NheI and BglII sequences, respectively. The amplified cDNA was digested with NheI and BglII, and the resulting cDNA fragment was inserted into the NheI-BamHI sites of pET-11a.

Table 2. Molecular properties of galectins

	cDNA insert-encoded protein			pET-11a vector (Tag-free protein)		pGEX-4T-2 vector (GST-fusion protein)		(GST-free protein[2])		Protein concentration (mg/mL, A280 nm = 1.0)[3]		
	Residues	M. W.	pI[1]	Residues	M. W.	Residues	M. W.	Residues	M. W.	Tag-free protein	GST-fusion protein	GST-free protein
Galectin-1												
G1	135	14,715	5.18	←[4]	←	361	41,008	137	14,860	1.69	0.697	1.57
CSG1	135	14,619	5.18	←	←	361	40,911	137	14,763	1.63	0.730	–
mG1	135	14,866	5.17	←	←	361	41,158	137	15,010	1.99	–	–
Galectin-2												
G2	132	14,644	5.92			358	40,936	134	14,788		0.739	1.36
mG2	130	14,880	7.39			356	41,172	132	15,024		–	–
Galectin-3												
G3	250	26,188	8.88	←	←	480	52,844	256	26,696	0.550	0.555	–
G3CRD	136	15,470	9.84	137	15,601	362	41,762	138	15,614	1.11	–	–
mG3	264	27,514	8.60			490	53,806	266	27,658		–	–
Galectin-4												
G4	323	35,940	9.44			553	62,597	329	36,449		–	0.852[5]
G4Null	292	32,609	9.58			522	59,265	298	33,117		–	0.844[5]
G4NT	166	18,709	8.88			396	45,366	172	19,218		–	0.683[5]
G4CT	143	15,783	9.98			373	42,439	149	16,291		0.630	1.38
mG4	326	36,371	9.40			552	62,663	328	36,515		–	–
Galectin-7												
G7	136	15,075	7.54	←	←	362	41,367	138	15,219	1.03	–	–
G7R54H	136	15,056	6.78			362	41,348	138	15,200		–	–
mG7	136	15,202	6.43	←	←	362	41,494	138	15,346	1.38	–	–
Galectin-8												
G8M	317	35,808	8.39	←	←	543	62,100	319	35,952	1.20	0.829	1.25
G8MR69H	317	35,788	7.97			543	62,081	319	35,933		–	–
G8MR233H	317	35,788	7.97			543	62,081	319	35,933		–	–
G8MR69,233H	317	35,769	7.44			543	62,062	319	35,914		–	–
G8L	359	40,397	9.33	←	←	585	66,689	361	40,541	1.33	0.829	–
G8LR69H	359	40,378	9.18			585	66,670	361	40,522		–	–
G8LR275H	359	40,378	9.18			585	66,670	361	40,522		–	–
G8LR69,275H	359	40,359	8.99			585	66,651	361	40,503		–	–
G8LR197A	359	40,312	9.17			585	66,604	361	40,456		–	–
G8Null	291	33,065	9.03	←	←	517	59,357	293	33,209	1.16	–	–
G8NT	168	18,893	8.18			394	45,186	170	19,038		0.661	1.55
G8CT	147	16,690	9.13			373	42,982	149	16,834		0.554	1.04
G8NCRD	153	17,345	9.15			379	43,637	155	17,489		–	–
G8CCRD	134	15,295	8.50			360	41,587	136	15,439		–	–
mG8M	316	36,161	9.25	←	←	542	62,453	318	36,305	–	0.675	1.02
mG8Null	290	33,321	9.27	←	←					–		
RaG8M	316	36,072	9.40	←	←	542	62,364	318	36,216	1.06	0.729	–
Galectin-9												
G9S	311	34,690	8.10	←	←	541	61,347	317	35,199	1.05	–	–
G9M	323	35,887	8.10	←	←	553	62,544	329	36,396	0.824	–	–
G9MR65H	323	35,868	7.78			553	62,525	329	36,377		–	–
G9MR239H	323	35,868	7.78			553	62,525	329	36,377		–	–
G9MR65,239H	323	35,849	7.40			553	62,506	329	36,358		–	–
G9L	355	39,517	9.37	←	←	585	66,174	361	40,026	–	–	–
G9Null	296	33,144	8.10	←	←					1.00		
G9NullR65D	296	33,103	7.36	←	←					–		
G9NullR212D	296	33,103	7.36	←	←					–		
ssG9	284	31,803	8.10	←	←					1.02		
G9NT	167	18,323	7.35			397	44,980	173	18,832		0.661	1.03
G9CT	164	18,476	8.49			394	45,133	170	18,985		0.622	0.981
G9NCRD	148	16,345	7.30			378	43,002	154	16,854		–	–
G9CCRD	146	16,549	8.43			376	43,205	152	17,057		–	–
mG9M[6]	322	36,543	8.75	325	36,833	548	62,835	324	36,687	–	–	–
mG9L[6]	353	40,021	9.56	356	40,310					–		
mG9Null	295	33,453	8.55	←	←					–		
mG9NullR64H[6]	295	33,434	8.35	298	33723					–		
mG9NullR211H[6]	295	33,434	8.35	298	33723					–		
RaG9M	322	36,340	8.56									
Galectin-10												
G10	142	16,481	7.23			368	42,773	144	16,625		0.643	0.853[5]

Table 2. Molecular properties of galectins

Galectin-10												
G10	142	16,481	7.23			368	42,773	144	16,625		0.643	0.853[5)]
Galectin-12												
LNG12	336	37,541	9.26			562	63,833	338	37,685		–	–
G12	314	35,220	9.02			540	61,513	316	35,365		–	–
G12NT	170	19,178	7.87			396	45,470	172	19,322		–	–
G12CT	141	15,612	10.26			367	41,904	143	15,756		–	–
Galectin-13												
G13	139	16,118	5.36			365	42,410	141	16,262		–	1.13
Galectin-14												
G14/PPL13	139	16,113	7.21			365	42,405	141	16,257		–	1.40[5)]
GST												
GST[7)]	224	26,166	6.10							0.549		

1) The theoretical isoelectric point values calculated using GENETYX (GENETYX Corporation).
2) Number of amino acid residues and molecular weight of recombinant proteins after removal of the GST tag with thrombin.
3) Protein concentration (mg/mL) of a solution exhibiting absorbance at 280 nm of 1.0.
4) The same as for the cDNA insert-encoded protein.
5) These preparations contain a significant amount of contaminants, mostly *E. coli* chaperones.
6) The cDNA was inserted in the NheI-BamHI sites of the pET-11a vector.
7) GST moiety obtained on thrombin digestion of GST-fusion proteins.

Table 3. Summary of the yields of human galectins expressed in *E. coli*

Galectin	Temp. (°C)	*E. coli* yield (g, wet weight)	Purified protein			Purified protein yield		
			Volume (mL)	Conc. (mg/mL)	Conc. (μM)	mg	mg/*E. coli* (g)	mg/L culture
Galectin-1 (pET-G1)								
	37	0.70	2.35	1.13	77.0	2.66	3.80	13.3
	37	0.69	2.40	1.22	83.0	2.93	4.24	14.6
	30	1.11	2.40	1.85	126.0	4.45	4.00	22.2
	30	1.00	2.55	1.89	128.5	4.82	4.82	24.1
	20	2.15	3.90	2.67	181.6	10.41	4.84	52.1 (11.3*)
	20	2.02	4.05	2.14	145.3	8.65	4.28	43.3 (13.9**)
CSGalectin-1 (pET-CSG1)								
	20	1.89	2.55	2.01	138.0	5.14	2.72	25.7
	20	1.85	4.00	1.39	95.2	5.56	3.01	27.8
Galectin-2 (pGEX-G2)								
	20	2.04	2.90	0.518	35.5	1.50	0.737	7.51
	20	2.00	3.00	0.505	34.6	1.51	0.757	7.57
Galectin-3 (pET-G3)								
	37	0.67	1.00	0.047	1.78	0.047	0.070	0.23
	37	0.60	1.00	0.031	1.20	0.031	0.052	0.16
	30	0.91	2.30	0.323	12.3	0.743	0.816	3.71
	30	0.86	2.30	0.332	12.7	0.764	0.888	3.82
	20	1.42	3.85	2.40	91.5	9.23	6.50	46.2
	20	1.24	4.00	1.72	65.7	6.89	5.55	34.4
Galectin-4 (pGEX-G4)								
	37	0.71	1.85	0.043	1.19	0.080	0.113	0.40
	37	0.70	2.95	0.048	1.31	0.141	0.201	0.70
	30	0.92	2.90	0.080	2.20	0.232	0.252	1.16
	30	0.94	2.95	0.137	3.77	0.405	0.430	2.02
	20	1.56	2.90	0.174	4.77	0.504	0.323	2.52
	20	1.64	2.95	0.181	4.99	0.535	0.326	2.68
Galectin-7 (pET-G7)								
	37	0.66	4.00	1.03	68.1	4.12	6.24	20.6
	37	0.66	3.95	0.91	60.6	3.61	5.47	18.1
	30	1.00	3.90	1.98	131.0	7.71	7.71	38.6
	30	0.83	3.90	2.11	140.0	8.24	9.93	41.2
	20	1.53	4.00	2.90	192.4	11.62	7.59	58.1
	20	1.44	4.00	2.46	162.9	9.84	6.83	49.2
Galectin-8M (pET-G8M)								
	37	0.56	2.40	0.792	22.1	1.90	3.39	9.50
	37	0.64	2.40	0.736	20.5	1.77	2.76	8.83
	30	0.76	2.50	0.974	27.2	2.44	3.21	12.2
	30	0.84	2.40	1.11	30.9	2.65	3.16	13.3
	20	1.30	3.80	1.98	55.3	7.52	5.79	37.6 (5.85*)
	20	1.32	3.80	2.15	60.0	8.16	6.18	40.8 (8.98**)
Galectin-9S (pET-G9S)								
	37	0.71	1.10	0.012	0.35	0.013	0.018	0.06
	37	0.70	1.10	0.020	0.58	0.022	0.031	0.11
	30	0.98	1.10	0.040	1.15	0.044	0.045	0.22
	30	1.03	2.35	0.059	1.70	0.138	0.134	0.69
	20	2.25	5.40	0.221	6.35	1.19	0.529	5.95
	20	1.90	5.35	0.212	6.11	1.13	0.597	5.67
Galectin-9Null (pET-G9Null)								
	20	1.73	5.50	0.380	11.5	2.09	1.21	10.45
	20	1.92	5.35	0.384	11.6	2.05	1.07	10.27
Galectin-10 (pGEX-G10)								
	20	1.81	2.90	0.155	9.41	0.450	0.249	2.25
	20	1.70	3.00	0.113	6.88	0.340	0.200	1.70
Galectin-12 (pGEX-G12)								
	20	1.16	–	–	–	–	–	–
	20	1.29	–	–	–	–	–	–
Galectin-13 (pGEX-G13)								
	20	1.79	3.10	1.73	106.0	5.36	2.99	26.8
	20	2.02	3.10	2.11	129.0	6.53	3.23	32.7
Galectin-14/PPL13 (pGEX-PPL13)								
	20	1.88	3.00	0.162	9.96	0.487	0.259	2.44
	20	1.98	2.90	0.195	11.9	0.564	0.285	2.82

*** Recovered from the unadsorbed fraction obtained on batch affinity adsorption.**
**** Recoverd from the column wash fraction.**

pET vector

Procedure

1) Inoculate 10 mL of LB-broth containing 100 μg/mL ampicillin with a loop full of frozen glycerol stock BL21(DE3)* cells carrying the expression plasmid.
   * In the case of pET vectors, an expression host containing a chromosomal copy of the T7 RNA polymerase gene under lacUV5 control, e.g., *E. coli* BL21(DE3), must be used.

↓

2) Incubate the inoculum at 37° C overnight with shaking (200 rpm) in a shaking incubator.

↓

3) Dilute 4.4 mL of the overnight culture with 220 mL of 2xYT medium containing 100 μg/mL ampicillin in a 1,000-mL flask.

↓

4) Incubate the inoculum at 37° C for ~ 2 h with shaking until the absorbance at 600 nm ($A_{600\ nm}$) reaches 0.6 - 0.7*.
   * Measurements of absorbance were done on samples collected at intervals by withdrawing culture liquid from the flask.

↓

5) Add 0.22 mL of 0.1 M IPTG (isopropyl- β -thiogalactopyranoside) solution to the culture (final concentration, 0.1 mM).

↓

6) Incubate the culture at 37° C for 2 h/at 30° C for 3 h/at 20° C for 16 h (overnight) with shaking (200 rpm).

↓

7) Transfer 200 mL of the overnight culture to centrifuge tubes and then spin down the cells at 6,000 xg for 10 min.

↓

8) Resuspend the pellet in 36 mL of 10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM PMSF (phenylmethylsulfonyl fluoride)*
   * 10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM PMSF, 1 mM DTT (dithiothreitol) was used for G1 and mG1.

↓

9) Transfer the cell suspension to a cooling cell and then lyse the *E. coli* cells using a sonicator*.
   * When using a rosette cooling cell and a Sonifier 250 (Branson) equipped with a 1/2" diameter horn, sonicate the cell suspension for 2 min x 2 and an additional 1 min (with 1-min intervals between the bursts) with the following settings: % duty cycle, 100; output control, 5.0. Place the cooling cell in an ice-water bath during the sonication.

↓

10) Add 4 mL of 10% (w/v) Triton X-100 to the cell lysate (final concentration, 1%).

↓

11) Mix the lysate for 30 min at 4° C with a magnetic stirrer.

↓

12) Transfer the lysate to a centrifuge tube and then spin down the cell debris at 18,000 xg for 30 min at 4° C.

↓

13) Transfer the supernatant to a 100-mL bottle.

↓

14) Add 2 mL of the lactose-agarose slurry (50%[v/v] in TBS [Tris-buffered saline: 20 mM Tris-HCl, pH 7.5, 0.15 M NaCl]) to the supernatant.

↓

15) Mix the lactose-agarose gel suspension by rotating the bottle for 1 h on a tube rotator*.
   * Do not use a magnetic stirrer to avoid destruction of the gel beads.

↓

16) Transfer the gel suspension to a 50-mL conical tube and then spin down the gel at 1,500 xg for 5 min.

↓

17) Discard the supernatant and then suspend the gel pellet in an appropriate volume of TBS, 0.03% CHAPS (3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate).

↓

18) Pack the gel into a column.

↓

19) Wash the gel with 20 gel-bed volumes of TBS, 0.03% CHAPS.

↓

20) Elute the recombinant protein from the gel with 4 mL (1 mL x 4) of TBS, 0.2 M lactose.

↓

21) Dialyze the eluate against PBS*.
   * 1st, 4 - 5 h (500 mL); 2nd, overnight (1,000 mL); 3rd, 4 - 5 h (500 mL)

↓

22) Transfer the dialyzed solution to a centrifuge tube, and then spin down the insoluble material at 25,000 xg for 20 min.

↓

23) Sterilize the supernatant with a sterile filter (0.2 μm).

↓

24) Store the sterilized solution at 4 ° C.

Fig. 1. Expression and purification protocol for galectins (tag-free proteins).

pGEX vector

Procedure

1) Inoculate 10 mL of LB-broth containing 100 μg/mL ampicillin with a loop full of frozen glycerol stock BL21 cells carrying the expression plasmid.
↓
2) Incubate the inoculum at 37° C overnight with shaking (200 rpm) in a shaking incubator.
↓
3) Dilute 4.4 mL of the overnight culture with 220 mL of 2xYT medium containing 100 μg/mL ampicillin in a 1,000-mL flask.
↓
4) Incubate the inoculum at 37° C for ~ 2 h with shaking until the absorbance at 600 nm ($A_{600\ nm}$) reaches 0.6 - 0.7*.
* Measurements of absorbance were done on samples collected at intervals by withdrawing culture liquid from the flask.
↓
5) Add 0.22 mL of 0.1 M IPTG solution to the culture (final concentration, 0.1 mM).
↓
6) Incubate the culture at 37° C for 2 h/at 30° C for 3 h/at 20° C for 16 h (overnight) with shaking (200 rpm).
↓
7) Transfer 200 mL of the overnight culture to centrifuge tubes and then spin down the cells at 6,000 xg for 10 min.
↓
8) Resuspend the pellet in 36 mL of 10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM PMSF*
* 10 mM Tris-HCl (pH 7.5), 0.5 M NaCl, 1 mM PMSF, 1 mM DTT was used for G1 (GST-G1) and mG1 (GST-mG1).
↓
9) Transfer the cell suspension to a cooling cell and then lyse the *E. coli* cells using a sonicator*.
* When using a rosette cooling cell and a Sonifier 250 (Branson) equipped with a 1/2" diameter horn, sonicate the cell suspension for 2 min x 2 and an additional 1 min (with 1-min intervals between the bursts) with the following settings: % duty cycle, 100; output control, 5.0. Place the cooling cell in an ice-water bath during the sonication.
↓
10) Add 4 mL of 10% (w/v) Triton X-100 to the cell lysate (final concentration, 1%).
↓
11) Mix the lysate for 30 min at 4° C with a magnetic stirrer.
↓
12) Transfer the lysate to a centrifuge tube and then spin down the cell debris at 18,000 xg for 30 min at 4° C.
↓
13) Transfer the supernatant to a 100-mL bottle.
↓
14) Add 2 mL of the glutathione-Sepharose slurry (50%[v/v] in TBS) to the supernatant.
↓
15) Mix the glutathione-Sepharose gel suspension by rotating the bottle for 1 h on a tube rotator*.
* Do not use a magnetic stirrer to avoid destruction of the gel beads.
↓
16) Transfer the gel suspension to a 50-mL conical tube and then spin down the gel at 1,500 xg for 5 min.
↓
17) Discard the supernatant and then suspend the gel pellet in an appropriate volume of TBS, 0.03% CHAPS.
↓
18) Pack the gel into a column.
↓
19) Wash the gel with 20 gel-bed volumes of TBS, 0.03% CHAPS.
↓
20) Wash the gel with 2 gel-bed volumes of PBS.
↓
21) Add 1 mL of thrombin solution (50 u/mL in PBS) to the gel in the column, and then mix.
↓
22) Incubate the gel suspension (in the column) at 25° C for 16 h (overnight) with shaking (120 rpm).
↓
23) Drain the GST-free recombinant protein solution from the column (eluate-1).
↓
24) Elute the GST-free recombinant protein from the gel with 2 mL (1 mL x 2) of PBS (eluate-2).
↓
25) Dialyze the combined eluates (eluate-1 and eluate-2) against PBS*.
* 1st, 4 - 5 h (500 mL); 2nd, overnight (1,000 mL); 3rd, 4 - 5 h (500 mL)
↓
26) Transfer the dialyzed solution to a centrifuge tube, and then spin down the insoluble material at 25,000 xg for 20 min.
↓
27) Sterilize the supernatant with a sterile filter (0.2 μm).
↓
28) Store the sterilized solution at 4 ° C.

Fig. 2. Expression and purification protocol for galectins (GST-fusion proteins).

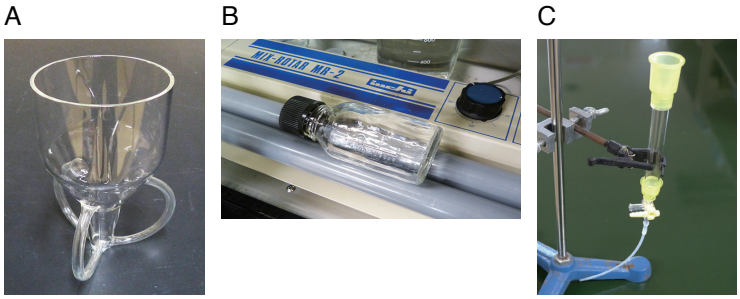

Fig. 3. Laboratory ware used in the purification of recombinant proteins.

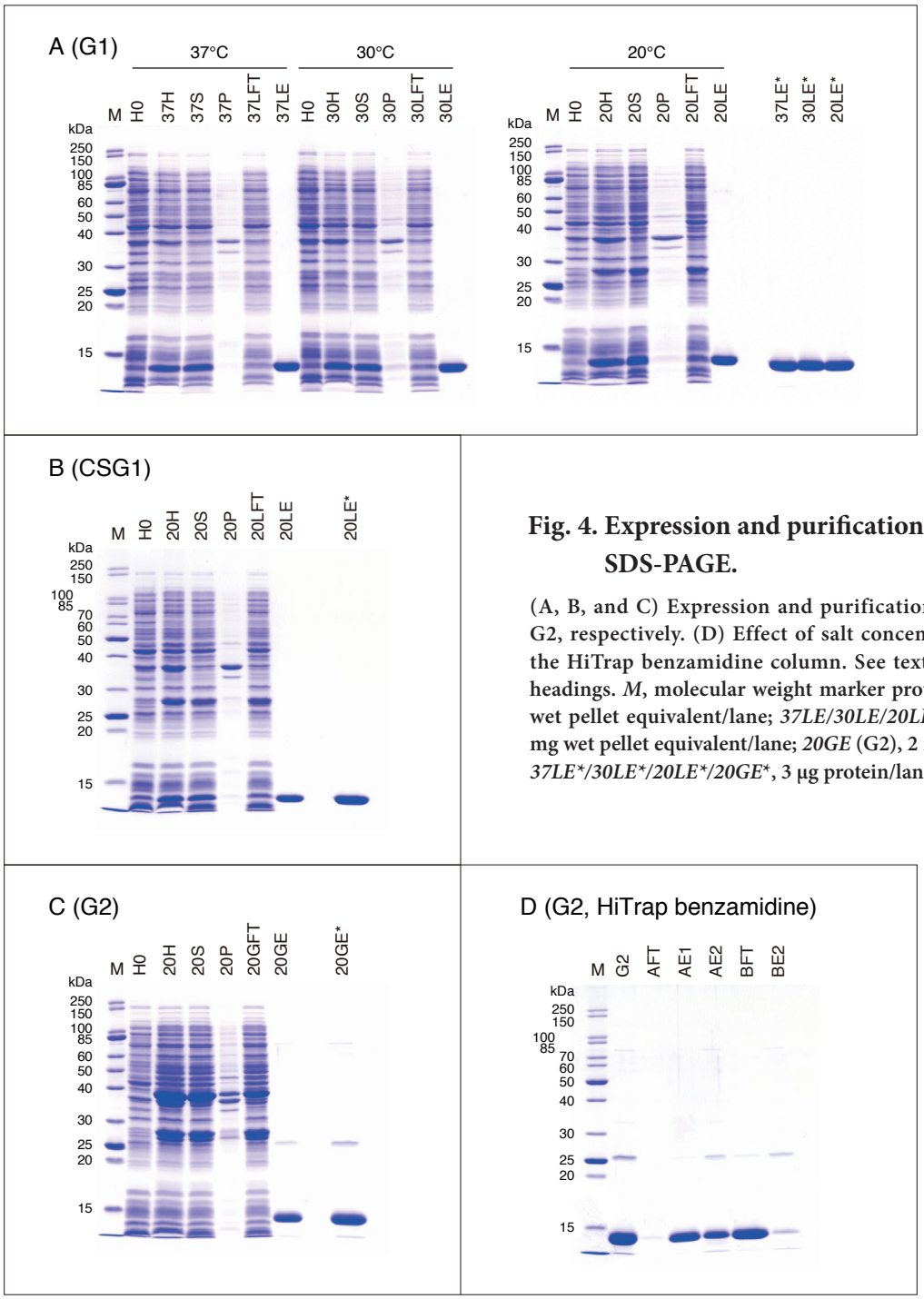

Fig. 4. Expression and purification profiles of galectins on SDS-PAGE.

(A, B, and C) Expression and purification profiles of G1, CSG1, and G2, respectively. (D) Effect of salt concentration on binding of G2 to the HiTrap benzamidine column. See text for explanation of the lane headings. *M*, molecular weight marker proteins. *H0* ~ *LFT/GFT*, 0.1 mg wet pellet equivalent/lane; *37LE/30LE/20LE* (G1) and *20LE* (CSG1), 0.5 mg wet pellet equivalent/lane; *20GE* (G2), 2 mg wet pellet equivalent/lane; *37LE*/30LE*/20LE*/20GE, 3 μg protein/lane.**

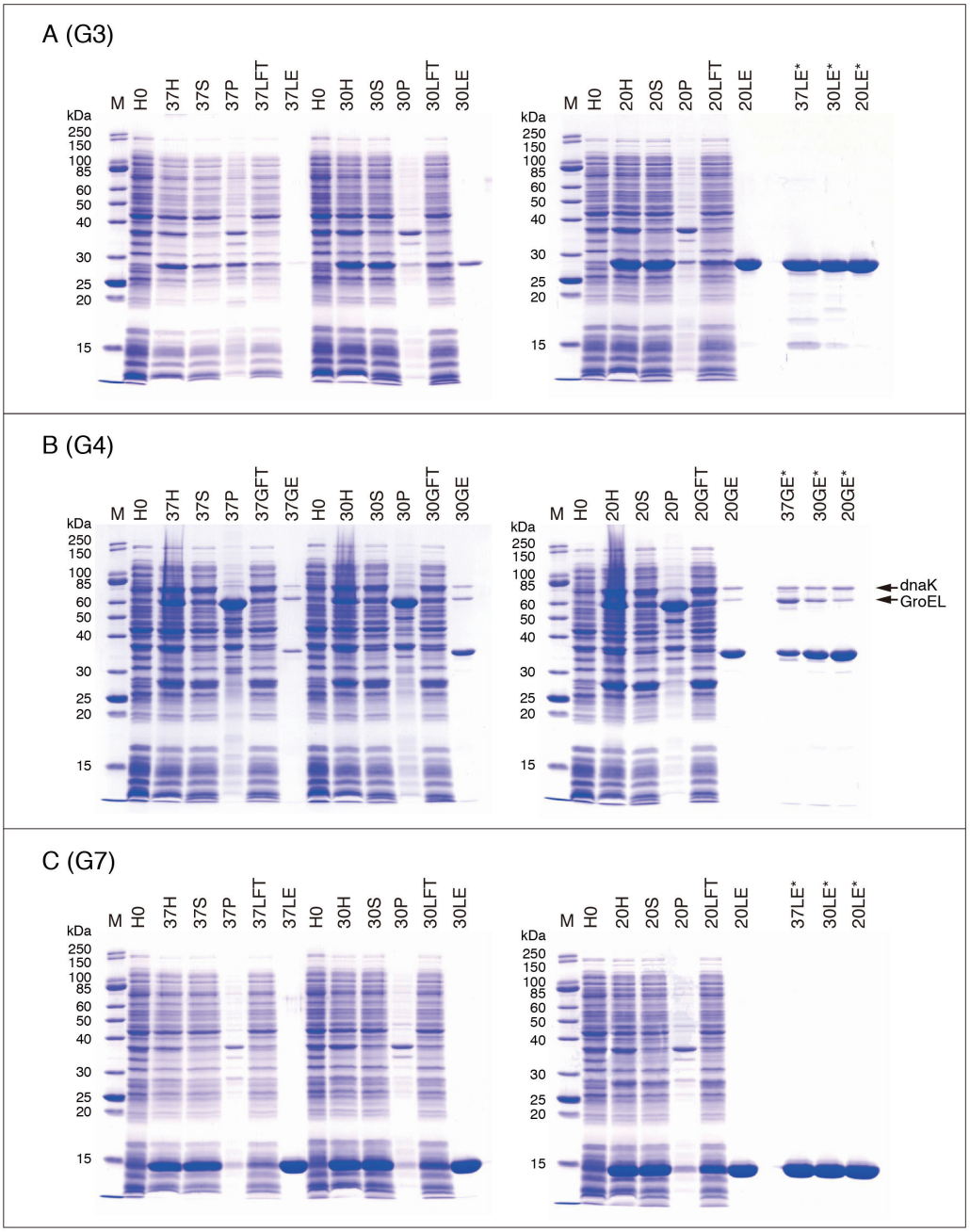

Fig. 5. Expression and purification profiles of galectins on SDS-PAGE.

(A, B, and C) Expression and purification profiles of G3, G4, and G7, respectively. See text for explanation of the lane headings. *M*, molecular weight marker proteins. *H0* ~ *LFT/GFT*, 0.1 mg wet pellet equivalent/lane; *37LE/30LE/20LE* (G3 and G7), 0.3 mg wet pellet equivalent/lane; *37GE/30GE/20GE* (G4), 5 mg wet pellet equivalent/lane; *37LE*/30LE*/20LE*/37GE*/30GE*/20GE, 3 μg protein/lane.**

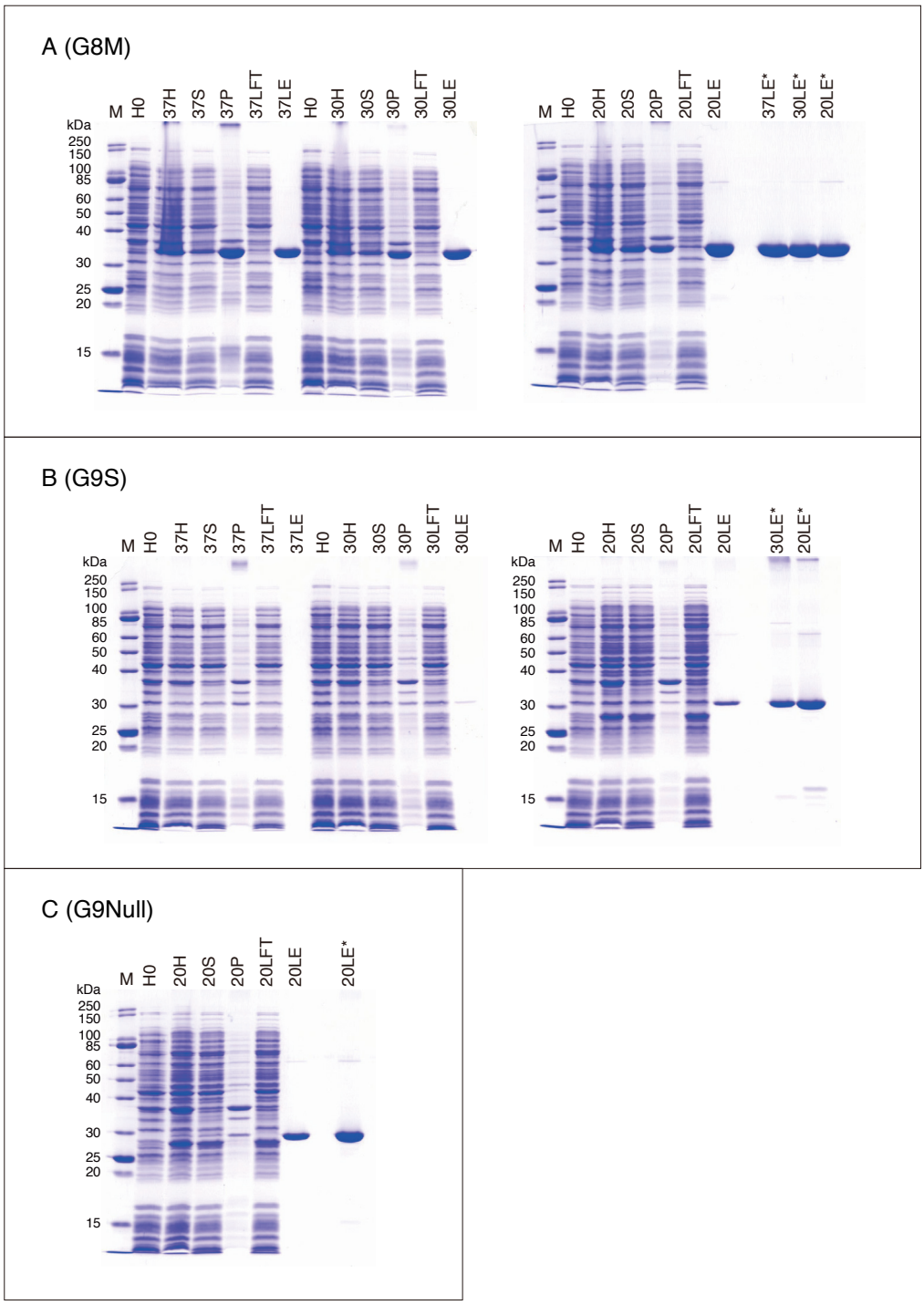

Fig. 6. Expression and purification profiles of galectins on SDS-PAGE.

(A, B, and C) Expression and purification profiles of G8M, G9S, and G9Null, respectively. See text for explanation of the lane headings. *M*, molecular weight marker proteins; *H0 ~ LFT*, 0.1 mg wet pellet equivalent/lane; *37LE/30LE/20LE* (G8M), 0.3 mg wet pellet equivalent/lane; *37LE/30LE/20L*E (G9S) and *20LE* (G9Null), 1 mg wet pellet equivalent/lane; *37LE*/30LE*/20LE**, 3 μg protein/lane.

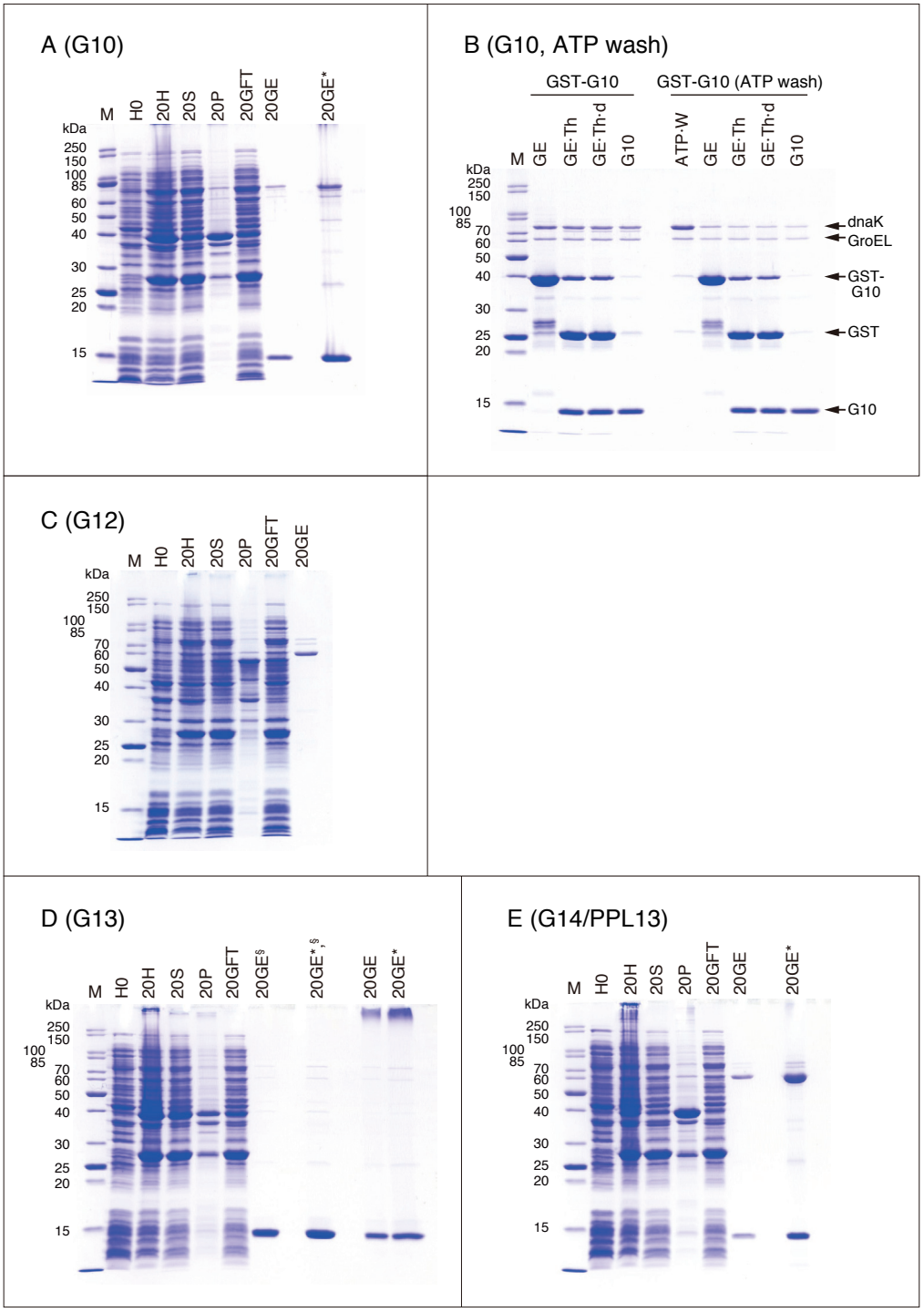

Fig. 7. Expression and purification profiles of galectins on SDS-PAGE.

(A, C, D, and E) Expression and purification profiles of G10, G12, G13, and G14/PPL13, respectively. (B) Effect of ATP washing on the content of *E. coli* chaperones in the G10 preparation. See text for explanation of the lane headings. *M*, molecular weight marker proteins; *H0 ~ GFT*, 0.1 mg wet pellet equivalent/lane; *20GE* (G10,G12, and G14/PPL13), 2 mg wet pellet equivalent/lane; *20GE* (G13), 0.5 mg wet pellet equivalent/lane; *20GE**, 3 μg protein/lane; *20GE*§ , *20GE**,§, heat treatment was not performed.

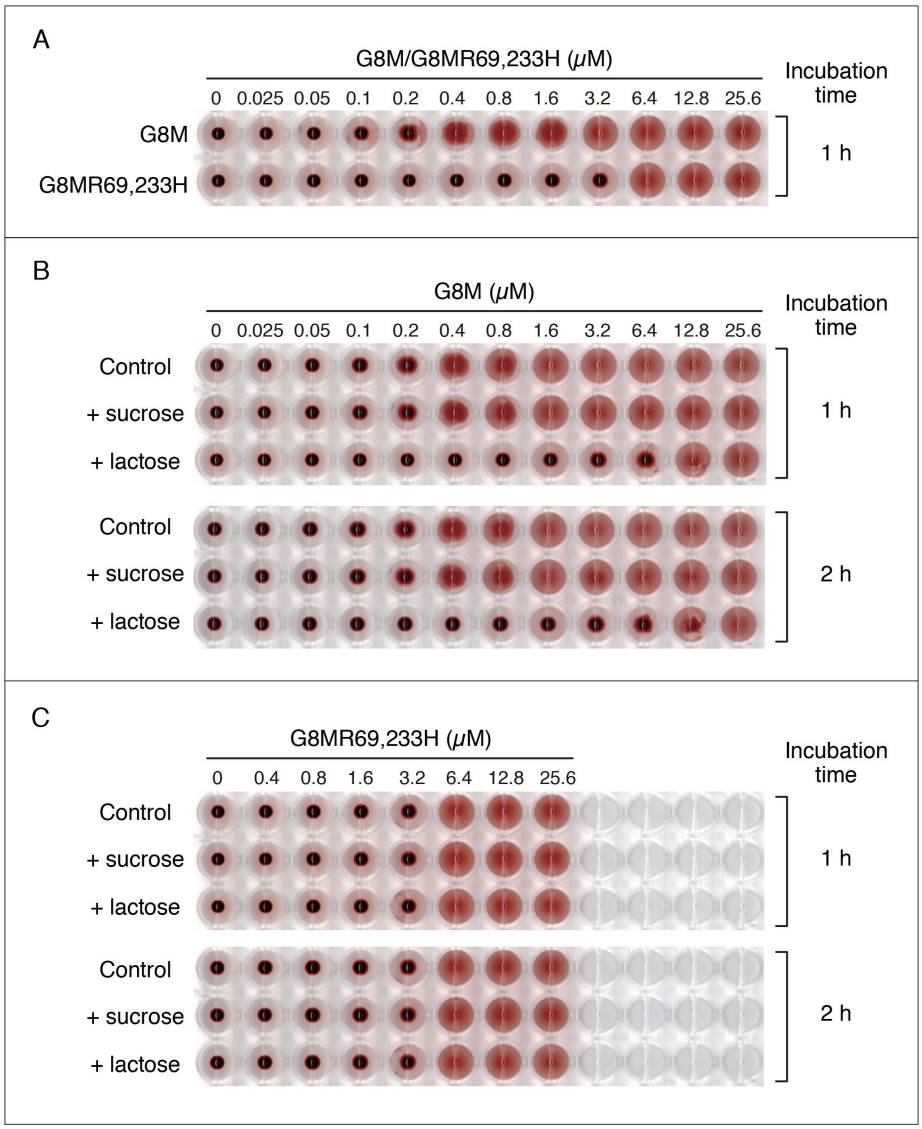

Fig. 8. Hemagglutination assay of G8M and G8MR69,233H.

(A) Twofold serial dilutions (in PBS, 10 mg/ml BSA, 0.05% NaN_3) of purified GST-free G8M and G8MR69,233H were prepared and then added to the wells of V-bottom 96-well plates (50 μL/well). Trypsinized, glutaraldehyde-fixed rabbit erythrocyte suspension (2%[v/v] in PBS, 10 mg/ml BSA, 0.05% NaN_3) was added to each well (50 μL/well) and then mixed by pipetting several times using a multichannel pipette set at 50 μL, followed by 1-h incubation at room temperature. (B and C) Effect of sucrose and lactose on the hemagglutination activity of G8M (*B*) and G8MR69,233H (*C*). Sucrose or lactose was added to both the sample solution and the erythrocyte suspension at a concentration of 50 mM. Hemagglutination was evaluated after 1- and 2-h incubation. The plate was scanned with a flat bed scanner.

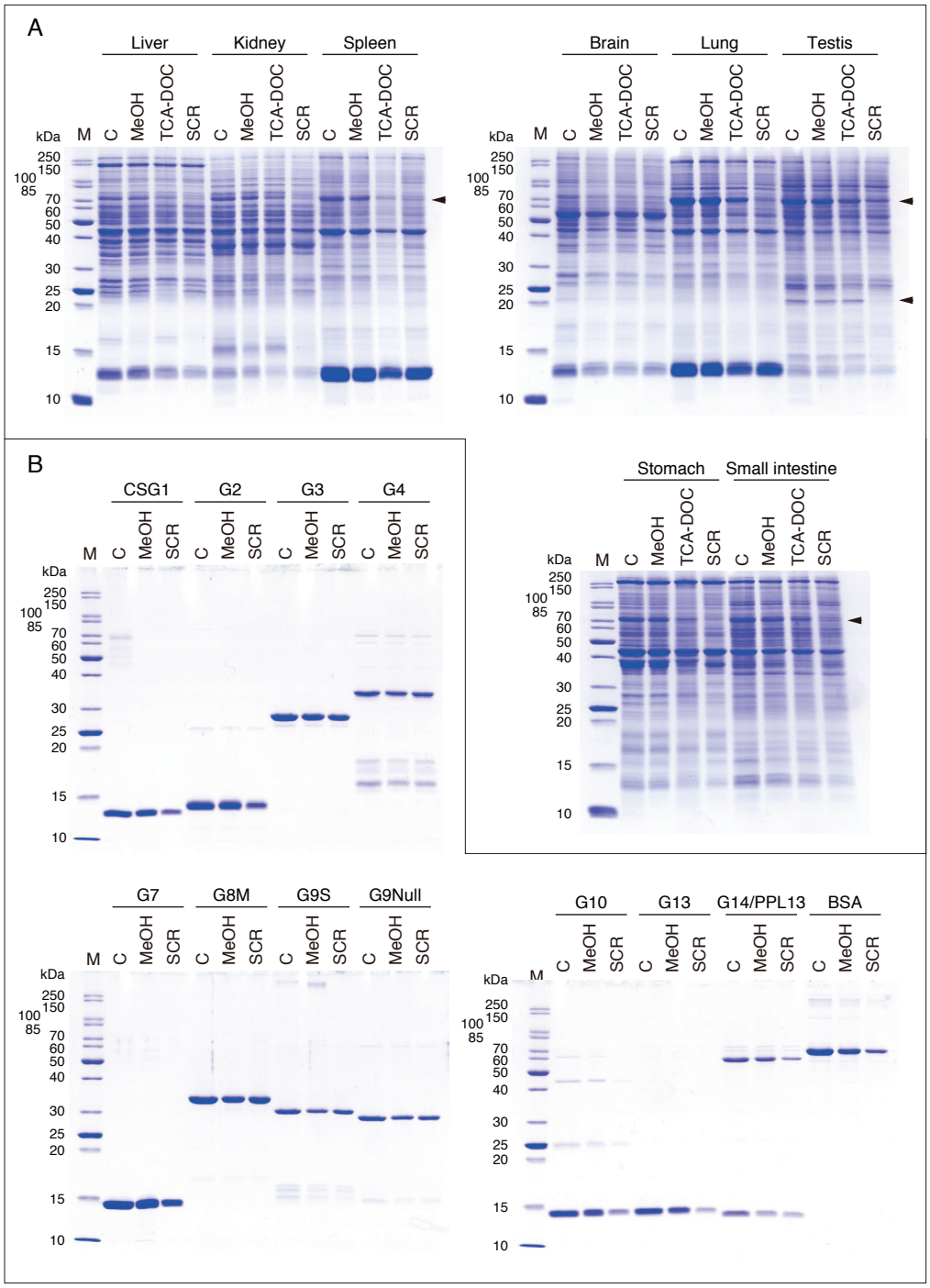

Fig. 9. Comparison of protein enrichment protocols for SDS-PAGE.

(A) Rat tissue extracts concentrated by the three methods (methanol method, TCA-DOC method, and SCR method) were analyzed by SDS-PAGE. (B) Purified galectin preparations and BSA concentrated by two methods (methanol method and SCR method) were analyzed by SDS-PAGE. *M*, molecular weight marker proteins; *C*, 7.5 μg protein/lane (rat tissue extracts) and 1 μg protein/lane (purified galectin preparations and BSA); MeOH, samples concentrated with methanol method; TAC-DOC, samples concentrated with TAC-DOC method; SCR, samples concentrated with SCR method.

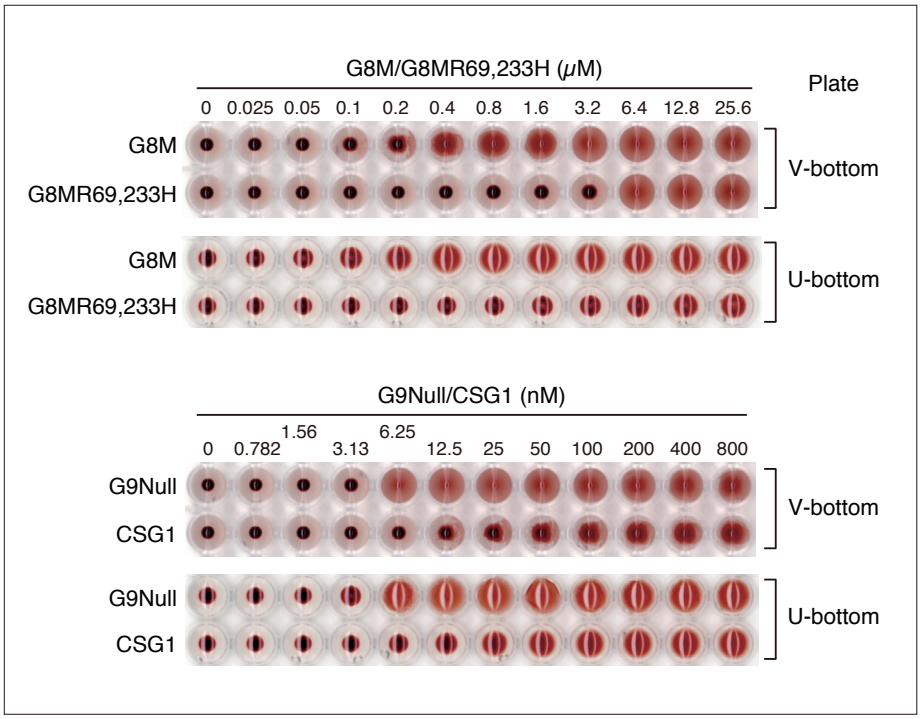

Fig. 10. Hemagglutination assay of G8M, G8MR69,233H, G9Null, and CSG1.

Hemagglutination assay was performed as described in the legend to Fig. 8, except that two types of 96-well plates (V-bottom and U-bottom) were used. G9Null and CSG1 were expressed as tag-free forms (pET vector).

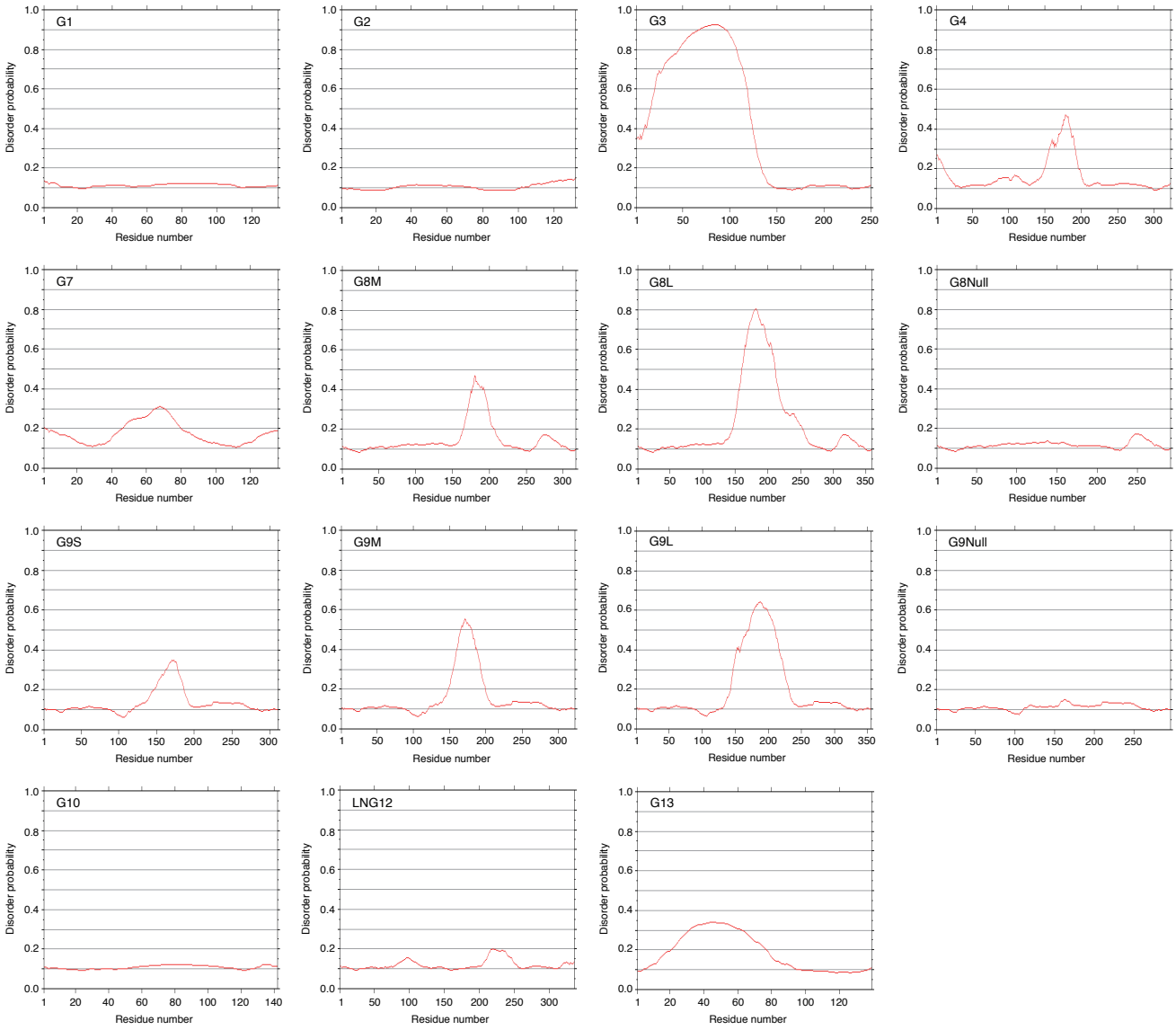

Fig. 11. The prediction of disordered regions of galectins with POODLE.