Journal of Chromatography, 597 (1992) 181–187
Elsevier Science Publishers B.V., Amsterdam

CHROM. 23 904

Arginine-tail method, an affinity tag procedure utilizing anhydrotrypsin agarose

Jun Hirabayashi* and Ken-ichi Kasai

Department of Biological Chemistry, Faculty of Pharmaceutical Sciences, Teikyo University, Sagamiko, Kanagawa 199-01 (Japan)

ABSTRACT

The arginine-tail method is a recently developed affinity tag procedure utilizing immobilized anhydrotrypsin for specific enrichment of a recombinant protein. Three model proteins (originally human β-galactoside-binding lectin with a relative sub-unit molecular mass of 14 000) were prepared by mutagenesis, each of which has a tail of either Arg, Gly–Arg, or Gly–Gly–Arg at the C-terminus. All of them retained their original sugar-binding activity and antigenicity, and became recognizable by anhydrotrypsin. They were adsorbed on an anhydrotrypsin–agarose column at pH 5 after treatment with 4 *M* urea or 10 m*M* HCl to expose the C-terminal tails. The adsorbed arginine-tailed lectins were eluted either specifically with benzoylglycylarginine (Bz–Gly–Arg) or with 5 m*M* HCl. Added arginine was removed by carboxypeptidase B, but very slowly. When urea-denatured *Escherichia coli* lysate containing one of the arginine-tailed lectins was applied to the column, the lectin was adsorbed together with some host proteins, which are expected to have arginine or lysine at their C-termini. However, more than a ten-fold enrichment was attained by this procedure. The described method would be useful for purifying various recombinant proteins, even those which are inactive as a result of mutagenesis. The smallest tag arginine at the C-terminus should have a minimum effect on both the structure and function of a target protein.

INTRODUCTION

An increasing number of affinity techniques called "affinity tag procedures" [1–10] have appeared during the last decade. These procedures are very useful when target proteins have lost their original biological activities such as catalytic activity, binding activity to some substances and antigenicity as a result of mutagenesis. Generally, the procedures consist of fusion of a target protein to another functional, well characterized protein (*e.g.*, β-galactosidase [1,2], protein A [3], glutathione-S-transferase [4] and maltose-binding protein [5]), and specific cleavage to remove these tag partners. As the fused proteins become recognizable by immobilized *p*-aminophenyl-β-thiogalactoside, immunoglobulin G, glutathione and amylose, respectively, specifc isolation of target proteins is greatly facilitated. Although these procedures have proved useful in some instances, it is possible that fusion to such large proteins may impair the structure or function of a target protein. In this respect, procedures utilizing smaller peptides have recently been developed [6–9], *e.g.*, oligocysteine for the use of a thiopropyl column [6], oligohistidine for metalchelate chromatography [7], oligophenylalanine for hydrophobic chromatography [6], oligoarginine for cation-exchange chromatography [8] and a specific haptenic peptide recognizable by Ca^{2+}-dependent antibody [9].

We recently developed a novel procedure utilizing immobilized anhydrotrypsin [10], which is an enzymatically inactive derivative of trypsin but shows a unique property of binding preferentially to product-type compounds [11–13]. Immobilized anhydrotrypsin has also proved to be useful for the isolation of C-terminal tryptic peptides which lack arginine and lysine [14–17]. The developed procedure (termed the arginine-tail method) consists of (i) introduction of an arginine residue at the C-terminus by mutagenesis, (ii) adsorption of the derived arginine-tailed protein on anhydrotrypsin agarose and (iii) elution of the arginine-tailed protein with a specific inhibitor of anhydrotrypsin, benzoylglycyl-

0021-9673/92/$05.00

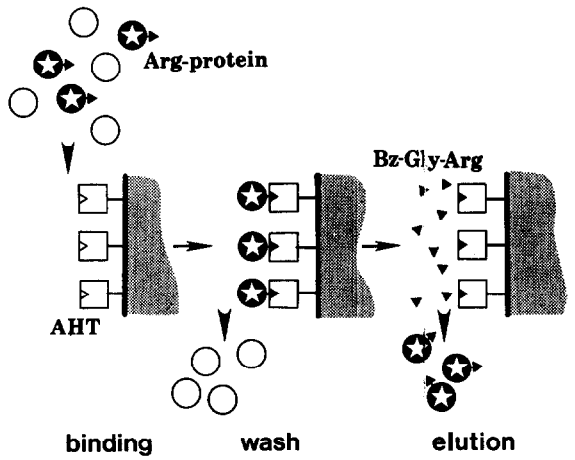

Fig. 1. The arginine-tail method consists of (1) specific adsorption on an anhydrotrypsin column of a target protein to which was added a tag of an arginine residue at the C-terminus, (2) washing of host proteins and (3) specific elution of the arginine-tailed target protein with Bz–Gly–Arg. If necessary, the arginine tail can be removed by carboxypeptidase B treatment. AHT, anhydrotrypsin.

arginine (Bz–Gly–Arg) (Fig. 1). If necessary, the added arginine residue is removable by carboxypeptidase B.

The procedure requires only a single arginine residue at the C-terminus as an affinity tag. In this work, we prepared three arginine-tailed recombinant proteins (human β-galactoside-binding lectin with a relative sub-unit molecular mass of 14 000, H-14 [18–22]) as model proteins, each of which has an Arg, Gly–Arg or Gly–Gly–Arg tail. They became recognizable by anhydrotrypsin retaining the original sugar-binding activity.

EXPERIMENTAL

Construction of arginine-tailed lectin genes and their expression

Construction of an expression plasmid for the mutant human lectin having an arginine residue at the C-terminus (designated H14–Arg) was described previously [10]. A single-stranded M13tv18 DNA, in which an antisense sequence of the H14–Arg gene is inserted, was used as a template for further mutagenesis. To construct mutant genes for lectins having C-terminal Gly–Arg and Gly–Gly–Arg (H14–Gly–Arg and H14–Gly–Gly–Arg, respectively), mutagenesis was performed by using *in vitro* mutagenesis system (Amersham) [23] and mutagenic oligonucleotides, 5′-GCC TTT GAC GGT CGC TGA AAT C-3′ and 5′-GCC TTT GAC GGT GGC CGC TGA AAT C-3′, corresponding to Ala–Phe–Asp–Gly–Arg and Ala–Phe–Asp–Gly–Gly–Arg, respectively (the original C-terminal amino acid sequence is Ala–Phe–Asp). Mutant clones were selected by dot–blot hybridization by using the 5′-^{32}P-labelled mutagenic oligonucleotides according to the manual provided by Amersham. *Bam*HI fragments (500 base pairs) from replicative form DNAs of the positive clones were inserted downstream of *tac* promotor of the original expression vector pUC540 (Kan^{R})[22].

Expression of H14–Gly–Arg and H14–Gly–Gly–Arg was performed essentially as described previously [10] except that *recA* − strain HB101 was used as an expression host instead of Y1090. The arginine-tailed lectins were purified by conventional asialofetuin–agarose column chromatography [18,19], as they retained sugar-binding acitivity. Protein concentration was determined by the method of Bradford [24]. Sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the method of Laemmli [25]. Amino acid compositions of the purified arginine-tailed lectins were analysed with a Hitachi Model 835 amino acid analyser.

Anhydrotrypsin–agarose chromatography

Anhydrotrypsin–agarose (85 nmol of anhydrotrypsin per millilitre of gel) was obtained from Takara (Kyoto, Japan). Chromatography was performed at 6°C with a Pierce disposable polystyrene column (gel bed volume 1 ml) at a flow-rate of 1 ml/min. Arginine-tailed lectins (0.1 or 0.2 mg) were dissolved in 50 m*M* sodium acetate (pH 5.0) containing 20 m*M* $CaCl_2$ (designated initial buffer), and were applied to the anhydrotrypsin–agarose column equilibrated with initial buffer. After extensive washing of the column with initial buffer, the adsorbed proteins were eluted with either initial buffer containing 10 m*M* Bz–Gly–Arg or 5 m*M* HCl.

For urea denaturation, the arginine-tailed lectins (0.2 mg) were dissolved and incubated in 0.2 ml of initial buffer containing 4 *M* urea at 24°C for 30 min prior to application to the column equilibrated with initial buffer.

For acid treatment, H14–Arg (0.1 mg) was dissolved in 0.1 ml of dilute HCl (100, 10 or 1 m*M*), allowed to stand at 0°C for 30 min and then applied to the anhydrotrypsin–agarose column equilibrated with initial buffer.

For chromatography of *Escherichia coli* crude lysates, *E. coli* cells (Y1090) producing H14–Arg were harvested from a 300-ml culture and disrupted by sonication (0°C, 2 min, three times) with initial buffer for chromatography. After centrifugation (15 000 *g*, 4°C, 25 min), the supernatant solution obtained, containing 5.6 mg of protein, was directly loaded on the column. Alternatively, 2.8 mg of the lysate protein were mixed with an equal volume of initial buffer containing 4 *M* urea. After incubation at 24°C for 30 min, the sample was applied to the column equilibrated with initial buffer. Elution was performed with 10 m*M* Bz–Gly–Arg in initial buffer and then with 5 m*M* HCl. Flow-through and adsorbed fractions were subjected to SDS-PAGE and Western-blotting analyses.

Carboxypeptidase B digestion

The added arginine tail of H14–Arg was removed by carboxypeptidase B digestion. Purified H14–Arg (50 μg) was incubated at 37°C with 1 μg of carboxypeptidase B (Worthington Biochemical, Freehold, NJ, USA) in 0.2 ml of 0.1 *M* sodium phosphate buffer (pH 7.0). At appropriate intervals, aliquots were taken and subjected to amino acid analysis.

RESULTS

Preparation of arginine-tailed lectins

Three arginine-tailed lectin genes were constructed by the procedure of site-directed mutagenesis, and were expressed under the control of *tac* promoter in *E. coli* cells. All of the lectins retained their original sugar-binding activity, and therefore could be purified by conventional affinity chromatography on asialofetuin agarose (purification yields 1–2 mg from 1-l of culture). Each of them showed a single band with a relative sub-unit molecular mass of 14 000 in SDS-PAGE (Fig. 2). They were also recognized by specific antiserum raised against wild-type human placenta lectin [18] in Western-blotting analysis (data not shown). These observations suggest that these arginine-tailed lectins retain their original protein structure. They were used as model proteins to evaluate arginine-tail method.

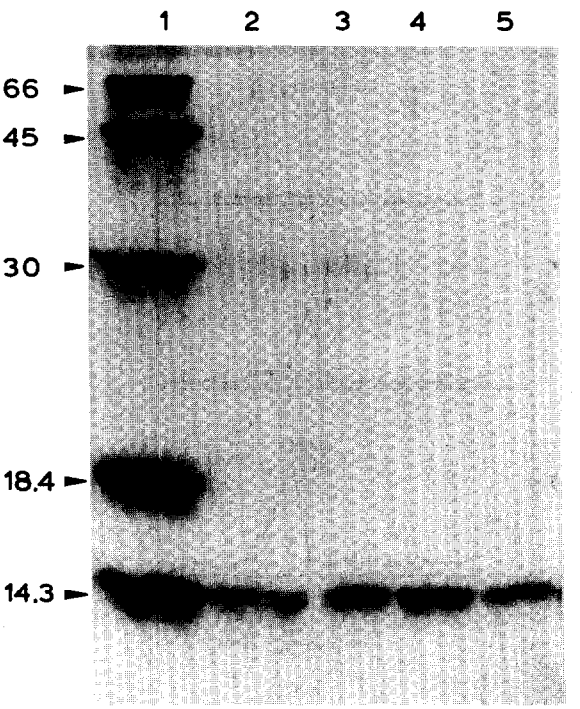

Fig. 2. SDS-PAGE of purified arginine-tailed lectins. Lanes: 1 = marker proteins; 2 = wild-type lectin; 3 = H14–Arg; 4 = H14–Gly–Arg; 5 = H14–Gly–Gly–Arg. Arrowheads indicate the positions of marker proteins; bovine serum albumin (molecular mass 66 · 10^3), chicken ovalbumin (45 · 10^3), bovine carbonic anhydrase B (30 · 10^3), bovine β-lactoglobulin (18.4 · 10^3) and chicken lysozyme (14.3 · 10^3). Protein was stained with Coomassie Blue.

Chromatography of H14–Arg

When H14–Arg was directly applied to an anhydrotrypsin–agarose column, it passed through the column (Fig. 3a). However, it was bound to the column when it was pretreated with 4 *M* urea. In this case, the column had been equilibrated with the initial buffer for chromatography [50 m*M* sodium acetate (pH 5.0)–20 m*M* $CaCl_2$]. The adsorbed lectin was eluted with either 10 m*M* Bz–Gly–Arg (Fig. 3b) or 5 m*M* HCl (Fig. 3c). In contrast, the wild-type lectin which has C-terminal aspartic acid was not adsorbed even after 4 *M* urea treatment (Fig. 3d). H14–Arg eluted with Bz–Gly–Arg was found to have recovered its original activity in terms of heamagglutination of trypsinized rabbit erythrocytes comparable to that of the wild-type lectin (the

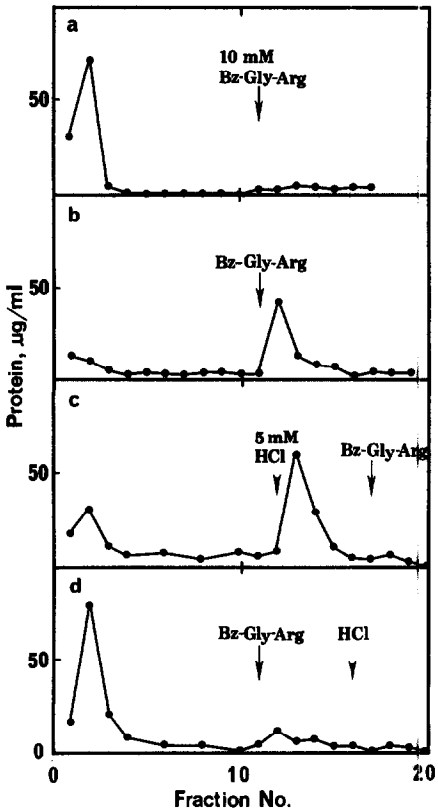

Fig. 3. Anhydrotrypsin–agarose chromatography of purified H14–Arg (a–c) and the wild-type lectin (d). The lectins were dissolved in initial buffer for chromatography [50 m*M* sodium acetace (pH 5)–20 m*M* $CaCl_2$] (a), or the initial buffer containing 4 *M* urea (b–d), and then applied to an anhydrotrypsin–agarose column (bed volume 1 ml). Flow-rate, 1 ml/min; fraction volume, 1 ml. Adsorbed protein was eluted with the initial buffer containing 10 m*M* Bz–Gly–Arg or 5 m*M* HCl. Arrows and arrowheads indicate the starts of the Bz–Gly–Arg and HCl elution, respectively.

minimum concentration required for heamagglutination was 1 μg/ml for each lectin).

We found that acid pretreatment also facilitated the binding. H14–Arg was treated with 100 or 10 m*M* HCl at 0°C for 30 min before application to the column equilibrated with initial buffer. It was bound (Fig. 4b and c) and eluted with 5 m*M* HCl. The acid-denatured H14–Arg was also eluted with 10 m*M* Bz–Gly–Arg (data not shown), suggesting that specific adsorption occured. On the other hand, only about half of the amount of H14–Arg was adsorbed when it was treated with 1 m*M* HCl (Fig. 4d), indicating that the C-terminal region of this protein is fully exposed at pH 3.

Chromatography of H14–Gly–Arg and H14–Gly–Gly–Arg

The above results suggest that the C-terminal arginine of H14–Arg is not accessible to immobilized anhydrotrypsin without denaturation by urea or acid. As H14–Gly–Arg and H14–Gly–Gly–Arg have a longer and more flexible tail than H14–Arg, more facilitated recognition of them by anhydrotrypsin was expected. However, they also failed to be bound to the column under non-denaturing conditions (Fig. 5a and d). It seemed possible that the poor accessibility of the lectins is due to a deformed protein structure during the experiments, because β-galactoside-binding lectins are known to cause oxidative inactivation under non-reducing conditions followed by drastic conformation change [26,27]. However, the addition of 4 m*M* β-mercaptoethanol to the sample solution to prevent oxidation had no effect on the binding (Fig. 5b and e).

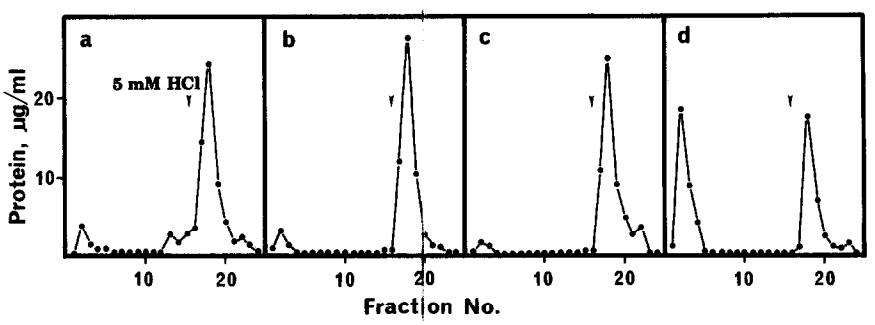

Fig. 4. Effect of acid treatment on adsorption of H14–Arg on anhydrotrypsin–agarose. Purified H14–Arg (0.1 mg) was dissolved in 100 m*M* HCl (pH 1) (b), 10 m*M* HCl (pH 2) (c) or 1 m*M* HCl (pH 3), and was incubated at 0°C for 30 min. They were applied to an anhydrotrypsin column equilibrated with the initial buffer used for chromatography. Adsorbed protein was eluted with 5 m*M* HCl (indicated by arrowheads). For comparison, H14–Arg treated with 4 *M* urea was also analysed (a). Other chromatographic conditions as in Fig. 3.

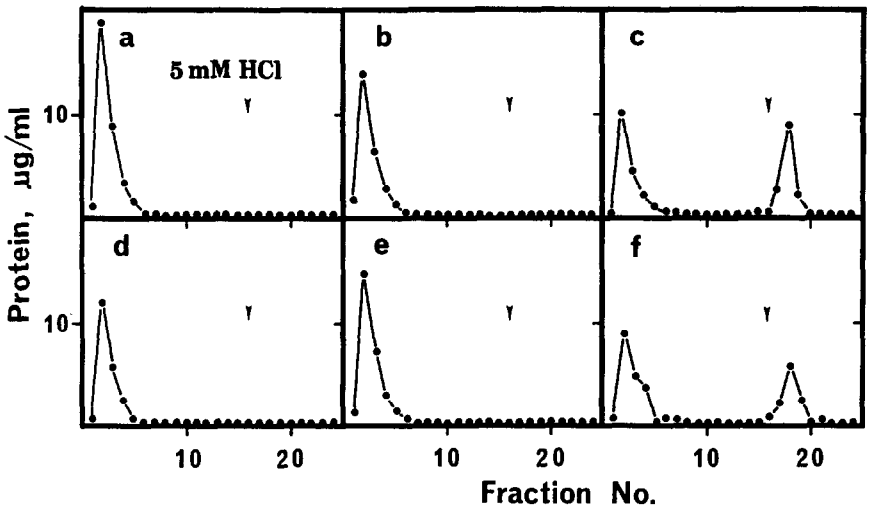

Fig. 5. Anhydrotrypsin–agarose chromatography of purified H14–Gly–Arg (a–c) and H14–Gly–Gly–Arg (d–f). The lectins (0.1 mg) were dissolved in the initial buffer for chromatography and then directly applied to an anhydrotrypsin–agarose column (a,d). Alternatively, the lectins were dissolved in initial buffer containing either 4 m*M* β-mercaptoethanol (b,e) or 4 *M* urea (c,f). After incubation at 24°C for 30 min, they were applied to the column equilibrated with initial buffer. The adsorbed protein was eluted with 5 m*M* HCl (indicated by arrowheads). Other chromatographic conditions as in Fig. 3.

Moreover, the effect of 4 *M* urea in unfolding the polypeptide chains was never complete: about half their amount passed through the column (Fig. 5c and e). The reason for this incomplete binding is not known.

Carboxypeptidase B digestion of H14–Arg

When carboxypeptidase B digestion was performed on the purified H14–Arg as described under Experimental, release of only arginine was observed, but it took a long time to reach stoichiometric removal; the amount of liberated arginine was 0.06, 0.29, and 0.89 mol/mol sub-unit after incubation for 1, 4 and 20 h, respectively. This result suggests that the C-terminal arginine residue of H14–Arg is not fully accessible to carboxypeptidase B.

Enrichment of H14–Arg from E. coli lysate

When expression lysate (5.6 mg) containing H14–Arg was directly applied to the anhydrotrypsin–agarose column, H14–Arg passed through the column together with most *E. coli* proteins (Figs. 6a and 7), consistent with the above result. Some *E. coli* proteins (0.2 mg) were absorbed, which should represent naturally occurring arginine or lysine-tailed *E. coli* proteins. They were eluted with 10 m*M* Bz–Gly–Arg, byt only a small amount of protein was additionally eluted with 5 m*M* HCl, suggesting that the binding is specific. On the other hand, when the lysate (2.8 mg) was applied after 4 *M* urea treatment, H14–Arg was adsorbed on the column. It was eluted with 10 m*M* Bz–Gly–Arg together with some *E. coli* proteins (0.2 mg) (Figs. 6b and 7). We could hardly detect the lectin in the pass-through fraction of the denatured lysate in Western-blotting analysis (data not shown), indicating that H14–Arg was almost completely adsorbed on the column. Thus, more than a ten-fold enrichment of H14–Arg was attained in this procedure.

Some of the adsorbed *E. coli* proteins are common between the denatured and non-denatured samples (Fig. 7). However, there was apparently an increased number of protein bands obtained for the denatured lysate (compare lanes 2 and 4 in Fig. 7). This observation implies that a significant number of *E. coli* proteins have difficulty in the recognition by anhydrotrypsin under non-denaturing conditions. In contrast, there are also some specific bands for the non-denatured sample which are not detected in the denatured sample. They themselves cannot have C-terminal arginine or lysine because no adsorption occurred after denaturation, but they may

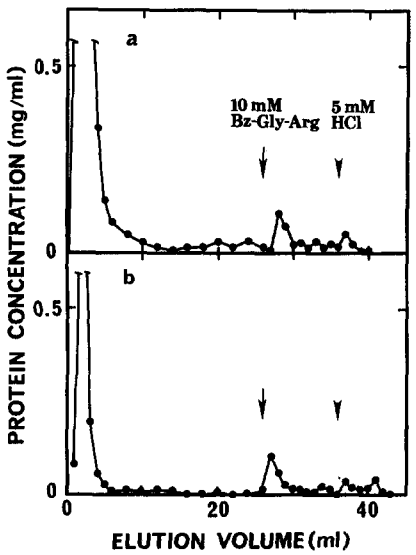

Fig. 6. Anhydrotrypsin–agarose chromatography of *E. coli* expression lysates containing H14–Arg. The lysate containing 5.6 mg of protein was dissolved in initial buffer for chromatography and was directly applied to an anhydrotrypsin-agarose column (a). Alternatively, the lysate containing 2.8 mg of protein was dissvolved and incubated (24 °C, 30 min) in the initial buffer containing 4 *M* urea and then applied to the column equilibrated with initial buffer (b). Adsorbed proteins were first eluted with 10 m*M* Bz–Gly–Arg (indicated by arrows) and then with 5 m*M* HCl (indicated by arrowheads). Other chromatographic conditions as in Fig. 3.

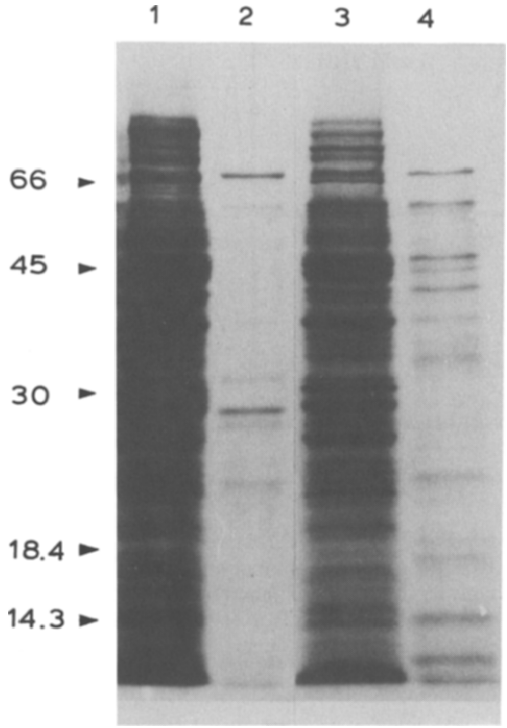

Fig. 7. SDS-PAGE of the flow-through (lanes 1 and 3) and the Bz–Gly–Arg fractions (lanes 2 and 4) obtained in Fig. 6. Lanes: 1 and 2 = samples obtained from the non-denatured lysate (Fig. 6a); 3 and 4 = samples obtained from the denatured lysate (Fig. 6b). Protein was stained with Coomassie Blue. Marker proteins were the same as those used in Fig. 2.

form oligomers under non-denaturing conditions with other subunits having those terminal groups, and hence should be adsorbed without denaturation. Such examples include human haemoglobin β-chain, which lack arginine and lysine at the C-terminus but is adsorbed on anhydrotrypsin–agarose together with α-chain having an arginine tail only under non-denaturing conditions [28].

DISCUSSION

Human β-galacotisde-binding lectin, which was used as a model protein in this work, became recognized by anhydrotrypsin by adding a single arginine residue at the C-terminus (H14–Arg). It was adsorbed on an anhydrotrypsin–agarose column after 4 *M* urea treatment, and was eluted specifically and mildly with 10 m*M* Bz–Gly–Arg. The adsorbed H14–Arg was also eluted with 5 m*M* HCl. Once denatured with urea, H14–Arg recovered its haemagglutinating activity after Bz–Gly–Arg elution comparable to that of the wild-type lectin. Acid pretreatment also facilitated recognition of H14–Arg by anhydrotrypsin as in the case of urea pretreatment. In both instances, anhydrotrypsin retains its binding activity during chromatography, because such denaturants are immediately diluted with the chromatographic buffer used. However, application of a smaller sample volume than the volume of a used column has been recommended [17].

The arginine-tail method should have advantages over previous affinity tag procedures in that the addition of only a single arginine residue at C-terminus may have a minimum effect on protein structure. In fact, the three arginine-tailed lectins described here retained both their original sugar-binding activity and antigenicity. Also, the method is

widely applicable to recombinant proteins, even though they are biologically inactive as a result of mutagenesis. From a practical point of view, the method is easily performed, because it does not require any special expression vector and the addition of arginine is easily attained by conventional mutagenesis procedure. Application of this method to naturally occurring proteins and peptides is also promising. Ishii *et al.* [28] reported that human haemoglobin α-chain, which originally has C-terminal arginine, was recognized by anhydrotrypsin.

In this work, however, neither of the arginine-tailed lectins having longer tails than H14–Arg was adsorbed on anhydrotrypsin agarose unless urea treatment was performed. In addition, it took a long time to liberate the C-terminal arginine from H14–Arg by carboxypeptidase B digestion. These observations imply that the C-terminal region of human β-galactoside binding lectin is not completely accessible to anhydrotrypsin and carboxypeptidase B. Because the lectin forms a non-covalent dimer under non-denaturing conditions [19,27], it is possible that the C-terminal region of this protein may be involved in the inter-sub-unit interaction, and thus hindered. However, the precise reason for this poor accessibility under non-denaturing conditions should be investigated.

Co-purification of some host proteins is a major disadvantage of this method, as seen in Fig. 7. Although the number of these co-purified proteins is limited, they must be removed for further purification. Hence much remains to be improved in the arginine-tail method. Nevertheless, the method should provide a useful alternative affinity tag procedure because of its simplicity, applicability and harmlessness. The use of immobilized antibody raised against some host proteins can theoretically remove such "background" proteins.

REFERENCES

1 J. Germino, J. Gray, H. Charbonneau, T. Vanaman and D. Bastia, *Proc. Natl. Acad. Sci. USA*, 80 (1983) 6848.
2 A. Ullman, *Gene*, 29 (1984) 27.
3 B. Nilsson, L. Abrahasen and M. Uhlen, *EMBO J.*, 4 (1985) 1075.
4 J. Smith, R. B. Derbyshire, E. Cook, L. Dunthorne, J. Viney, S. J. Brewer, H. M. Sassenfeld and L. D. Bell, *Gene*, 32 (1984) 321.
5 C. Guan, O. Li, P. D. Riggs and H. Inouye, *Gene*, 67 (1988) 21.
6 M. Persson, M. G. Bergstrand, L. Bulow and K. Mosbach, *Anal. Biochem.*, 172 (1988) 330.
7 C. Ljungquist, A. Breitholtz, H. Brink-Nilsson, T. Moks, M. Uhlen and B. Nilsson, *Eur. J. Biochem.*, 186 (1989) 563.
8 H. M. Sassenfeld and S. J. Brewer, *Biotechnology*, 2 (1984) 76.
9 K. S. Prickett, D. C. Amberg and T. P. Hopp, *Biotechniques*, 7 (1989) 580.
10 J. Hirabayashi and K. Kasai, *J. Mol. Recog.*, 3 (1990) 204.
11 H. Ako, R. J. Foster and C. A. Ryan, *Biochem. Biophys. Res. Commun.*, 47 (1972) 1402.
12 K. Kasai and S. Ishii, *J. Biochem. (Tokyo)*, 74 (1973) 631.
13 H. Yokosawa and S. Ishii, *J. Biochem. (Tokyo)*, 81 (1977) 647.
14 H. Yokosawa and S. Ishii, *Biochem. Biophys. Res. Commun.*, 72 (1976) 1443.
15 H. Yokosawa and S. Ishii, *Anal. Biochem.*, 98 (1979) 198.
16 S. Ishii, H. Yokosawa, T. Kumazaki and I. Nakamura, *Methods Enzymol.*, 91 (1983) 378.
17 T. Kumazaki, K. Terasawa and S. Ishii, *J. Biochem. (Tokyo)*, 102 (1987) 1539.
18 J. Hirabayashi and K. Kasai, *Biochem. Biophys. Res. Commun.*, 122 (1984) 938.
19 J. Hirabayashi, H. Kawasaki, K. Suzuki and K. Kasai, *J. Biochem. (Tokyo)*, 101 (1987) 987.
20 J. Hirabayashi and K. Kasai, *J. Biochem. (Tokyo)*, 104 (1988) 1.
21 J. Hirabayashi, H. Ayaki, G. Soma and K. Kasai, *Biochim. Biophys. Acta*, 1008 (1989) 85.
22 J. Hirabayashi, H. Ayaki, G. Soma and K. Kasai, *FEBS Lett.*, 250 (1989) 161.
23 K. L. Nakamaye and F. Eckstein, *Nucleic Acid Res.*, 14, (1980) 9679.
24 M. M. Bradford, *Anal. Biochem.*, 78 (1976) 248.
25 U. K. Laemmli, *Nature (London)*, 227 (1970) 680.
26 P. Whitney, J. T. Powell and G. L. Sanford, *Biochem. J.*, 238 (1986) 683.
27 J. Hirabayashi and K. Kasai, *J. Biol. Chem.*, 266 (1991) 23 648.
28 S. Ishii, T. Kumazaki, A. Fujitani, and K. Terasawa, *Macromol. Chem., Macromol. Symp.*, 17 (1988) 281.