

Purification and cDNA cloning of cytokinin-specific binding protein from mung bean (*Vigna radiata*)

Yasuyuki FUJIMOTO¹, Ryuji NAGATA², Hiroshi FUKASAWA², Keiichi YANO³, Masaki AZUMA⁴, Akihiro IIDA³, Seiji SUGIMOTO³, Koichi SHUDO² and Yuichi HASHIMOTO¹

¹ Institute of Molecular and Cellular Biosciences, University of Tokyo, Japan

² Faculty of Pharmaceutical Sciences, University of Tokyo, Japan

³ Tokyo Research Laboratories, Kyowa Hakko Kogyo Co., Ltd, Japan

⁴ Technical Research Laboratories, Kyowa Hakko Kogyo Co., Ltd, Japan

(Received 22 June 1998) – EJB 98 0809/1

Synthetic urea derivatives such as *N*-phenyl-*N'*-(4-pyridyl)urea (4PU) and *N*-(2-chloro-4-pyridyl)-*N'*-phenylurea (4PU30) have strong cytokinin activities. Using tritiated 4PU30 as a probe, we previously established the presence of a cytokinin-specific binding protein (CSBP) of high affinity (K_a for 4PU30 = $4 \times 10^{10} \text{ M}^{-1}$) in the soluble fraction of etiolated mung bean seedlings [Nagata, R., Kawachi, E., Hashimoto, Y. & Shudo, K. (1993) *Biochem. Biophys. Res. Commun.* 191, 543–549]. In this report, we purified CSBP by the use of 4PU-Sepharose 4B, an affinity gel liganded with 4PU. We determined partial amino acid sequences of CSBP and isolated its cDNA by reverse-transcription (RT) PCR. The cDNA encoded a protein with a calculated molecular mass of 17 kDa. A data base homology search revealed that CSBP is a novel member of a major pollen allergen/pathogenesis-related protein family. Recombinant CSBP was expressed in *Escherichia coli* and was confirmed to bind specifically to cytokinins.

Keywords: cytokinin; binding protein; purification; cDNA cloning; recombinant protein.

Cytokinins are plant growth hormones that control the proliferation and differentiation of plants [1, 2]. Natural cytokinins include purine analogues such as zeatin and *N*⁶-(Δ^2 -isopentenyl)-adenine, which exert potent and specific biological activities at nanomolar concentrations. We have found a series of synthetic urea derivatives with very potent cytokinin activity, such as *N*-phenyl-*N'*-(4-pyridyl)urea (4PU) [3], *N*-(2-chloro-4-pyridyl)-*N'*-phenylurea (4PU30, forchlorfenuron) [3] and *N*-(2,6-dichloro-4-pyridyl)-*N'*-phenylurea (4PU1080) [4]. 4PU30 possesses activity almost equivalent to that of natural purine-type cytokinins, while 4PU1080 possesses much stronger activity. In the tobacco callus growth test, the optimum concentrations of 4PU, 4PU30 and 4PU1080 were 0.5 μM , 4 nM and < 1 nM, respectively [4]. At first glance, the chemical structures of urea-type and purine-type cytokinins are different. However, both urea-type and purine-type cytokinin antagonists inhibited the biological responses of both types of cytokinins [5], suggesting that both structural types of cytokinins share a common molecular mechanism of action. Lineweaver-Burk analysis of the inhibitory effects of these antagonists and the quantitative structure/activity relationships

[6] also suggested that urea-type and purine-type cytokinins have a common site of action. However, despite the potent and specific biological activities of these cytokinins, and the accumulated data on the structure/activity relationships of a variety of cytokinin structures, the molecular mechanisms of their biological actions have not been fully elucidated.

Various kinds of soluble cytokinin-binding proteins have been found in various plant species using purine-type cytokinins as probes [7–14]. However, most of these proteins binds potent cytokinins with association constants (K_a 's) of less than 10^8 M^{-1} , far less than the values expected from their effective concentrations. Moreover, the binding activity of these purine-type cytokinin-binding proteins toward urea-type cytokinins is obscure. We have set the following criteria for a cytokinin-specific binding protein(s): (a) a protein with specific binding activity for cytokinins, with association constants (K_a) that match the effective concentrations, and (b) a protein which binds specifically to both purine- and urea-type cytokinins, as well as their antagonists, in a mutually competitive manner. In order to search for cytokinin-specific binding proteins meeting these criteria, we prepared tritiated 4PU30 with a high specific radioactivity (2.83 TBq/mol). Using this, as well as purine-type cytokinins, we previously showed the presence of binding activity for 4PU30/purine-type cytokinins in the soluble fraction of etiolated mung bean seedlings [15]. The association constants (K_a 's) were of the order of 10^9 – 10^{10} M^{-1} for both types of potent cytokinins [15]. This binding was protease sensitive and subject to competition from many kinds of cytokinins of both types, but not cytokinin ribosides or ineffective analogues, including adenine. These results suggested the existence of a cytokinin-specific binding protein (CSBP) which satisfies our criteria. The binding activity was eluted as a single peak on HPLC with Mono Q, Superose 12 and phenyl Superose columns [15]. Thus, there seemed to be

Correspondence to Y. Hashimoto, Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo, Japan 113-0032

Fax: +81 3 5684 8629.

E-mail: hashimot@imcbns.iam.u-tokyo.ac.jp

Abbreviations. 4PU, *N*-phenyl-*N'*-(4-pyridyl)urea; 4PU30, *N*-(2-chloro-4-pyridyl)-*N'*-phenylurea; RT-PCR, reverse transcription-polymerase chain reaction; GST, glutathione S-transferase; WCE, whole cell extract; DPU, *N,N'*-diphenylurea.

Note. The novel nucleotide sequence data reported in this paper have been submitted to the DDBJ/EMBL/GenBank nucleotide sequence databases and are available under accession numbers AB012218, AB012219 and AB012220. Y. Fujimoto and R. Nagata contributed equally to this publication.

only one kind of CSBP in the soluble fraction. Its molecular mass was estimated to be 21 kDa from the elution position on Superose 12 column chromatography [15].

In this study, we prepared an affinity gel liganded with 4PU for affinity purification of CSBP. By the use of this affinity gel, we purified CSBP until it could be identified as a band of 20 kDa in SDS/PAGE. We determined partial amino acid sequences of CSBP and isolated its cDNA by means of reverse-transcriptase PCR (RT-PCR).

MATERIALS AND METHODS

Binding assay. Generally, assay was performed as described previously [15], except that the concentrations of 4PU30 and [³H]4PU30 were 200 μ M and 1 μ M, respectively, in some experiments.

Preparation of 4PU-Sepharose 4B. 3-(*N'*-(4-Pyridyl)ureido)benzoic acid was synthesized as follows. An aqueous solution of sodium nitrite (3.9 g/10 ml) was added dropwise to an ice-cold solution of isonicotinic acid hydrazide (7.0 g in 70 ml 2 M HCl). The mixture was stirred for 15 min, then neutralized with potassium carbonate, and the resultant precipitate of isonicotinic acid azide was collected by filtration. The azide was dissolved in benzene, and the solution was dried over anhydrous sodium sulfonate, then 7.0 g of 3-aminobenzoic acid was added to it. The mixture was refluxed for 3 h, then cooled to afford 3-(*N'*-(4-pyridyl)ureido)benzoic acid as a precipitate, which was collected by filtration (6.22 g). The collected precipitate was purified by recrystallization from aqueous HCl (pH 1) to give colorless needles [melting point 224–226°C; infrared (KBr): 3250, 2820, 1718, 1638, 1613, 1568, 1506 cm^{-1} ; ¹H-NMR (*d*₆-dimethylsulfoxide, 400 MHz, δ): 7.42 (dd, *J* = 9 Hz, 8 Hz, 1H), 7.46 (d, *J* = 6 Hz, 2H), 7.59 (d, *J* = 8 Hz, 1H), 7.65 (d, *J* = 9 Hz, 1H), 8.14 (s, 1H), 8.37 (d, *J* = 6 Hz, 2H), 9.26 (s, 1H), 9.35 (s, 1H), 12.6 (s, 1H)]. 3-(*N'*-(4-Pyridyl)ureido)benzoic acid thus obtained (1.10 g) was dissolved in 800 ml 50% (by vol.) ethylene glycol in aqueous HCl (pH 4.5) by heating in boiling water. To this solution, 20 ml 10.1% (mass/vol.) *N*-ethyl-*N'*-(3-dimethylaminopropyl)carbodiimide hydrochloride in aqueous HCl (pH 4.5) and 45 ml washed EAH Sepharose 4B were added, and the coupling reaction was carried out with shaking at room temperature for 16 h. After the reaction, the gel was washed successively with 500 ml 50% ethylene glycol in aqueous HCl (pH 4.5), aqueous HCl (pH 4.5), 0.1 M NaOH and deionized water. The gel was further washed three times with 250 ml 0.5 M NaCl in 0.1 M sodium acetate, pH 4.0, and 0.5 M NaCl in 0.1 M Tris/HCl, pH 8.0.

Preparation of whole cell extract. Batches (each about 7 kg) of washed 6-day-old etiolated mung bean seedlings were homogenized in 10 l 20 mM Tris/HCl, pH 8.0, containing 30 mM KCl, 10 mM 2-mercaptoethanol, 1 mM EDTA and 1 mM phenylmethylsulfonyl fluoride, using a Robot-coupe cutting mixer of 30 l scale (R-40B, TK Food Machinery, Inc.). This was performed 14 times on a total of 95 kg etiolated seedlings. Powdered cellulose (KC-flock W-50, Nippon Paper Industries Co.) was added to the homogenates as a filter aid at a final concentration of 2% (mass/vol.) and the solution was filtered through a filter pad (NA-050, Advantec) coated with KC-flock using a pressure module (18-D-4, Nissen Co.). The filtrate was centrifuged at 15000 rpm (50–60 l/h) at room temperature (JA-Q2543-O-3H, Penwalt Co.). The supernatant was further centrifuged at 8000 rpm for 30 min at 4°C (Himac 20 and R-10A rotor, Hitachi Co.). The supernatant thus obtained was filtered through a mixed cellulose ester membrane filter (MF-Millipore; 0.8 μ M pore size). This filtrate (whole cell extract)

was concentrated to 10 l by ultrafiltration with a hollow fiber cartridge (H10P10, Amicon) and the concentrate was again centrifuged at 8000 rpm for 30 min at 4°C (Himac 20 and R-10A rotor, Hitachi Co.) to afford a final supernatant (concentrated whole cell extract).

Purification of CSBP. The concentrated whole cell extract was applied to a column of Q-Sepharose Fast Flow (bed volume 3 l) equilibrated with 10 mM 2-mercaptoethanol; 20 mM Tris/HCl, pH 8.0. Adsorbed materials were eluted with 0.3 M NaCl in the same buffer and fractionated. Each fraction was assayed for cytokinin-specific binding activity. The active fractions were collected and phenylmethylsulfonyl fluoride at a final concentration of 1 mM was added to the fractions. The active fractions were concentrated to 250 ml by ultrafiltration (H10P10, Amicon) and the solution was stocked in –80°C until use. The concentrated solution was incubated with 4PU-Sepharose 4B affinity gel in the presence of 2.5% (by vol.) dimethyl sulfoxide for more than 1 hour at 6–10°C with shaking (50 rpm). The protein/gel volume ratio was adjusted to 2.60–62.4 mg/20 μ l. After the incubation, the gel was washed three times with 20 mM Tris/HCl, pH 8.0, containing 0.6 M NaCl and 2.5% dimethyl sulfoxide. Bound proteins were eluted by incubation with SDS sample buffer for 20 min. For the determination of binding specificity, incubation with 4PU-Sepharose 4B was done in the presence of various competitive reagents at the concentration of 200 μ M (except 4PU80: 20 μ M). These reagents were also added to the washing solution at the same concentration. For further purification, the 4PU-Sepharose 4B eluate was applied to a Mini Prep Cell™. The acrylamide content was 4% for the stacking gel (2.5 cm), and 14% for the separating gel (5.0 cm). Electrophoresis was performed at a constant voltage of 200 V. Elution was monitored by measuring the ultraviolet absorbance at 280 nm, and the eluate was fractionated. Each fraction was analyzed by SDS/PAGE and finally, a fraction containing a 20-kDa protein, was obtained.

Electrophoresis. Electrophoresis on SDS/polyacrylamide slab gels was carried out by the method of Laemmli [16] using Ready Gel J (10–20%, Bio-Rad). Samples were denatured by heating them for 20 min at 75°C in 1% SDS and 2% 2-mercaptoethanol. After electrophoresis, the gels were stained by the method of Fairbanks et al. [17].

Determination of amino acid sequences of partial peptides from CSBP. CSBP purified by the Mini Prep Cell™ was concentrated and separated on a 14% SDS/PAGE gel. The protein band was excised with a laser and digested with lysyl endopeptidase (*Achromobacter* protease I, Wako Chem.) in 0.1 M Tris buffer, pH 9.0, containing 4 M urea at 37°C for 16 h. The resulting peptides were separated by reverse-phase HPLC using a SMART system (Pharmacia Biotech) equipped with a μ RPC C2/C₁₈ column (100 mm×2.1 mm internal diameter, Pharmacia Biotech). The column was developed (0.1 ml/min) with a linear gradient from 1% to 60% (0.67% increase/min) acetonitrile in H₂O, containing 0.065% trifluoroacetic acid. Peptides were fractionated and collected for amino acid sequence analysis by automated Edman sequencing using a model PPSQ-10 protein sequencer (Shimadzu) according to the standard protocols recommended by the supplier.

Isolation of CSBP cDNA. CSBP cDNA was isolated by RT-PCR. mRNA was extracted from 3-day-old etiolated mung bean seedlings using a QuickPrep Micro mRNA Purification Kit (Pharmacia Biotech). First, we amplified a DNA fragment by means of the 3' RACE method as follows. Reverse transcription was performed using this mRNA as a template and the synthetic nucleotide 5'-AACTGGAAGAATTCGCGGCCGCAGGAAT₁₈-3' as a reverse transcription primer (Ready·To·Go T-primed First-Strand Kit; Pharmacia Biotech). PCR was done using this

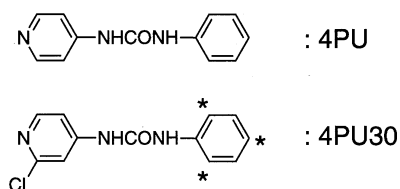
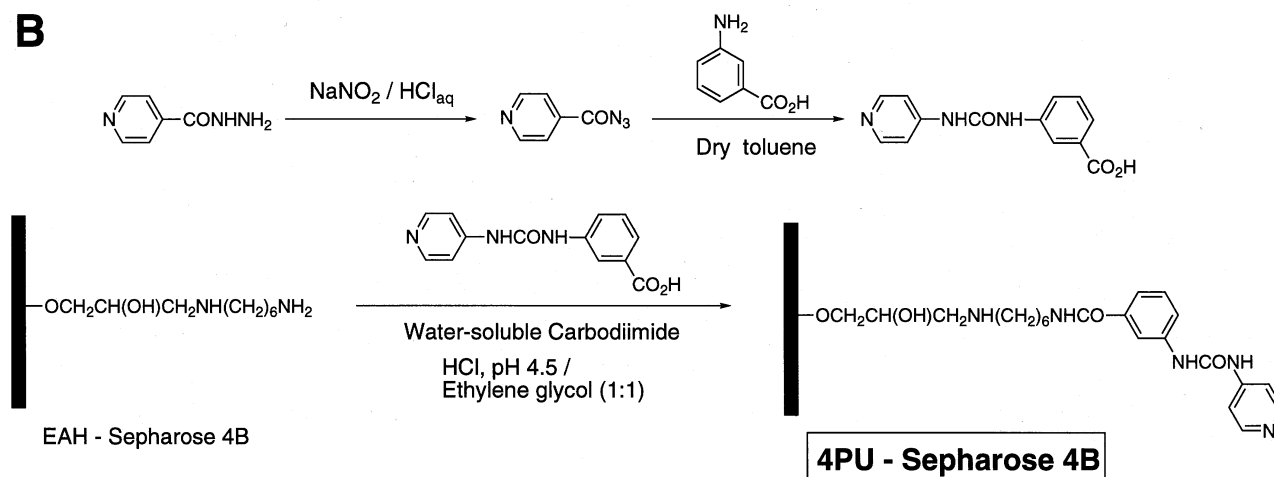
A**B**

Fig. 1. Preparation of 4PU-Sepharose. (A) Structure of 4PU and 4PU30. In $[^3\text{H}]$ 4PU30, tritium was introduced at the positions indicated by asterisks. (B) Scheme of preparation of 4PU-Sepharose 4B. 3-(*N*-(4-Pyridyl)ureido)benzoic acid was synthesized and coupled to EAH-Sepharose 4B.

first-strand cDNA as a template. As primers for first amplification, 5'-ATHTCITAYGAYCAYGAYTCIGAYATHGA-3' (designed based on the partial peptide ISYDHDSDIEEK) and 5'-TGGAAGAATTCGCGGCCGCAG-3' (corresponding to a part of the reverse transcription primer) were used. For nested amplification, 5'-TAYGAYCAYGAYTCIGAYATHGARGARAA-3' (designed based on the same partial peptide) and 5'-TGGAAGAATTCGCGGCCGCAG-3' (the same as shown above) were used. The amplified DNA fragment was cloned with a TA cloning kit (Invitrogen) and sequenced. This DNA fragment encoded the carboxy-terminal region of the 20-kDa protein. Based on the determined nucleotide sequence, we further extended the cDNA in the 5'-upstream direction by means of the 5' RACE method as follows. Double-stranded cDNA was synthesized using a TimeSaver cDNA Synthesis Kit (Pharmacia Biotech) and the Marathon cDNA adaptor was ligated to it [MarathonTM cDNA Amplification Kit (Clontech, USA)]. PCR was done using this double-stranded cDNA as a template. For first amplification, Marathon adaptor primer 1 and 5'-TCAAGCAGAGCCATTCGACAAGTATCTTTC-3' (corresponding to the amino acid sequence ERYLSNGSA at the carboxy-terminus of the 20-kDa protein) were used. For nested amplification, Marathon adaptor primer 1 and 5'-TAGACGCCTAAGATACATTAAAGTGGACTGTG-3' (corresponding to the amino acid sequence SQSTLMYLRL in the carboxy-terminus of the 20-kDa protein) were used. The amplified DNA fragment was cloned and sequenced. DNA sequencing was done by the dideoxy chain termination method of Sanger et al. [18] using an ABI 373A DNA sequencer with a Taq DyeDeoxy terminator cycle sequencing kit (Applied Biosystems). The nucleotide sequences of both strands were determined.

Isolation of CSBP genomic DNA. Genomic DNA of CSBP was isolated by an inverse PCR method [19, 20] as follows. Genomic DNA was extracted from etiolated mung bean seedlings by the method of Murray and Thompson [21]. This genomic DNA was digested with *Bam*HI, *Bgl*II, *Hind*III, *Kpn*I, *Sac*I, *Nco*I, *Nhe*I (Boehringer Mannheim) or *Pst*I (Takara Shuzo Co.). Each digest was self-ligated by T4 DNA ligase (Takara Shuzo Co.). PCR was done using these genomic DNA fragments as templates. Primers used were 5'-GGAACCACTGTTATGAAATCCTTAGATAGCAC-3' and 5'-TCTCTTACGACCATGATTCCGACATAGAAG-3'. The amplified DNA fragment was cloned and sequenced.

Preparation of recombinant CSBP fused with glutathione S-transferase (GST). The open reading frame of CSBP was amplified by PCR using 5'-CGGGATCCCATATGGTAAAGGAATCAACACCCAAAC-3' and 5'-CCGCTCGAGTCAAGCAGAGCCATTCGACAAGTATCTTTC-3'. The former and latter primers contain a *Bam*HI site and a *Xho*I site, respectively. The amplified DNA fragment was double-digested with *Bam*HI and *Xho*I and ligated into the pGEX-4T3 vector (Pharmacia Biotech), which was also double-digested with *Bam*HI and *Xho*I. This construct was cloned and transformed into *Escherichia coli* strain BL21(DE3) LysS (Novagen). After induction of expression with isopropyl- β -D-thiogalactopyranoside, bacteria were recovered and lysed by sonication. The supernatant of centrifugation at $17\,000\times g$ was subjected to binding assay.

RESULTS

Purification of CSBP. For the preliminary study, we attempted to purify CSBP from whole cell extract prepared from 1 kg etio-

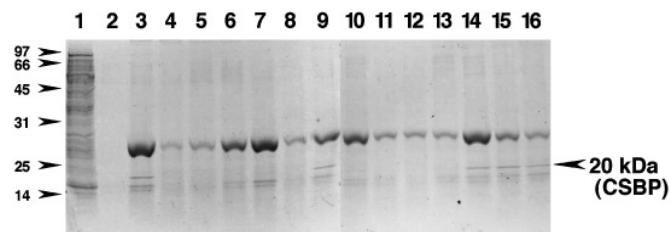


Fig. 2. SDS/PAGE analysis of proteins bound to 4PU-Sepharose 4B in the presence of various competitors. The eluate from Q-Sepharose Fast Flow was incubated with 4PU-Sepharose 4B in the presence of various cytokinins and related compounds. Proteins bound to 4PU-Sepharose 4B were eluted with SDS sample buffer and analyzed by SDS/PAGE (stained with Coomassie brilliant blue). Lane 1, applied sample; lane 2, eluate from Sepharose 4B (base gel without affinity ligand, as control); lanes 3–16, eluate from 4PU-Sepharose 4B. Competitors were: lane 3, without competitor; lane 4, 4PU; lane 5, 4PU30; lane 6, 4PU1080; lane 7, *N,N'*-diphenylurea; lane 8, thidiazuron [28]; lane 9, adenine; lane 10, zeatin; lane 11, *N'*-benzyladenine; lane 12, *N*⁶-(*A*²-isopentenyl)adenine; lane 13, kinetin (*N*⁶-furfuryl)adenine; lane 14, zeatin riboside; lane 15, *N*⁶-benzyladenine riboside; lane 16, *N*⁶-(*A*²-isopentenyl)adenine riboside. The position of the 20-kDa protein is shown by an arrowhead. Molecular-mass markers are also shown (values in kDa).

lated mung bean seedlings by means of a combination of column chromatographies, including Q-Sepharose Fast Flow, Butyl-Toyopearl, Mono Q, Superose HR and Heparin column chromatographies, because these columns were shown to be effective in our previous studies [15]. By this method, CSBP could be purified approximately 200-fold, but the recovery after these steps was as low as 0.5%. This seemed insufficient for further purification, because the content of CSBP in the seedlings is very low [15].

Therefore, we improved the purification procedures in two respects. First, we increased the amount of starting materials to 100 kg. Second, to simplify the purification, we adopted affinity chromatography using 4PU as an affinity ligand (Fig. 1).

Whole cell extract from 100 kg etiolated seedlings was applied to a Q-Sepharose Fast Flow column and the binding activity was recovered by stepwise elution. The pooled active fraction was incubated with the 4PU-Sepharose 4B affinity gel, resulting in quantitative absorption of the binding activity on the gel. The eluate from the affinity gel contained several kinds of proteins (lane 3 in Fig. 2). Of these proteins, only a protein of 20 kDa seemed to bind specifically, because the binding of this 20-kDa protein was subject to competition by both urea-type cytokinins (lanes 4–8 in Fig. 2) and purine-type cytokinins (lanes 10–13 in Fig. 2), but not by cytokinin ribosides or adenine (lanes 14–16 and lane 9, respectively, in Fig. 2). The pattern of competition

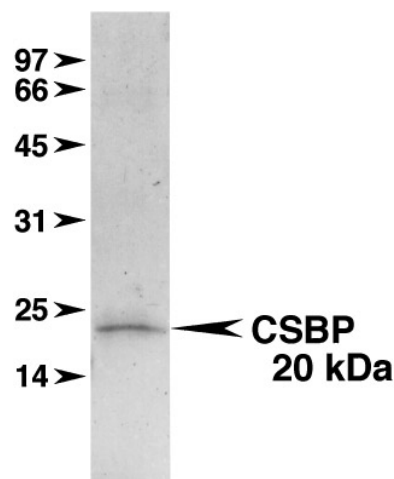


Fig. 3. SDS/PAGE analysis of purified CSBP. CSBP eluted from the Mini Prep Cell was analyzed by SDS/PAGE (stained with Coomassie brilliant blue). The position corresponding to 20 kDa is shown by an arrowhead. Molecular-mass markers are also shown (values in kDa).

precisely coincided with the result of binding assay previously performed using whole cell extract and [³H]4PU30 [15]. This was not the case for the other proteins bound by the affinity gel. For example (Fig. 2, lane 7), the binding of the 20-kDa protein was blocked by *N,N'*-diphenylurea, whereas that of other proteins was not. Moreover, the molecular mass of the protein is consistent with the value of 21 kDa, previously predicted for CSBP on the basis of Superose-12 column chromatography under nondenaturing conditions. Thus, we concluded this 20-kDa protein to be the cytokinin-specific binding protein (CSBP).

For determination of the amino acid sequence of this 20-kDa protein, we further purified it using a Mini Prep Cell (Fig. 3). We finally obtained 20 µg 20-kDa protein (Table 1).

Isolation of cDNA for CSBP. The purified 20-kDa protein was digested with lysyl endopeptidase, and the amino acid sequences of three partial peptides were determined. The determined sequences were ISYDHDSIEEK, VLPPIVK and XVXXIEGDGGV (X could not be unequivocally determined). Based on these amino acid sequences, we designed several degenerate oligonucleotide primers for RT-PCR analysis. Using primers corresponding to the peptide ISYDHDSIEEK and a portion of the reverse-transcription primer, we succeeded in amplifying a DNA fragment which seemed to encode the carboxy-terminal region of 20-kDa protein. We further extended this cDNA in the 5'-upstream direction, using the 5' RACE method. 26 independently amplified PCR products showed similar nucleotide

Table 1. Summary of purification of CSBP. At the steps of whole cell extract, concentrated whole cell extract and Q-Sepharose Fast Flow, binding assay was performed. Amount of CSBP was determined as the amount of [³H]4PU30 specifically bound to CSBP, assuming that binding ratio is 1/1 in 4PU30/CSBP complex. At the steps of 4PU-Sepharose 4B and Mini Prep Cell, binding assay was not performed because proteins recovered in these steps were denatured with SDS (see Procedures). In these steps, amount of CSBP was determined as the amount of 20-kDa protein. Specific content was defined as (amount of CSBP)/(amount of total protein) in any step.

Step	Protein	Amount of CSBP	Recovery	Specific content	Purification
	g	nmol	%	fmol/mg	-fold
Whole cell extract (WCE)	184	16.8	100	91.1	1.0
Concentrated WCE	122	14.4	86	118	1.3
Q-Sepharose Fast Flow	27.8	13.2	79	475	5.2
4PU-Sepharose 4B	0.00570	2.77	17	486 000	5 300
Mini Prep Cell	0.0000205	1.03	6	50 200 000	551 000

A

```

      10      20      30      40      50      60      70      80      90
TGTTAAGAGCCATGGTAAAGGAATTCAACACCCAAACAGAGCTTAGTGTGAGATTGGAAGCCCTTTGGGAGTGCTATCTAAGGATTTC
  M V K E F N T Q T E L S V R L E A L W A V L S K D F I 27

     100     110     120     130     140     150     160     170     180
TAACAGTGGTTCCAAAGGTACTTCCTCATATTGTGAAAGATGTGCAACTGATTGAAGGTGATGGAGGGGTTGGAACAATCCTCATCTTTA
  T V V P K V L P H I V K D V Q L I E G D G G V G T I L I F N 57

     190     200     210     220     230     240     250     260     270
ATTTTCTGCCGTGAGGTGTCAACCAAGCTACCAAGGGAGGAGATCACAGAGTTTGATGAGAGTTACATGAGATAGGTTTGCAAGTGATTG
  F L P E V S P S Y Q R E E I T E F D E S S H E I G L Q V I E 87

     280     290     300     310     320     330     340     350     360
AAGGAGGTTATCTAAATCAAGGGTGTTCATATTACAGACAACTTTTAACTATCTGAAATTGAAGAGGATAAGACTTTGGTCAACGTGA
  G G Y L N Q G L S Y Y K T T F K L S E I E E D K T L V N V K 117

     370     380     390     400     410     420     430     440     450
AAATCTCTTACGACCATGATTCGACATAGAAGAAAAGGTCACACCTACGAAGACATCAGTCCACTTTAATGTATCTTAGGCGTCTAG
  I S Y D H D S D I E E K V T P T K T S Q S T L M Y L R R L E 147

     460     470     480
AAAGATACTTGTGCAATGGCTCTGCTTGAAG---polyA
  R Y L S N G S A * 155

```

B

```

41 VQLIEG 46      105 SEIEEDKT 112
   *.***          *.*** *
83 LQVIEG 88      124 SDIEEKVT 131

```

C**5'-upstream**

```

-329: TGTATTTATTTGACTTGACGTATATGTTGAAATCTTCTA: -290
-289: CGAAGCAAGTATGATAAGTGAAGCATATATAATATGTTATTTATTTGTAATGATATTGTTTTTAATGTTGTTATTTTACCAGCAGTATT: -200
-199: TTCCTTTAGGGCATCAAACTCTATAATGTATCCAACTTATTTTATTAACACATGTTGCATGAAAGAACATATCCTTCACATAAAAA: -110
-109: ACAAAGCATATTTTTCATATCTATATGTGGATGAGATCATAAAGATTCCCTATCGAAGTACTACATATACACATAGCAGTGCATGTC: -20
-19: TCTGTGTAAGAGCAGGTTTGTTAAGAGCCATGGTAAAGGAATTCAACACCCAAACAGAGCTTAGTGTGAGATTGGAAGCCCTTTGGGCA: 71
      *                               M V K E F N T Q T E L S V R L E A L W A

```

3'-downstream

```

394: AAAAGGTCACACCTACGAAGACATCAGTCCACTTTAATGTATCTTAGGCGTCTAGAAAGATACCTTGTGCAATGGCTCTGCTTGAAGAA: 483
      K V T P T K T S Q S T L M Y L R R L E R Y L S N G S A *
484: AAAATATTTTCGTATTGAAGTGGATAACGATTCCTTTACGGCATTTTTTTATAATATTTTAACATTATTTACGTGTTATTATTTGATTGG: 573
574: TTTATATTTTACGAACATGATATAGACCAATTACAGAATGACATGTAATAATGTTAAATATTTGTTAAAAAACATTGCAAAATATCAT: 663
664: TATCCTGTTAAATATAAAAAATAAAGAATGATTAGTTGTTGTTTTTACTATTTTCTATTATACTTAAGAATATATAATAAAAGG: 753
754: AGAAATGAAATACACTATATTTTCTAAAGCCAAACATCTCCTACTATTGTAATATGTGTGGGTTAGGGTAGTTGCGTCTCTGTTTTTAA: 843
844: TAAAGTTTACTATTTTCAGTCGAGTAAGGAAAATACTGTTAATTGGCCCTCTATAATGTAAGGGTCCCGTTACTTTGACACCCATTTTTT: 933
934: AACACTCATTTGACAACGCTACATGTCAAGTTTCCATTGGTGATTTTAAATTTTGGTTTTTAAATACAAACTTTTCTCGAAAGTGTA: 1023
1024: ATGATTTTCAAGATATTGCATTTCATCTCTATCCCTGCTCT: 1074

```

Fig. 4. Nucleotide sequence of the CSBP gene. (A) Nucleotide sequence of CSBP cDNA and the deduced amino acid sequence. Underlines, partial peptides from CSBP; asterisk, translation termination codon; open letters, short repetitive sequences which are also shown in (B). (C) Nucleotide sequence of CSBP genomic DNA. The regions 5'-upstream and 3'-downstream of the protein coding region were amplified and sequenced. The regions in black-back correspond to the 5'- and 3'-terminal regions of CSBP cDNA. Deduced amino acid sequences are shown below. In-frame stop codons, putative TATA boxes and poly adenylation signals are framed. Nucleotides are numbered continuously, with the positions of the 5'- and 3'-termini of the protein coding region as 12 and 476, respectively.

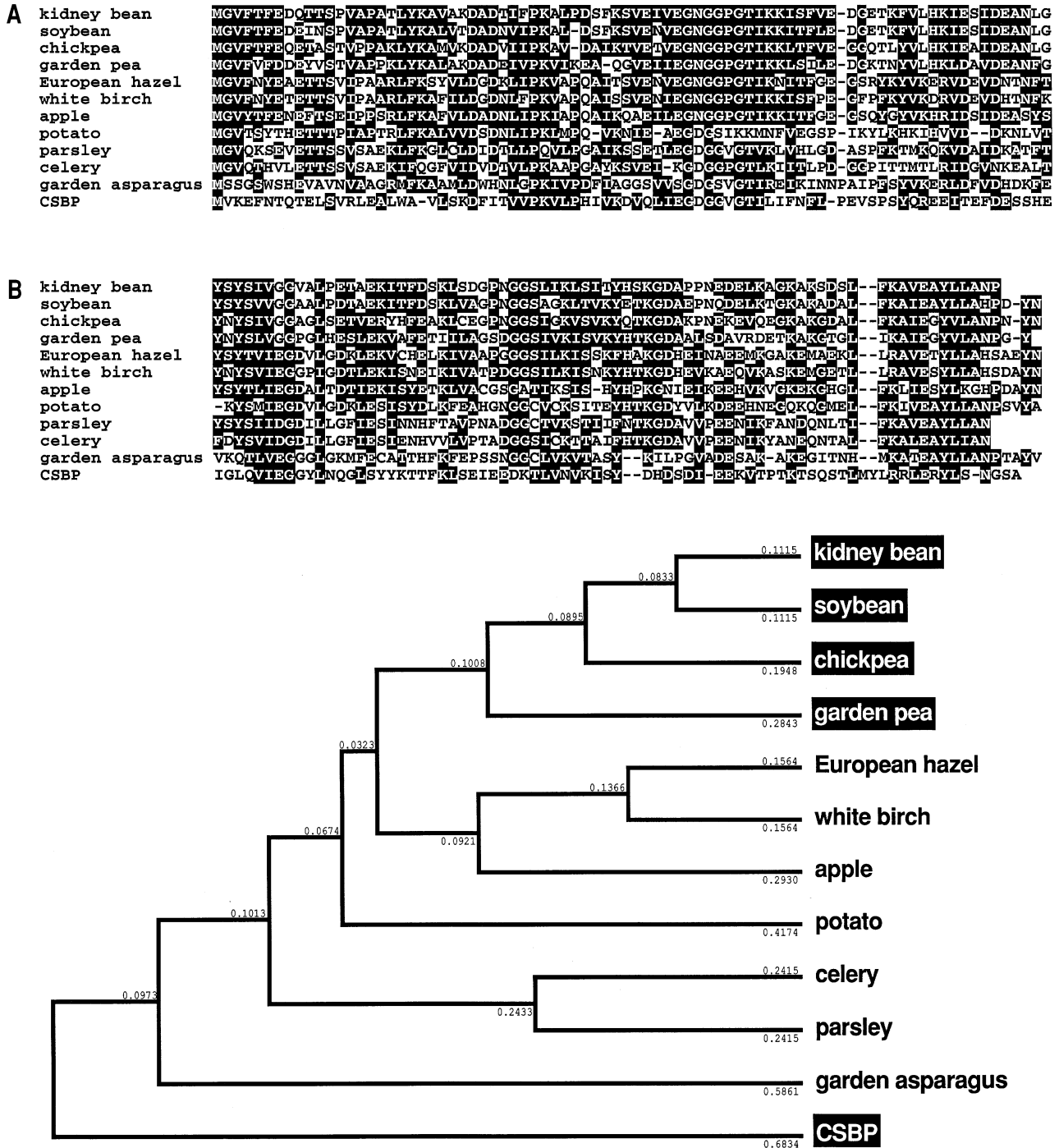


Fig. 5. Comparison of the amino acid sequence of CSBP with those of major pollen allergens and pathogenesis-related proteins from various species. (A) The amino acid sequence of CSBP is aligned with those of major pollen allergens and pathogenesis-related proteins from kidney bean [29], soybean [30], chickpea (pir: S47140), garden pea [31], European hazel [32], European white birch [33], apple [34], potato [35], parsley [36], celery [37], and garden asparagus [38]. Amino acid residues conserved in more than four proteins are shown in black-back. Gaps have been introduced to obtain maximal sequence similarity. (B) Proteins shown in (A) were phylogenetically analyzed based on their sequence similarity using the UPGMA method [39]. Numbers denote genetic distances.

lengths. The nucleotide sequence and deduced amino acid sequence of the longest cDNA are shown in Fig. 4A.

The protein encoded by this cDNA contained all three partial peptides from the 20-kDa protein. Thus, we concluded that this cDNA encoded the 20-kDa protein, i.e. CSBP. We confirmed that a pair of primers designed based on the 5'- and 3'-termini could amplify the sequence in Fig. 4A, and the sequences obtained in two independent amplifications were identical.

A data base homology search revealed that the 20-kDa protein is a novel member of a major pollen allergen/pathogenesis-related protein family (Fig. 5A). In this family, not only amino acid sequences, but also lengths of peptides are well conserved. The most upstream methionine residue detected in CSBP (Fig. 4A) corresponds to the initiating methionine residue of the others (Fig. 5A), so this methionine residue seems to be the initiating methionine of CSBP, although the cDNA contained no

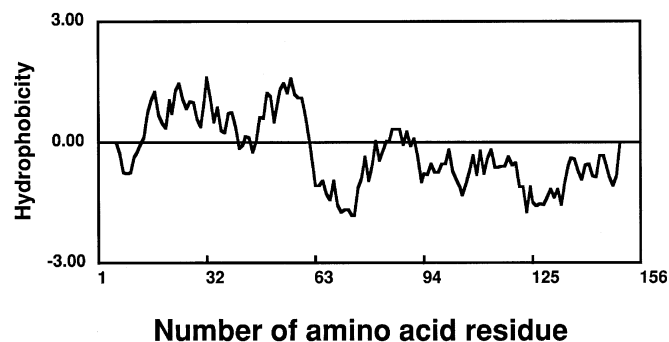


Fig. 6. Hydropathy profiles of CSBP. The distribution of hydrophobic and hydrophilic domains of CSBP was analyzed by the method of Kyte and Doolittle [40]. The span number used for averaging was 13. The numbers of amino acid residues are shown at the bottom of the panel. Hydrophobic and hydrophilic portions are plotted above and below the central horizontal line, respectively.

in-frame stop codon upstream of the methionine residue (Fig. 4A). If this is the initiating methionine residue, the protein encoded by the cDNA consists of 155 amino acid residues. The calculated molecular mass of the protein is 17 kDa which is compatible with the observed value of 20 kDa.

Although CSBP is a member of the above-mentioned protein family, the degree of similarity is not high. In phylogenetic analysis, CSBP is distant even from its counterparts from bean species (Fig. 5B). Two pairs of short repetitive sequences were also found in the 20-kDa protein (Fig. 4A, B). The hydropathy profile showed that the 20-kDa protein contains no amino-terminal signal sequence or obvious transmembrane domains (Fig. 6). This provides support for the idea that the 20-kDa protein is a soluble intracellular protein.

We noticed that the sequence of the obtained cDNA seems not to contain a true poly(A) tail. Although a poly(A) sequence was detected at the 3'-terminus of the cDNA, this poly(A) was almost adjacent to the translation termination codon at the 3'-terminal and no poly-adenylation signal was found (Fig. 4A). We performed amplification from two independent reverse transcripts, but the products had identical sequences in their 3'-termini.

To obtain information about the 5'- and 3'-untranslated regions of CSBP cDNA, we determined the genomic sequences 5' upstream and 3' downstream of the coding region. In the 5'-upstream region, a stop codon was detected 13 bp upstream of the region corresponding to the 5'-terminal of the cDNA (Fig. 4C). Unless there is an intron-exon junction in this 13-bp region, this stop codon is located in-frame and the most upstream methionine residue (Fig. 4A) is the initiating methionine. Two TATAAT sequences were found at positions -271 and -179 (Fig. 4C). In the 3'-downstream region, the sequence AGAAAAAA was detected just after the 3'-terminal stop codon. Probably, reverse transcription started at this AGAAAAAAA on mRNA. Two AATAAA sequences were found at positions 747 and 842 (Fig. 4C). These may be poly-adenylation signal(s). An intron(s) seems to be located in the genome region corresponding to the open reading frame (data not shown) suggesting that the above-mentioned RT-PCR products were not derived from contaminating genomic DNA.

We attempted to extend the cDNA further upstream, but without success. Taking into account that the nucleotide length of the RT-PCR product is only about 500 bp, the 5'-terminal end of the RT-PCR product (Fig. 4A) may be the 5'-terminal end of the mRNA.

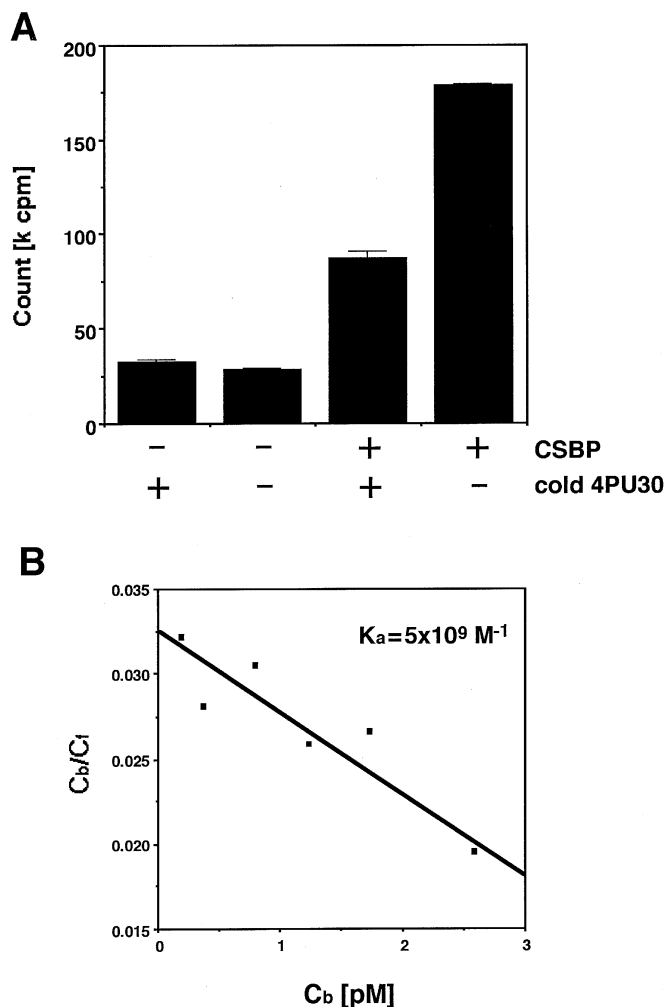


Fig. 7. Binding of recombinant CSBP and 4PU30. (A) Lysate of an *E. coli* strain transformed with GST-CSBP or the non-transformed host strain was incubated with 1 μ M [^{3} H]4PU30 in the presence or absence of 200 μ M 4PU30 (non-labeled). The amount of bacterial lysate was 40 μ g (protein amount) in each 200 μ l incubation mixture. Specific binding = (amount of [^{3} H]4PU30 bound in the absence of 200 μ M 4PU30) - (amount of [^{3} H]4PU30 bound in the presence of 200 μ M 4PU30). (B) The association constant (K_a) was determined by Scatchard analysis of [^{3} H]4PU30/GST-CSBP complex. Lysate of *E. coli* strain carrying GST-CSBP was incubated with various amounts of [^{3} H]4PU30. The amount of lysate was 24 μ g as protein in each 200 μ l incubation mixture. C_b , concentration of bound [^{3} H]4PU30; C_i , concentration of free [^{3} H]4PU30. Only high affinity binding is shown.

Cytokinin binding activity of recombinant CSBP. To confirm that the protein encoded by this cDNA indeed binds to cytokinin, we prepared the recombinant 20-kDa protein, and its binding activity was tested. The recombinant protein was designed as a fusion protein with GST so as to be expressed efficiently in soluble form. The lysate of the *E. coli* strain expressing recombinant CSBP showed specific binding (Fig. 7A), while that of the host strain showed no specific binding. The binding affinity was estimated by Scatchard analysis (Fig. 7B), giving an association constant (K_a) of 5×10^9 M^{-1} , which is consistent with the K_a value of native CSBP [15]. We also tested the binding specificity using N^6 -(Δ^2 -isopentenyl)adenine and N^6 -benzyladenine as competitors. As was the case for native CSBP, binding was subject to competition from both of these purine-type cytokinins, but not adenine or N^6 -(Δ^2 -isopentenyl)adenine riboside (Table 2).

Table 2. Inhibition of cytokinin-specific binding of recombinant CSBP by various compounds. Lysate of *E. coli* strain carrying GST-CSBP was incubated with 1 μ M [^{3} H]4PU30 in the presence of 200 μ M *N*⁶-benzyladenine [13], *N*⁶-(*A*²-isopentenyl)adenine, *N*⁶-(*A*²-isopentenyl)adenine riboside, or adenine. Total (100%) cytokinin-specific binding = (amount of [^{3} H]4PU30 bound in the absence of other added compounds) – (amount of [^{3} H]4PU30 bound in the presence of 200 μ M 4PU30).

Compound	Inhibition of specific binding
	%
Benzyladenine	28.6
Isopentenyladenine	34.2
Isopentenyladenine riboside	not observed
Adenine	not observed

DISCUSSION

We have established that a series of *N*-phenyl-*N'*-(4-pyridyl)-urea derivatives possesses potent cytokinin activity [3, 4]. Using 4PU30, a member of this series, we detected a high-affinity binding site for cytokinin in the soluble fraction of etiolated mung bean seedlings [15]. This binding site appeared to be located on a protein with an apparent molecular mass of 21 kDa [15]. We designated this putative protein as CSBP, cytokinin-specific binding protein.

In this study, we purified the protein from the soluble fraction using a 4PU-coupled affinity gel. The binding specificity of the protein to the gel precisely corresponded to the result of binding assay using [^{3} H]4PU30 as a ligand. Moreover, the molecular mass of the protein was 20 kDa, which is almost equal to the apparent molecular mass found previously, strongly suggesting that this protein is responsible for the major cytokinin binding activity initially detected. We concluded that this 20-kDa protein is CSBP.

Although urea-type and purine-type cytokinins are structurally different, their binding site is thought to be common. Based on calculations of the quantitative structure/activity relationships, Iwamura et al. [6] suggested that both types of cytokinins and cytokinin antagonists bind to a common site. They proposed a model of the cytokinin binding site [22]. CSBP probably contains the common binding site, since there was mutual competition for binding to CSBP by both types of cytokinins. Thus, taking into account that cytokinin-binding activity was originally detected as a single peak, our CSBP seems to contain the only specific cytokinin-binding site in the soluble fraction.

The affinity chromatography using 4PU-Sepharose 4B was highly effective for the purification of CSBP. In designing the affinity gel, we considered the steric hindrance to ligand-protein interaction caused by the spacer. For this, the receptor map [22] was valuable. Fitting the structures of urea-type cytokinins to the map showed that steric hindrance could be avoided if a spacer was bound at the *meta*-position of the phenyl ring in urea-type cytokinins. So, we coupled the spacer to urea-type cytokinins at the *meta*-position. First, we prepared affinity gels using DPU, 4PU and 4PU30. In affinity column chromatography, the efficiency of absorption of the cytokinin-binding activity of the gels liganded with 4PU and 4PU30 was much higher than that of the gel liganded with DPU. The adsorbed binding activity could be eluted by 4PU30, but not by 1–2 M NaCl. The elution efficiency was better in the case of the affinity gel liganded with 4PU than the affinity gel liganded with 4PU30. These data (not shown in this report) mean that binding was ligand-dependent and the affinity decreased in the order of 4PU30 $\geq$ 4PU >

DPU. This order is consistent with those in binding assay and bioassay for cytokinin activities.

We determined the amino acid sequence of the protein, which indicated that CSBP is a novel member of a major pollen allergen/pathogenesis-related protein family, though the sequence similarity is not so high. Thus, the 20-kDa protein may not simply be a homologue of these proteins, but rather a distant relative of them. Although various cytokinin-binding proteins have been reported, most of them remain to be identified [7–14, 23, 24]. As identified cytokinin-binding proteins, only ribosome and *S*-adenosyl-L-homocysteine hydrolase are known [25, 26]. Of the reported cytokinin-binding proteins, ours seems to have the highest affinity for cytokinins.

However, the function of the CSBP is still unknown. As with signal transducers, CSBP has a high affinity, and its content is low. But CSBP does not resemble known signal transducers or regulators. Recently, a membrane-spanning histidine kinase was reported as a candidate for a functional cytokinin receptor [27]. This protein was implicated in cytokinin signal transduction, because *Arabidopsis* calli overexpressing this kinase develop in the absence of cytokinin. This kinase is similar to the receptor of ethylene. Binding analysis of this protein and cytokinins, including 4PU30, is in progress. However unlike this protein, CSBP appears to be an intracellular protein. Further studies are needed to elucidate the biological function of CSBP, for example, by preparing transgenic plants.

REFERENCES

1. Skoog, F. & Armstrong, D. J. (1970) *Annu. Rev. Plant Physiol.* **21**, 359–384.
2. Hall, R. H. (1973) *Annu. Rev. Plant Physiol.* **24**, 415–444.
3. Takahashi, S., Shudo, K., Okamoto, T., Yamada, K. & Isogai, Y. (1978) Cytokinin activity of *N*-phenyl-*N'*-(4-pyridyl)urea derivatives, *Phytochem.* **17**, 1201–1207.
4. Okamoto, T., Shudo, K., Takahashi, S., Kawachi, E. & Isogai, Y. (1981) 4-pyridylureas are surprisingly potent cytokinins. The structure-activity relationship, *Chem. Pharm. Bull.* **29**, 3748–3750.
5. Kurosaki, F., Takahashi, S., Shudo, K., Okamoto, T. & Isogai, Y. (1981) Structural and biological links between urea- and purine-cytokinins, *Chem. Pharm. Bull.* **29**, 3751–3753.
6. Iwamura, H., Fujita, T., Koyama, S., Koshimizu, K. & Kumazawa, Z. (1980) Quantitative structure-activity relationship of cytokinin-active adenine and urea derivatives, *Phytochem.* **19**, 1309–1319.
7. Yoshida, K. & Takegami, T. (1977) Isolation of cytokinin binding protein from tobacco leaves by bioaffinity chromatography and its partial characterization, *J. Biochem.* **81**, 791–799.
8. Chen, C., Melitz, D. K., Petschow, B. & Eckert, R. L. (1980) Isolation of cytokinin-binding protein from plant tissues by affinity chromatography, *Eur. J. Biochem.* **108**, 379–387.
9. Hamaguchi, N., Iwamura, H. & Fujita, T. (1985) Fluorescent anticytokinin as a probe for binding. Isolation of cytokinin-binding proteins from the soluble fraction and identification of a cytokinin-binding site on ribosomes of tobacco callus cells, *Eur. J. Biochem.* **153**, 565–572.
10. Romanov, G. A., Taran, V. Y. & Venis, M. A. (1990) Cytokinin-binding protein from maize shoots, *J. Plant Physiol.* **136**, 208–212.
11. Jayabaskaran, C. (1990) Isolation and characterization of a cytokinin-binding protein from cucumber cotyledons, *Plant Growth Regulation* **9**, 9–17.
12. Momotani, E. & Tsuji, H. (1992) Isolation and characterization of a cytokinin-binding protein from the water-soluble fraction of tobacco leaves, *Plant Cell Physiol.* **33**, 407–412.
13. Sakai, S. & Kamei, N. (1992) Purification of a soluble cytokinin-binding protein from etiolated mung bean seedlings, *Biosci. Biotech. Biochem.* **56**, 504–507.

14. Mitsui, S. & Sugiura, M. (1993) Purification and properties of cytokinin-binding proteins from tobacco leaves, *Plant Cell Physiol.* **34**, 543–547.
15. Nagata, R., Kawachi, E., Hashimoto, Y. & Shudo, K. (1993) Cytokinin-specific binding protein in etiolated mung bean seedlings, *Biochem. Biophys. Res. Commun.* **191**, 543–549.
16. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature* **227**, 680–685.
17. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. (1971) Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane, *Biochemistry* **10**, 2606–2617.
18. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) DNA sequencing with chain-terminating inhibitors, *Proc. Natl Acad. Sci. USA* **74**, 5463–5467.
19. Ochman, H., Gerber, A. S. & Hartl, D. L. (1988) Genetic applications of an inverse polymerase chain reaction, *Genetics* **120**, 621–623.
20. Triglia, T., Peterson, M. G. & Kemp, D. J. (1988) A procedure for in vitro amplification of DNA segments that lie outside the boundaries of known sequences, *Nucleic Acids Res.* **16**, 8186.
21. Murray, M. G. & Thompson, W. F. (1980) Rapid isolation of high molecular mass plant DNA, *Nucleic Acids Res.* **8**, 4321–4325.
22. Iwamura, H., Murakami, S., Koshimizu, K. & Matsubara, S. (1985) Quantitative structure-activity relationships in cytokinin agonistic and antagonistic pyrido[2,3-d]pyrimidine derivatives: insights into receptor topology, *J. Med. Chem.* **28**, 577–583.
23. Kulaeva, O. N., Karavaiko, N. N., Moshkov, I. E., Selivankina, S. Yu. & Novikova, G. V. (1990) Isolation of a protein with cytokinin-receptor properties by means of anti-idiotypic antibodies, *FEBS Lett.* **261**, 410–412.
24. Kulaeva, O. N., Karavaiko, N. N., Selivankina, S. Yu., Zemlyachenko, Ya. V. & Shipilova, S. V. (1995) Receptor of trans-zeatin involved in transcription activation by cytokinin, *FEBS Lett.* **366**, 26–28.
25. Fox, J. E. & Erion, J. L. (1975) A cytokinin binding protein from higher plant ribosomes, *Biochem. Biophys. Res. Commun.* **64**, 694–700.
26. Mitsui, S., Wakasugi, T. & Sugiura, M. (1993) A cDNA encoding the 57 kDa subunit of a cytokinin-binding protein complex from tobacco: the subunit has high homology to S-adenosyl-L-homocysteine hydrolase, *Plant Cell Physiol.* **34**, 1089–1096.
27. Kakimoto, T. (1996) CKI1, a histidine kinase homolog implicated in cytokinin signal transduction, *Science* **274**, 982–985.
28. Mok, M. C., Mok, D. W. S., Armstrong, D. J., Shudo, K., Isogai, Y. & Okamoto, T. (1982) Cytokinin activity of *N*-phenyl-*N'*-1,2,3-thiadiazol-5-ylurea (Thidiazuron), *Phytochem.* **21**, 1509–1511.
29. Walter, M. H., Liu, J. W., Grand, C., Lamb, C. J. & Hess, D. (1990) Bean pathogenesis-related (PR) proteins deduced from elicitor-induced transcripts are members of a ubiquitous new class of conserved PR proteins including pollen allergens, *Mol. Gen. Genet.* **222**, 353–360.
30. Crowell, D. N., John, M. E., Russell, D. & Amasino, R. M. (1992) Characterization of a stress-induced, developmentally regulated gene family from soybean, *Plant Mol. Biol.* **18**, 459–466.
31. Iturriaga, E. A., Leech, M. J., Barrett, D. H. P. & Wang, T. L. (1994) Two ABA-responsive proteins from pea (*Pisum sativum* L.) are closely related to intracellular pathogenesis-related proteins, *Plant Mol. Biol.* **24**, 235–240.
32. Breiteneder, H., Ferreira, F., Hoffmann-Sommergruber, K., Ebner, C., Breitenbach, M., Rumpold, H., Kraft, D. & Scheiner, O. (1993) Four recombinant isoforms of *Cor a* I, the major allergen of hazel pollen, show different IgE-binding properties, *Eur. J. Biochem.* **212**, 355–362.
33. Breiteneder, H., Pettenburger, K., Bito, A., Valenta, R., Kraft, D., Rumpold, H., Scheiner, O. & Breitenbach, M. (1989) The gene coding for the major birch pollen allergen *Betv1*, is highly homologous to a pea disease resistance response gene, *EMBO J.* **8**, 1935–1938.
34. Vanek-Krebitz, M., Hoffmann-Sommergruber, K., Laimer da Camara Machado, M., Susani, M., Ebner, C., Kraft, D., Scheiner, O. & Breiteneder, H. (1995) Cloning and sequencing of Mal d 1, the major allergen from apple (*Malus domestica*), and its immunological relationship to Bet v 1, the major birch pollen allergen, *Biochem. Biophys. Res. Commun.* **214**, 538–551.
35. Matton, D. P., Prescott, G., Bertrand, C., Camirand, A. & Brisson, N. (1993) Identification of cis-acting elements involved in the regulation of the pathogenesis-related gene *STH-2* in potato, *Plant Mol. Biol.* **22**, 279–291.
36. Somssich, I. E., Schmelzer, E., Kawalleck, P. & Hahlbrock, K. (1988) Gene structure and in situ transcript localization of pathogenesis-related protein 1 in parsley, *Mol. Gen. Genet.* **213**, 93–98.
37. Breiteneder, H., Hoffmann-Sommergruber, K., O'Riordain, G., Susani, M., Ahorn, H., Ebner, C., Kraft, D. & Scheiner, O. (1995) Molecular characterization of Api g 1, the major allergen of celery (*Apium graveolens*), and its immunological and structural relationships to a group of 17-kDa tree pollen allergens, *Eur. J. Biochem.* **233**, 484–489.
38. Warner, S. A. J., Scott, R. & Draper, J. (1992) Characterisation of a wound-induced transcript from the monocot asparagus that shares similarity with a class of intracellular pathogenesis-related (PR) proteins, *Plant Mol. Biol.* **19**, 555–561.
39. Sokal, R. R. & Michener, C. D. (1958) *University of Kansas Sci. Bull.* **28**, 1409–1438.
40. Kyte, J. & Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* **157**, 105–132.