Phytochemistry, Vol. 26, No. 1, pp. 41–45, 1987.
Printed in Great Britain.

0031 9422/87 $3.00 + 0.00

STRUCTURAL HOMOLOGY AMONG THE MAJOR 7S GLOBULIN SUBUNITS OF SOYBEAN SEED STORAGE PROTEINS

H. HIRANO, H. KAGAWA, Y. KAMATA* and F. YAMAUCHI*

Department of Molecular Biology, National Institute of Agrobiological Resources, Yatabe, Tsukuba, Ibaraki 305, Japan; *Faculty of Agriculture, Tohoku University, Sendai, Miyagi 980, Japan

(*Revised received* 9 *May* 1986)

Key Word Index –*Glycine max*; Leguminosae; soybean; seed storage protein; β-conglycinin; γ-conglycinin; basic 7S globulin; amino acid sequence.

Abstract—The soybean seed 7S globulin subunits, i.e. α, α′, β and γ-subunits of β-conglycinin, the γ-conglycinin subunit and the HI/HII and LII subunits of basic 7S globulin were purified and the NH_2-terminal amino acid sequences of all these subunits except the γ-subunit of β-conglycinin were determined. Only the NH_2-terminal regions of the α and α′-subunits showed high sequence homology. However, sequencing of tryptic peptides from the seven subunits revealed that internal region sequences were highly homologous among the four subunits of β-conglycinin. In contrast to the β-conglycinin subunits, no sequence homology was found among the other subunits. On the basis of these results, the major 7S globulin fraction is considered more heterogeneous in primary structure than another major globulin fraction, 11S globulin (glycinin), in soybean seeds.

INTRODUCTION

The 7S globulin, one of the major seed storage proteins in soybean [*Glycine max* (L.) Merrill], is separated into three major fractions with different physicochemical properties, designated β-conglycinin, γ-conglycinin and basic 7S globulin. β-Conglycinin has a trimeric structure with a M_r of 140–170 k and consists of three subunits: α (57 k), α′ (58 k) and β (42 k); a relatively minor 42 k protein designated the γ-subunit copurifies with the β-conglycinin subunits [1]. γ-Conglycinin is a trimer with a M_r of 170 k [2] consisting of three identical subunits [3]. The basic 7S globulin, which has a M_r of 168 k [4], is composed of four pairs of high M_r subunits (HI and HII) (26 k) and low M_r subunits (LI and LII) (16 k) linked together via disulphide bond(s) [5].

The NH_2-terminal amino acids of the α (Val), α′ (Val) and β (Leu) subunits of β-conglycinin [1] and the γ-conglycinin subunit (Ile) [3] have been identified. The NH_2-terminal amino acid sequence of the β-subunit of β-conglycinin has also been determined [6]. Recently, partial amino acid sequences for the α and α′-subunits of β-conglycinin were derived from partial nucleotide sequences of the DNAs encoding these subunits [7, 8]. However, so far little is known on the amino acid sequences of most of the major 7S globulin subunits in soybean seeds.

We have now purified the major 7S globulin subunits, i.e. the α, α′, β and γ-subunits of β-conglycinin, the γ-conglycinin subunit and the HI/HII and LII subunits of the basic 7S globulin, and determined the NH_2-terminal amino acid sequences of all these subunits except the γ-subunit of β-conglycinin. The NH_2-terminal sequences were found to be homologous only between the α and α′-subunits of β-conglycinin in the 7S globulin. The seven subunits were cleaved enzymatically and the resultant peptides sequenced. This revealed that the sequences in the internal region are highly homologous among the four subunits of β-conglycinin. However, no sequence homology was found among the other subunits of the 7S globulin.

RESULTS AND DISCUSSION

The seven subunits of 7S globulin were partially purified from soybean (cv. Raiden) seeds as described in refs. [1, 3, 5] and lyophilized. The partially purified subunits were suspended in 0.0625 M Tris·HCl buffer (pH 6.8) containing 5% 2-mercaptoethanol and 2.3% sodium dodecyl sulphate (SDS) and heated at 95° for 5–10 min. The solution was subjected to reverse-phase high performance liquid chromatography (HPLC) for further purification and desalting. The α, α′, β and γ-subunits of β-conglycinin, the γ-conglycinin subunit and the LII subunit of basic 7S globulin were purified to homogeneity by this method. The 7S globulin subunits were electrophoretically at least 95% pure (Fig. 1). These purified subunits were used for peptide mapping and protein sequencing. The HI and HII subunits of the basic 7S globulin were not separated completely by HPLC (Fig. 1). Therefore, the HI and HII subunits were used for peptide mapping and sequencing as a mixture (HI/HII subunit).

The subunits eluted with acetonitrile gradients in trifluoroacetic acid (TFA) from the HPLC column were applied directly to a gas-phase protein sequencer [9] for determination of the NH_2-terminal amino acid sequences without prior lyophilization. This was necessary since lyophilized polypeptides of 7S globulin are insoluble in the solvents commonly used in protein sequencing.

Figure 2 (a) shows the NH_2-terminal amino acid sequences of the 7S globulin subunits. The NH_2-terminal residues of the α, α′ and β-subunits of β-conglycinin agree

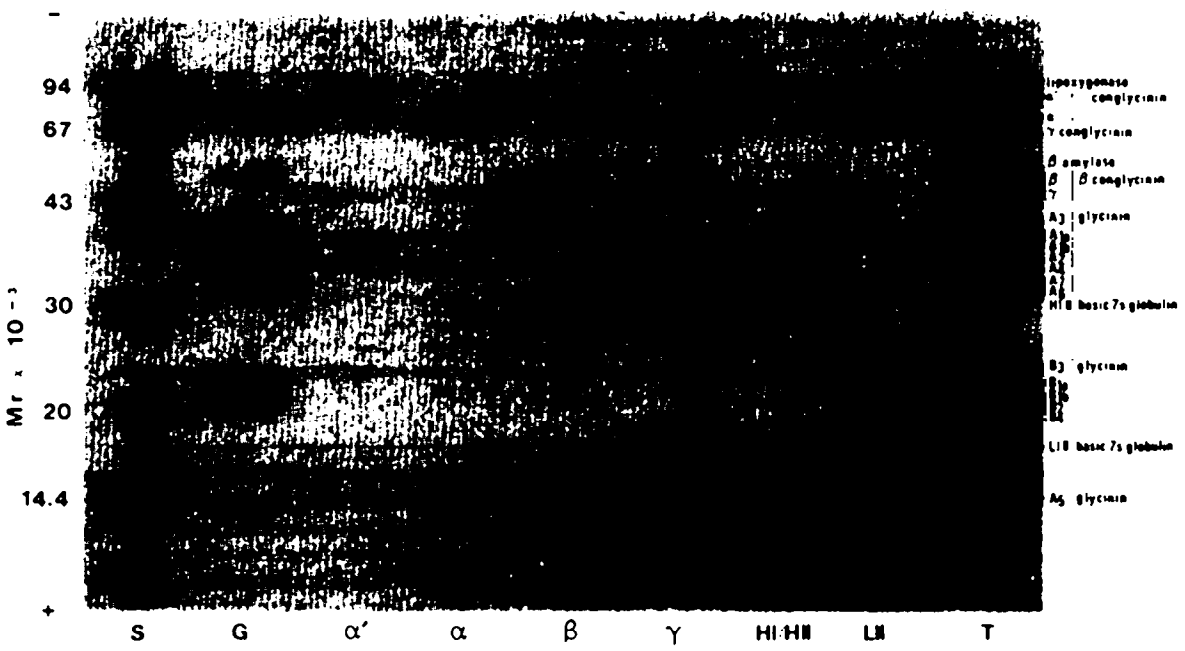

Fig. 1. SDS gel electrophoresis (17% gel) of the purified 7S globulin subunits. α′, α′-Subunit of β-conglycinin; α, α-subunit of β-conglycinin; β, β-subunit of β-conglycinin; γ, γ-conglycinin subunit; HI/HII, HI/HIII subunit of basic 7S globulin; LII, LII subunit of basic 7S globulin; G, crude 11S globulin (glycinin) (cv. Bonminori); S, MW calibration proteins (Pharmacia).

with those published previously [1]. However, the NH_2-terminal residue (Lys) of γ-conglycinin was different from that (Ile) published before [3]. This may be due to microheterogeneity of the γ-conglycinin subunit. No released PTH-amino acids could be identified after ten cycles of Edman degradation of the γ-subunit of β-conglycinin. This indicates a blocked NH_2-terminus.

As shown in Fig. 2 (a), the NH_2-terminal sequences are homologous only between the α and α′-subunits of β-conglycinin in the 7S globulin.

Staswick *et al.* [6] have determined the NH_2-terminal amino acid sequence (1st to 15th residues) of the β-subunit of β-conglycinin. The sequence is identical to that determined here. We confirmed that the amino acid sequence of the β-subunit is highly homologous to that of the 47 k subunit of the 7S globulin (vicilin) in pea seeds [Fig. 2 (b)] [10]. When aligned for maximal homology, 12 residues of 29 of the NH_2-terminal sequences of the β-subunit are identical to those of the vicilin subunit. However, the NH_2-terminal sequences of the other subunits of the soybean 7S globulin are different from that of the vicilin 47 k subunit.

The basic 7S globulin is composed of four subunit pairs, each pair consisting of high M_r and low M_r subunits which are linked by disulphide bond(s). Two high M_r subunits, HI and HII, and two low M_r subunits, LI and LII, with slightly different M_rs, respectively, have been identified [5]. Since we could not separate the HI and HII subunits in the present study, we analysed the NH_2-terminal sequence of a mixture of the HI and HII subunits, designated the HI/HII subunit. The HI/HII subunit revealed a single NH_2-terminal sequence [Fig. 2 (a)]. Apparently, the HI and HII subunits have an identical NH_2-terminal sequence. The LI subunit was not available for NH_2-terminal sequence analysis.

We examined the homology of the internal sequences of the 7S globulin subunits by tryptic peptide mapping using HPLC. The 7S globulin subunit fractions eluted by HPLC were lyophilized and digested with trypsin. The digests were separated on a reverse-phase HPLC column.

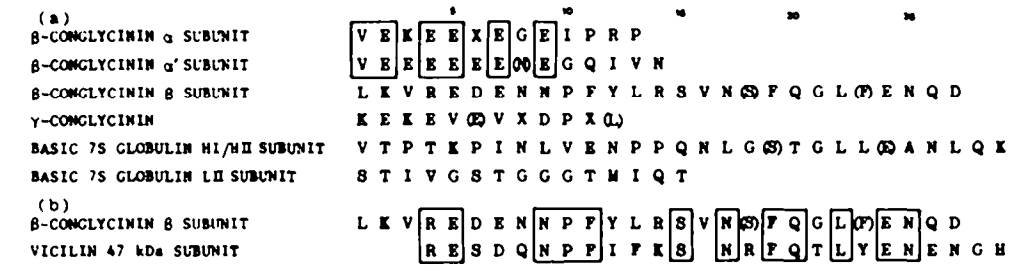

Fig. 2. (a) NH_2-terminal amino acid sequences of the 7S globulin subunits. Boxed residues are identical in the α and α′-subunits of β-conglycinin. (B) Structural homology between the β-subunit of β-conglycinin and the pea vicilin 47 k subunit [10]. Boxed residues are identical in the subunits.

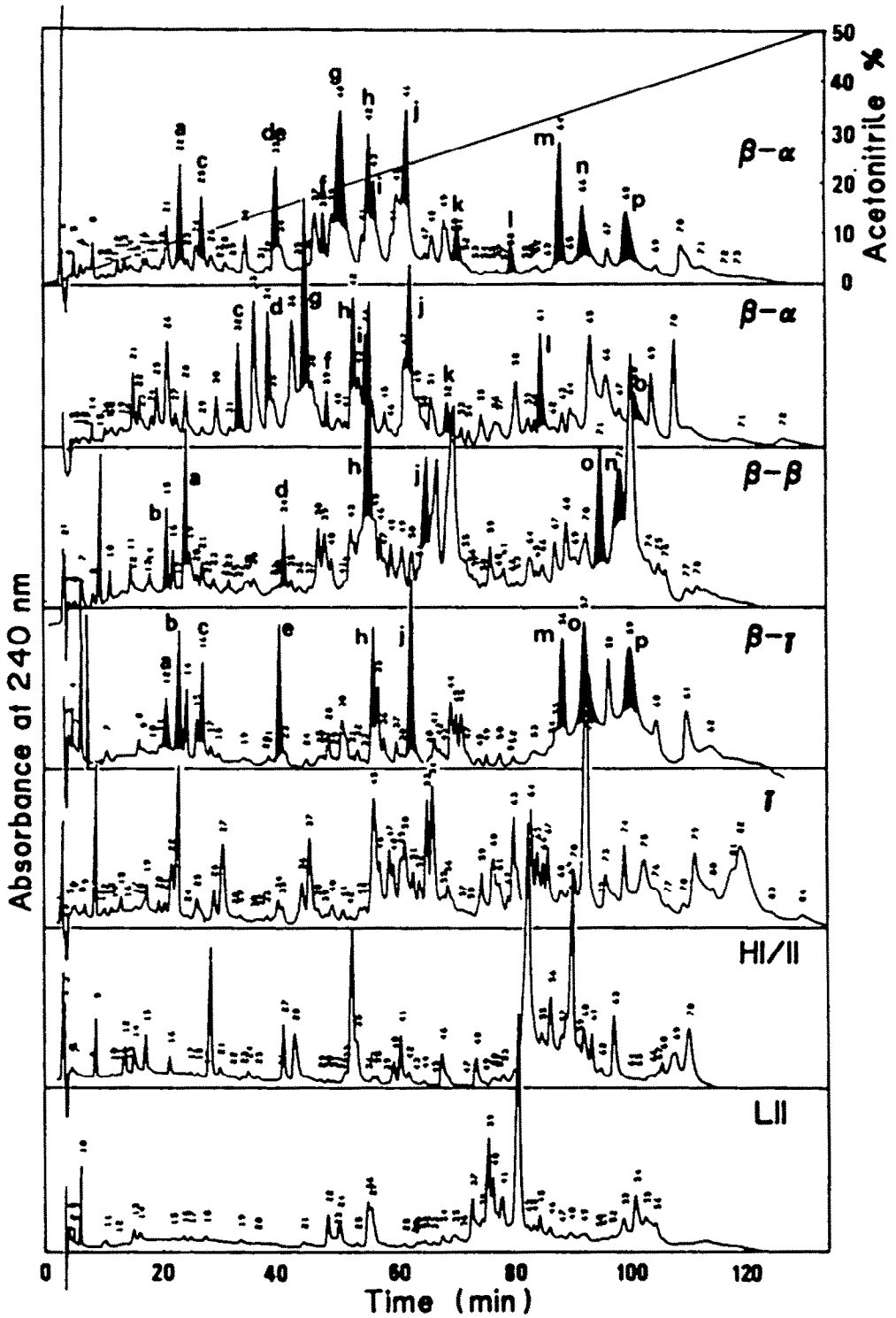

Fig. 3. HPLC elution profiles of the tryptic peptides of the seven subunits of 7S globulin. The peptides were chromatographed on a reverse-phase column with an acetonitrile gradient in 0.1% aqueous TFA. The flow rate was 1.0 ml/min and the column temperature 25°. Solid peaks (a–p) represent peptides which have an identical or similar sequence between two or more subunits. Their sequences are given in Table 1. Numerals above the peaks represent the order of elution in each chromatographic experiment.

Figure 3 shows the HPLC elution profiles of the digests of the 7S globulin subunits. The HPLC elution profiles were found to be similar among the four subunits of β-conglycinin. However, there was no similarity in the profiles between the β-conglycinin and γ-conglycinin subunits, between the γ-conglycinin and the 7S globulin subunits, and between the basic 7S globulin and β-conglycinin subunits. Also, the HPLC elution profile of the HI/HII subunit was different from that of the LII subunit in the basic 7S globulin. The HPLC elution profiles of the major 7S globulin subunits suggest that only the subunits of β-conglycinin are homologous.

To confirm this conclusion, we analysed the amino acid sequences of the tryptic peptides from all of the subunits by manual microsequence analysis using 4-*N*,*N*-dimethylaminoazobenzene 4'-isothiocyanate (DABITC) [11] or by the gas-phase protein sequencer. As suggested by the HPLC elution profiles, many, but not all, identical or similar sequences of peptides were found among the subunits of β-conglycinin. The homologous sequences of peptides are shown in Fig. 3 and Table 1. However, the amino acid sequences of the tryptic peptides from the subunit of γ-conglycinin and the HI/HII and LII subunits of the basic 7S globulin examined were unique to the respective subunits (H. Kagawa, F. Yamauchi and H. Hiraino, unpublished results).

Thanh and Shibasaki [1] have suggested that the α and α'-subunits of β-conglycinin should be homologous in their primary structures since they are immunologically related and have similar amino acid compositions. It was found in the present study that these subunits have highly homologous sequences not only in the NH_2-terminal regions but also in the internal regions. They [1] have also indicated that the α and α'-subunits are immunologically different from the β-subunit. However, the results of our peptide sequence determination show a high sequence homology among these subunits.

Schuler *et al.* [7, 8] have deduced the amino acid sequences in the COOH-terminal regions of the α/α'-subunit from the nucleotide sequences. The actual sequences of some of the peptides determined here agree with the sequences predicted by the nucleotide sequences. The nucleotide sequence of the α/α'-subunit has been shown to be highly homologous to those of French bean phaseolin and pea vicilin [12, 13]. Since the α and α'-subunits are homologous to the β and γ-subunits in β-conglycinin as shown here, the genes encoding all these subunits are suggested to be closely related evolutionarilly with the phaseolin and vicilin genes.

We have constructed synthetic oligonucleotides corresponding to the amino acid sequences unique to each subunit. Using these oligonucleotides as probes, we are cloning the DNA encoding the 7S globulin subunits. The complete amino acid sequences of these subunits will be deduced from the nucleotide sequences.

Recently, we have determined the complete amino acid sequences of the subunits of another major seed storage protein, 11S gobulin (glycinin) in soybean, by protein sequencing [14, 15] or have predicted them by sequencing the cDNA for specific proteins [16, 17]. We have found that there is a great similarity in their amino acid sequences among the glycinin subunits. In the present study, we confirmed that in the 7S globulin the sequence homology is high only among the β-conglycinin subunits, but the other subunits have unique sequences. Based on these results, the 7S globulin fraction seems to be more heterogeneous than the 11S globulin fraction.

EXPERIMENTAL

Purification of the 7S globulin subunits. The β-conglycinin, γ-conglycinin and basic 7S globulin subunits were partially purified from soybean (cv. Raiden) seeds as described in refs. [1], [3] and [5], respectively and lyophilized. These partially purified subunits (0.5 mg) were suspended in 100 μl 0.0625 M Tris HCl buffer (pH 6.8) containing 10% glycerol, 5% 2-mercaptoethanol and 2.3% SDS, heated at 95° for 5–10 min and applied to a reverse-phase HPLC column (Nucleosil, C18, 5 μm).

Table 1. Homology of the amino acid sequences of the tryptic peptides among the four subunits of β-conglycinin

	Subunit			
Peak	α	α'	β	γ
a	DYGQQQGEQR		DYGQQQGEQR	KYGQQQGEQR
b			YLSK	YSK
c	NQYGR	NQYGHVR		NQYGR
d	SRDPIYSNK	SRDPIYSNK	SRDPIYSNK	
e	LQSGVALNEISK			LQSQVALNK
f	EQQQEQQQEEQPLEVR	QQQQEQQQEEQPLEVR		
g	FPQLQNLK	SPQLQNLK		
h	ESYFVDAQPK	SQYESYFVDAQPK	ESYFVDAQPK	ESYFVDAQPK
i	NILEASYDTK	NILEASYDTK		
j	VPAGTTYX	VPAGTTYX	VPAGTTYX	VPAGTTYYEX
k	LITLAIPVNK	MITLAIPVNK		
l	FESFFLSX	FESFFLSX		
m	NFLAGSQDNVX			NFLAG(L)QDNX
n	KELFFLDIFVX		AELFLX	
o		ALLFEQX	AVLFEQK	AVLFEQK
p	ILEFNXK			PLEFNK

Numbers of peptides determined unambiguously their amino acid sequences which were 25, 24, 21, 19, 25, 15 and 8 for the α, α', β and γ-subunits of the β-conglycinin, the γ-conglycinin and the HI/HII and LII subunits of basic 7S globulin, respectively.

The column was run at a flow rate of 1.5 ml/min and the subunits were eluted with a 20–100% acetonitrile gradient in 0.1% aq. TFA over 50 min at a column temp. of 45°. The effluent was monitored by *A* at 280 nm.

Gel electrophoresis. SDS gel electrophoresis (17% gel) was carried out as described in ref. [18].

Amino acid sequence analysis. The NH_2-terminal amino acid sequences of the subunits (0.5 nmol) were determined by a gas-phase protein sequencer (Applied Biosystems 470A) [9]. The purified subunits (10 nmol) were digested with trypsin (Sigma) (1 : 50 enzyme/protein in 0.1 M ammonium bicarbonate for 18 hr at 15°) [19]. The digest was dissolved in 100 μl 0.1% aq. TFA, applied to a reverse-phase HPLC column (Varian MCH-5, C18, 5 μm), and eluted at a flow rate of 1 ml/min with a 0–50% acetonitrile gradient in 0.1% aq. TFA over 130 min at a column temp. of 25°. The effluent was monitored by *A* at 240 nm. The eluted individual peptides were collected manually, lyophilized and used for protein sequencing. This reverse-phase HPLC did not resolve all the components in the initial separation and some peptides were rechromatographed on the same system using shallower gradients. The purified peptides were sequenced manually using DABITC (Dojin) [11] or by the gas-phase protein sequencer.

Acknowledgement—We thank Dr. M. Walsh for his critical reading of the manuscript.

REFERENCES

1. Thanh, V. H. and Shibasaki, K. (1977) *Biochim. Biophys. Acta* **490**, 370.
2. Sato, W., Kamata, Y., Fukuda, M. and Yamauchi, F. (1984) *Phytochemistry* **23**, 1523.
3. Yamauchi, F., Sato, W. and Kamata, Y. (1985) *Phytochemistry* **24**, 1503.
4. Yamauchi, F., Sato, K. and Yamagishi, T. (1984) *Agric. Biol. Chem.* **48**, 645.
5. Sato, W., Yamagishi, T. and Yamauchi, F., *Phytochemistry*, submitted.
6. Staswick, P. E., Broué, P. and Nielsen, N. C. (1983) *Plant Physiol.* **72**, 1114.
7. Schuler, M. A., Schmitt, E. S. and Beachy, R. N. (1982) *Nucleic Acids Res.* **10**, 8225.
8. Schuler, M. A., Ladin, B. F., Pollaco, J. C., Freyer, G. and Beachy, R. N. (1982) *Nucleic Acids Res.* **10**, 8245.
9. Hewick, R. W., Hunkapiller, M. W., Hood, L. E. and Dreyer, W. J. (1981) *J. Biol. Chem.* **256**, 7990.
10. Lycett, G. W., Delauney, A. J., Gatehouse, J. A., Gilroy, J., Croy, R. R. D. and Boulter, D. (1983) *Nucleic Acids Res.* **11**, 2367.
11. Chang, J. Y., Brauer, D. and Wittmann-Liebold, B. (1978) *FEBS Letters* **93**, 205.
12. Schuler, M. A., Doyle, J. J. and Beachy, R. N. (1983) *Plant Mol. Biol.* **2**, 119.
13. Gatehouse, J. A., Croy, R. R. D. and Boulter, D. (1984) *CRC Crit. Rev. Plant Sci.* **1**, 287.
14. Hirano, H., Fukazawa, C. and Harada, K. (1984) *J. Biol. Chem.* **259**, 14371.
15. Hirano, H., Fukazawa, C. and Harada, K. (1985) *FEBS Letters* **181**, 124.
16. Fukazawa, C., Momma, T., Hirano, H., Harada, K. and Udaka, K. (1985) *J. Biol. Chem.* **260**, 6234.
17. Momma, T., Negoro, T., Hirano, H., Matsumoto, A., Udaka, K. and Fukazawa, C. (1985) *Eur. J. Biochem.* **149**, 491.
18. Laemmli, U. K. (1970) *Nature (London)* **227**, 680.
19. Hirano, H., Gatehouse, J. A. and Boulter, D. (1982) *FEBS Letters* **145**, 99.