Glycoconjugate Journal (1994) **11**:299–303

Characterization of isolectins in *Tetracarpidium conophorum* seeds (Nigerian walnut)

THERESA ANIMASHAUN[1], R. ADETORO TOGUN[2] and R. COLIN HUGHES[1]*

[1] *National Institute for Medical Research, The Ridgeway, Mill Hill, London, NW7 1AA, UK*
[2] *Department of Biochemistry, University of Obafemi-Awolowo, Ife-Ife, Nigeria*

Received 2 May 1994, revised 10 June 1994

A lectin preparation obtained from *Tetracarpidium conophorum* (Nigerian walnut) by affinity chromatography of seed extracts on lactose-agarose has been shown to contain two components by gel filtration on Sephadex G150. The larger component *Tetracarpidium conophorum* agglutinin I (TCAI) is a disulphide-bonded 70 kDa homodimer whereas the second component TCAII is a 34 kDa monomeric protein. Amino terminal aminoacid sequencing shows identity in TCAI and TCAII for the first fifteen residues after which the sequences diverge. The *N*-terminal sequences of TCAI and TCAII show identity with sequences in the B-chains of ricin and *Ricinus communis* agglutinin I (RCAI) in eleven of the initial fifteen residues. Thereafter TCAI appears to be homologous to the ricin B chain whereas TCAII is more homologous with the B chain of RCAI. A limited screening of the carbohydrate-binding specificity of TCAII by affinity chromatography of defined oligosaccharides on TCAII Sepharose columns shows that the binding specificity reported earlier for affinity purified *Tetracarpidium conophorum* isolectins (Sato S, Animashaun T, Hughes RC (1991) *J Biol Chem* **266**:11485–94) reflects the binding properties of TCAII which is the major isolectin in unfractionated lectin preparations.

Keywords: Tetracarpidium conophorum; isolectins

Introduction

In previous reports [1, 2] we described a galactoside-binding activity in extracts of seeds of the Nigerian walnut *Tetracarpidium conophorum*, a member of the *Euphorbiaceae* which includes other plants such as *Ricinus communis* containing well-characterized galactose-specific isolectins [3–7]. Affinity chromatography of *Tetracarpidium conophorum* seed extracts on lactose Sepharose gave an active fraction by elution with lactose and SDS-polyacrylamide gel electrophoresis under reducing conditions gave a single polypeptide band of about 34 000 molecular mass which yielded a single *N*-terminal sequence in the first five residues indicating a unique component [1]. Detailed studies of the carbohydrate-binding specificity of this lectin fraction showed several useful properties [2]. In particular the lectin activity showed preference for simple, unbranched oligosaccharides containing a terminal Galβ1 → 4GlcNAc sequence over a Galβ1 → 3GlcNAc sequence and terminal substitution by sialic acid or fucose significantly reduced binding activity. The highest preference in branched oligosaccharides was found for tri-antennary glycans carrying Galβ1 → 4GlcNAc substituents on C-2 and C-4 of Manα1 → 3 and C-2 of Manα1 → 6 core residues. More recently we have detected some polypeptide heterogeneity in *T. conophorum* agglutinin (TCA) preparations. SDS-polyacrylamide gel electrophoresis under non-reducing conditions showed variable but significant amounts of a polypeptide migrating with molecular mass of ~70 kDa in addition to the major ~34 kDa component. In this paper we show that affinity purified lectin fractions of *T. conophorum* seeds contains two galactoside-binding isolectins TCAI and TCAII. The carbohydrate-binding specificity of TCAII is reported and compared with our earlier conclusions [2] obtained using unfractionated isolectin preparations.

Materials and methods

Isolation of TCA isolectins

Saline extracts of *T. conophorum* seeds were applied to a column (3 ml) of lactose Agarose (Sigma Chemical Corp.) and after extensive washing with phosphate-buffered saline (PBS) to remove unbound proteins, a lectin fraction was eluted with 0.2 M lactose in PBS. The lectin fraction was dialysed against water, freeze-dried and reconstituted in 50 mM Tris HCl pH 7.2, 150 mM NaCl (TBS). Aliquots

* To whom correspondence should be addressed.

(5 mg protein, 0.25 ml) were adjusted to 200 mM lactose and applied to a column (1 cm × 86 cm) of Sephadex G150 previously equilibrated with TBS, 200 mM lactose. Fractions (1 ml) were eluted with TBS, 200 mM lactose and analysed for absorbance at 280 nm. Peak fractions were pooled, dialysed against water, freeze-dried and reconstituted in H_2O or TBS at approximately 1 mg ml^{-1}. The Sephadex G150 column was calibrated with Blue dextran to determine the void volume and with protein standards of known molecular mass, alcohol dehydrogenase (150 kDa), bovine serum albumin (66 kDa) and carbonic anhydrase (29 kDa). Gel filtration on columns (1 cm × 80 cm) of Biogel P150 and ACA54 was also carried out under the same conditions.

Oligosaccharides

The following oligosaccharides were gifts from the late Dr A. S. R. Donald, Research Sugars, Long Crendon, Aylesbury, UK: lacto-*N*-tetraose, LNT, Galβ1 → 3GlcNAcβ1 → 3Galβ1 → 4Glc; lacto-*N*-neotetraose LNNT, Galβ1 → 4GlcNAc β1 → 3Galβ1 → 4Glc; lacto-*N*-fucopentaose III, LNFPIII, Galβ1 → 4[Fucα1 → 3]GlcNAc β1 → 3Galβ1 → 4Glc. The structure of other oligosaccharides are shown in Fig. 4 and were obtained as described previously [2]. Oligosaccharides were labelled by reduction with NaB^3H_4 (600 mCi $mmol^{-1}$, Du Pont New England Nuclear).

Gel electrophoresis

SDS-polyacrylamide gel electrophoresis was carried out under reducing or non-reducing conditions as described [8]. Lectin samples containing up to 20 µg of protein were loaded on to minigels of 10% polyacrylamide with protein standards. After electrophoresis, proteins were located by staining with Coomassie blue.

Automated peptide sequencing

For sequencing, aliquots of the lectins (10–20 nmol) were analysed with Applied Biosystems 470A gas-phase and 477A pulsed liquid-phase sequencers with Applied Biosystems 120A on-line phenylthiohydantoin analysers [9]. Data collection and analysis were performed with an Applied Biosystems 900A module calibrated with 25 pmol phenyl thiohydantoin standards.

Affinity chromatography

Purified lectin was coupled to CNBr-activated Sepharose 4B using conditions recommended by the manufacturers (Pharmacia LKB). Columns of approximately 1 ml bed volume were packed with lectin Sepharose containing approximately 5 mg of lectin per ml. The columns were washed with 10 mM Tris HCl pH 7.5, 0.1 M NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 1 mM $MnCl_2$. Radioactive oligosaccharide standards were applied to the columns in elution buffer (0.1–0.2 ml) and incubated for 1 h at room temperature to ensure maximal adsorption. The columns were then washed sequentially with elution buffer, 10 mM lactose, 100 mM lactose and 250 mM lactose all in elution buffer. Fractions (1 ml) were collected and assayed for radioactivity. Other details of affinity chromatography are discussed later.

Haemagglutination assay

Haemagglutination and hapten inhibition was measured as described [10].

Results and discussion

Purification of isolectins

The main point of the present study is to show that *Tetracarpidium conophorum* seeds contain two distinct sugar binding proteins as shown by gel filtration of the lectin fraction obtained after affinity chromatography of seed extracts on lactose agarose and specific elution with lactose. The two protein peaks obtained after gel filtration on Sephadex G150 were eluted with apparent molecular masses of about 70 kDa for TCAI and 34 kDa for TCAII (Fig. 1). The aminoacid compositions of TCAI and TCAII were similar to each other and to the unfractionated TCA lectin mixture (results not shown). The relative yields of TCAI and TCAII appeared to be variable between different preparations for unknown reasons and in some cases TCAI accounted for as much as 25% of total lectins (see Fig. 1). Both lectin fractions were found to have similar levels of haemagglutinating activity. The minimum concentration of TCAI and TCAII required for agglutination of A erythrocytes was 16 µg ml^{-1}. Haemagglutination of TCAI was inhibited by simple haptens in the order:

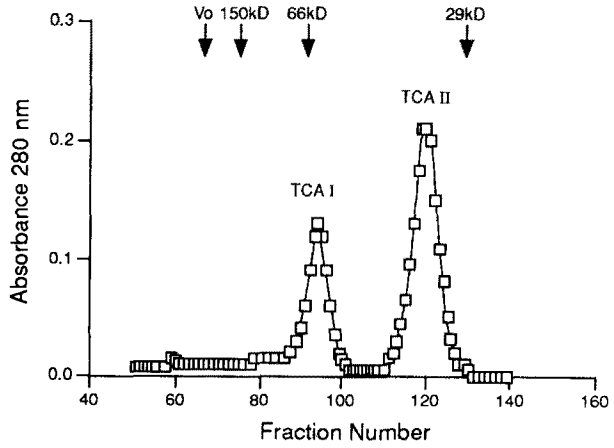

Figure 1. Gel filtration of isolectins purified from extracts of *Tetracarpidium conophorum* seeds. The lectin fraction isolated from seed extracts by affinity chromatography on lactose Agarose was applied to Sephadex G150 and peak fractions containing isolectins TCAI and TCAII were pooled for analysis. The void volume (V_0) and elution of protein standards of known size ($M_r \times 10^{-3}$) are indicated.

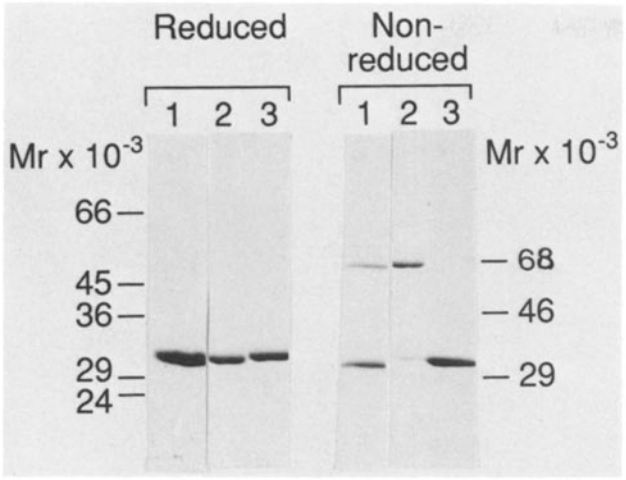

Figure 2. SDS-polyacrylamide gel electrophoresis of *Tetracarpidium conophorum* isolectins. Track 1, mixture of lectins from lactose Agarose affinity chromatography of seed extracts. Track 2, TCAI; Track 3, TCAII. Electrophoresis was carried out under non-reducing or reducing conditions and proteins were detected by Coomassie Blue staining. Migrations of protein standards ($M_r \times 10^{-3}$) are shown.

N-acetylgalactosamine > methyl β galactoside > lactose > galactose = methyl α galactoside. Haemagglutination by TCAII was inhibited by the same haptens in the order: lactose > methyl β galactoside > methyl α galactoside = galactose = *N*-acetyl-galactosamine (results not shown). The results obtained for TCAII are similar to those reported earlier for the unfractionated *T. conophorum* isolectins [1]. We conclude therefore that isolectins TCAI and TCAII have rather distinct carbohydrate-binding properties at this level of analysis. Gel filtration was also carried out on agarose Biogel P150 or agarose-acrylamide ACA54 columns. In both cases TCAI and TCAII were eluted as proteins of molecular masses 36–42 kDa and 18 kDa respectively, indicating interactions of both lectins with the chromatography matrix although gel filtration on the agarose-based column matrices was carried out in buffers containing 200 mM lactose (results not shown).

Subunits of TCAI and TCAII

During SDS-polyacrylamide gel electrophoresis TCAI migrated as a single band of apparent molecular size 65–70 kDa under non-reducing conditions and as a single band of approximately 34 kDa under reducing conditions (Fig. 2). By contrast, TCAII behaved during SDS-polyacrylamide gel electrophoresis as a ~34 kDa component under reducing or non-reducing conditions (Fig. 2). Electrophoresis of the crude lectin fraction showed under non-reducing conditions both the 65–70 kDa and 34 kDa bands in ratio 1:4 in the particular sample shown in Fig. 2. Upon reduction only a heavy 34 kDa band was detected (Fig. 2) as previously described [1]. These data suggest that TCAI contains two disulphide-bonded subunits each of 34 kDa

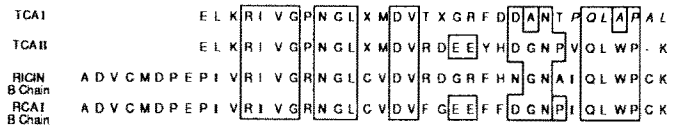

Figure 3. *N*-terminal aminoacid sequencers of *Tetracarpidium conophorum* isolectins and comparison with other galactose-specific lectins. Residues 12 and 17 in TCAI and TCAII were not positively identified but are assumed to be cysteine. Tentative residue assignments in TCAI are indicated by *italicized* letters. Data for ricin and RCAI are from Roberts *et al.* [6] and Araki and Funatsu [7].

molecular mass and similar in size to the polypeptide of the monomeric TCAII lectin.

In order to determine the relationship between the ~34 kDa polypeptide subunits of TCAI and TCAII we performed amino-terminal protein sequencing. These analyses gave essentially unambiguous sequences for the first 32 aminoacid residues except for residues 12 in TCAI and TCAII and residue 17 in TCAI (Fig. 3). We believe these residues to be cysteine. Interestingly TCAI and TCAII contain the same unique sequence for the first fifteen aminoacid residues, in agreement with a previous preliminary finding that the crude lectin preparation obtained after affinity chromatography on lactose agarose but without the subsequent gel filtration step gave a single, unique *N*-terminal sequence spanning the first five amino-acid residues (T. Animashaun, unpublished result). In the present analysis, we find that the sequences of TCAI and TCAII diverge after residue Val_{15} although there persists a significant homology in that five out of seventeen residues after Val_{15} are identical. The fact also that we can detect only one sequence in TCAI, at least over the first 32 residues from the amino-terminus, suggests strongly that TCAI consists of two identical ~34 kDa subunits that are disulphide-bonded in the ~65–70 kDa dimer rather than this lectin being made up of two dissimilar disulphide-bonded polypeptides each of 34 kDa molecular mass. However, more extensive sequence data will be required to substantiate this conclusion.

Homologies with ricin and RCAI

As Fig. 3 shows the *N*-terminal aminoacid sequences of TCAI and TCAII are very homologous to the galactose-binding subunits of ricin [3–5, 7, 11] and RCAI [6, 12]. Alignment of the *N*-terminus sequences of TCAI and TCAII with the ricin and RCAI sequences starting at Pro_9 shows thirteen identities between all chains within the next thirty-two residues and a similar number of conservative changes. It is apparent that similarly to the *N*-terminal sequences of TCAI and TCAII, the first 23 B chain *N*-terminal residues of ricin and RCAI as deduced from nucleotide or protein sequencing are identical but the following residues differ drastically in sequence between

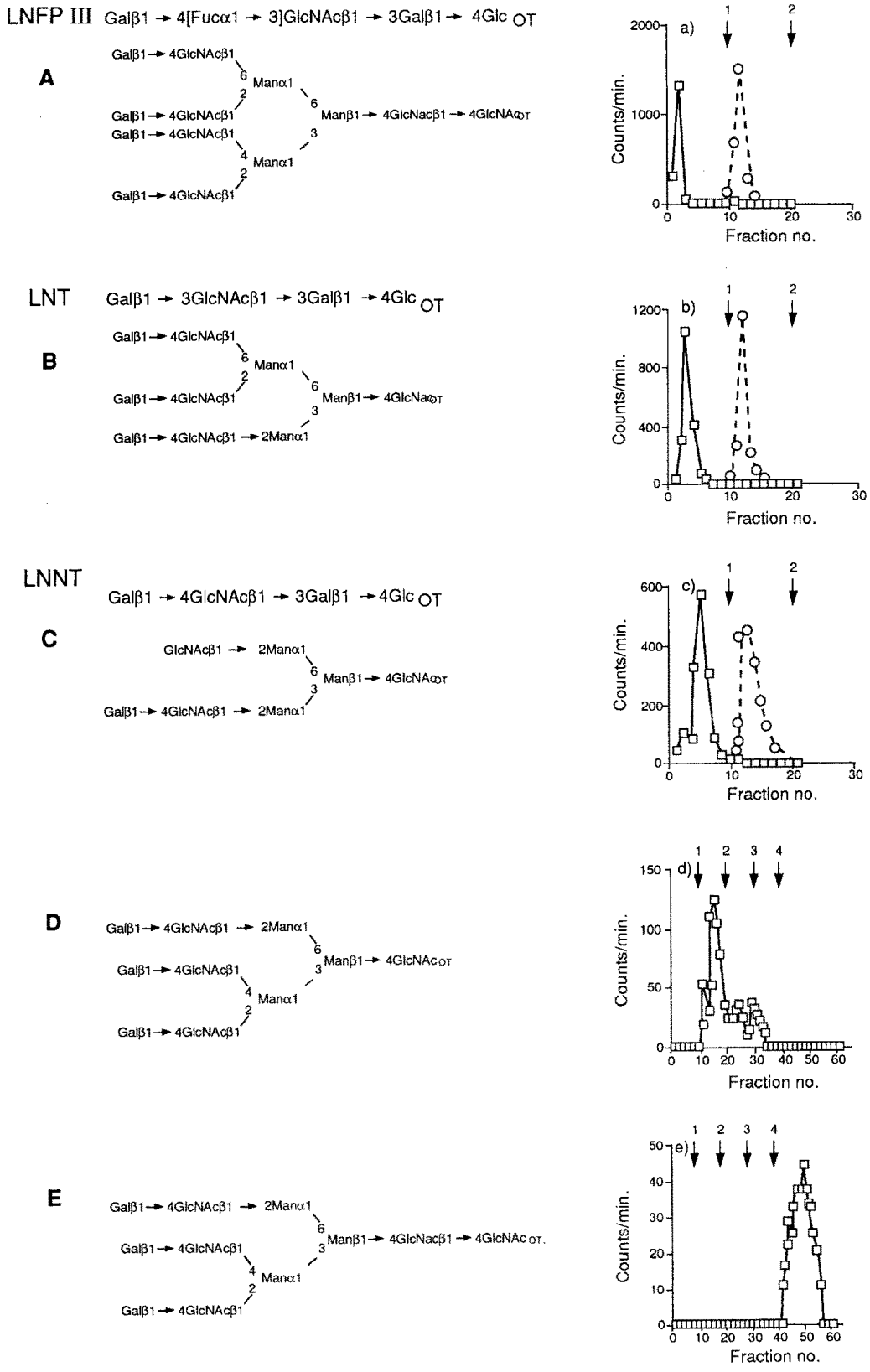

Figure 4. Affinity chromatography of oligosaccharides on TCAII Sepharose. See text for details. Columns were eluted sequentially with buffer at room temperature followed by 10 mM lactose (1) or 100 mM lactose (2) at room temperature, 100 mM (3) and 250 mM (4) lactose at 37 °C. (a) LNFPIII (□) and compound A (○); (b) LNT (□) and compound B (○); (c) LNNT (□) and compound C (○); (d) compound D; (e) compound E.

ricin and RCAI [6, 7]. Interestingly when we compared the sequences within this region it became clear that there exists a relatively higher homology between the pairs TCAI plus ricin B chain and TCAII plus RCAI B chain respectively (Fig. 3). Thus the triplet Gly_{18} Arg Phe (TCAI numbering) also appears in ricin B chain but not RCAI B chain (Fig. 3). By contrast the doublet $Glu_{18}Glu$ and Pro_{25} (TCAII numbering) occur in both TCAII and RCAI B chain (Fig. 3). The significance of these similarities between *T. conophorum* and *R. communis* isolectins remains to be determined. It will also be interesting to determine if the sequence homology between TCAI and TCAII polypeptides is as high as the homology (84%) between the B chains of ricin and RCAI [6], which nevertheless do show significantly different carbohydrate-binding properties, as do TCAI and II based on the relative inhibitory effects of simple haptens on haemagglutinating activities as discussed earlier.

Carbohydrate-binding specificity of TCAII

In order to determine what effect if any the presence of variable but significant amounts of TCAI isolectin would have on the conclusions obtained previously [2] for the carbohydrate-binding specificity of the isolectin mixture consisting mainly of the TCAII isolectin, we coupled TCAII purified by gel filtration and free of TCAI to a matrix support and examined the behaviour of defined oligosaccharides during affinity chromatography. We selected key oligosaccharides that in the earlier study [2] showed up several unique aspects of binding specificity. Each oligosaccharide was labelled at the reducing end with sodium borotritiide and applied to a TCAII Sepharose column which was washed sequentially with sample buffer followed by lactose solutions of different concentrations. As was shown earlier [2] elution of oligosaccharides from TCA Sepharose is extremely sensitive to temperature and hence certain chromatographic steps were performed at 37 °C rather than room temperature to assist in eluting strongly bound oligosaccharides. In general, the results obtained using TCAII Sepharose (Fig. 4) are very comparable with our earlier data [2] using the TCA isolectin mixture. Thus, a straight chain oligosaccharide LNNT (Fig. 4c) containing a Type II Galβ1 → 4GlcNAc terminal sequence was bound more tightly to TCAII Sepharose than LNT terminating in a Type I Galβ1 → 3GlcNAc sequence (Fig. 4b), Substitution of the penultimate *N*-acetylglucosamine residue with a Fucα1 → 3 residue as in LNFPIII absolished any affinity for TCAII Sepharose (Fig. 4a). In branched oligosaccharides highest affinity was shown by a 2.2.4-tri-antennary structure containing a chitobiose core (Fig. 4e) which required elution at 37 °C by high lactose concentrations. Removal of the first *N*-acetylglucosamine residue of the chitobiose core to give compound D (Fig. 4d) resulted in drastic loss of TCAII binding activity as shown earlier for unfractionated TCA [2]. Isomeric tetra-antennary, 2.2.6 tri-antennary and bi-antennary compounds A, B and C respectively (Fig. 4a–c) bound less tightly, as reported earlier [2]. Thus, TCAII is likely to be useful for the separation of glycopeptides or glycoproteins containing the 2.2.4 tri-antennary glycan structure such as fetuin [13] or porcine thyroglobulin [14].

In summary, the results reported here indicate that the carbohydrate-binding specificity reported earlier for *T. conophorum* lectins is clearly reflected by the binding specificity of the major isolectin TCAII of the seeds. We have not determined the binding specificity of the minor isolectin TCAI as yet but it appears, on the basis of simple hapten inhibition of haemagglutination assays, to have preference for *N*-acetylgalactosamine over β galactosides, as does ricin B chain with which TCAI shows greatest homology in *N*-terminal aminoacid sequences.

Acknowledgements

We thank Professor Akintola Aboderin, Ife-Ife and Dr Sachiko Sato, N.I.M.R., for helpful advice.

References

1. Togun RA, Animashaun T, Benzie CR, Kay JE (1988) In *Lectins: Biology-Biochemistry, Clinical Biochemistry.* (Bøg-Hansen TC, Freed DLJ, eds). Vol. 6, pp. 369–72. Sigma Chemical Co.
2. Sato S, Animashaun T, Hughes RC (1991) *J Biol Chem* **266**:11485–94.
3. Funatsu G, Yoshitake S, Funatsu M (1978) *Agric Biol Chem* **42**:501–3.
4. Funatsu G, Kimura M, Funatsu M (1979) *Agric Biol Chem* **43**:2221–4.
5. Bull H, Li S, Fowler E, Lin T (1980) *Int J Pept Protein Res* **16**:208–18.
6. Roberts LM, Lamb FI, Pappin DJC, Lord JM (1985) *J Biol Chem* **260**:15682–6.
7. Araki T, Funatsu G (1987) *Biochim Biophys Acta* **911**:191–200.
8. Animashaun T, Hughes RC (1989) *J Biol Chem* **264**:4657–63.
9. Aitken A, Geisow MJ, Findlay JBC, Holmes C, Yarwood A (1989) In *Protein Sequencing: A Practical Approach* (Geisow MJ, Findlay JBC, eds). pp. 46–68. Oxford: IRL Press.
10. Animashaun T (1983) *Lectins*, Vol. 3 (Bøg-Hansen TC, Spengler GA, eds). pp. 639–44. Berlin: Walter de Gruyter & Co.
11. Araki T, Funatsu G (1985) *FEBS Lett* **191**:121–4.
12. Araki T, Yoshioka Y, Funatsu G (1986) *Biochim Biophys Acta* **872**:277–85.
13. Takasaki S, Kobata A (1986) *Biochemistry* **25**:5706–15.
14. Yamamoto K, Tsuji T, Itimura T, Osawa T (1981) *Biochem J* **195**:701–13.