Sequence studies of peanut agglutinin

M. Lauwereys, A. Foriers*, N. Sharon+ and A.D. Strosberg†

*Protein Chemistry, Institute of Molecular Biology, Free University Brussels (VUB), 65 Paardenstraat, Sint Genesius Rode B-1640, *Medical and Special Biochemistry, Pharmaceutical Institute, Free University Brussels (VUB), 103 Laarbeeklaan, B-1090 Brussels, Belgium, +Department of Biophysics, Weizmann Institute of Science, Rehovot, Israel* and †*Molecular Immunology, Jacques Monod Institute, CNRS and University Paris VII 2, place Jussieu, F75251 Paris Cedex 5, France*

Received 28 November 1984

Sequence studies have been performed on affinity purified peanut agglutinin, a galactose binding lectin. 161 residues have been compared to homologous residues in soybean agglutinin and favin. Extensive similarities have been uncovered, confirming the conservation of lectin sequences among all legume lectins. Evidence is presented for the existence of internal duplications and/or isolectins.

Lectin *Sequence* *Homology* *Legume* *Peanut agglutinin*

1. INTRODUCTION

Lectins from the seeds of leguminous plants constitute a group of homologous sugar binding proteins (review [1]). Among the members of this group the following have been completely sequenced: concanavalin A [2], fava bean lectin [3], lentil lectin [4] and sainfoin [5]. The primary structures of several other lectins namely: soybean agglutinin (SBA) [6], pea lectin [7] and phytohemagglutinin (PHA) [8] have been deduced from the nucleotide sequence of cDNA reversely transcribed from specific mRNA. All these sequences have been found to be homologous, regardless of the existence of one single or two polypeptide chains in the monovalent subunit of the lectin. Other lectins have only been sequenced at their amino terminal portions, which were found to be homologous to those of the lectins sequenced completely [1,9]. Thus, PNA was shown to have an amino terminal sequence which, over the first 32 positions, is 47% identical to that of SBA.

Abbreviations: PNA, peanut agglutinin; SBA, soybean agglutinin; TPCK, L-1-tosylamido-2-phenylethyl-chloromethyl ketone

Because PNA is extensively used to characterize and separate lymphocytes (review [10]) and since it is one of the few lectins for which preliminary X-ray crystallography data are available [11,12] we decided to determine its amino acid sequence. We present here the sequence of 161 residues of PNA lined up with homologous regions of SBA. Despite vigorous efforts, the missing parts could not be characterized, probably because of their hydrophobic nature or because of microheterogeneity due to the existence of isolectins [13].

2. MATERIALS AND METHODS

2.1. *PNA*

PNA extracted from the seeds of the 'Shulamit' variety of *Arachis hypogea* was purified by affinity chromatography on a column of *N*-(.-aminocaproyl) *β*-D-galactopyranosylamine coupled to Sepharose [14] and displayed a single band with a molecular mass of 27 kDa on SDS–PAGE (polyacrylamide gel electrophoresis).

2.2. *Tryptic hydrolysis*

Various digestions were performed:

(A) PNA was digested at 37°C with TPCK-

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/85/$3.30

treated trypsin added to a 5 mg/ml solution of PNA in 1% ammonium bicarbonate. The enzyme/lectin final ratio was 1/100 on a weight to weight basis. Trypsin was added three times over a period of 24 h [15].

(B) The same digestion as in (A) was performed in the presence of 0.1% SDS.

(C) PNA was digested at 37°C with 50 μg TPCK-treated trypsin added to a 5 mg/ml solution of PNA (pre-denatured with 6 M urea) in 1% ammonium bicarbonate and 2 M urea.

2.3. *Chemical hydrolysis by cyanogen bromide in heptafluorobutyric acid/formic acid*

PNA was hydrolyzed in a solution of HFBA/HCOOH (1:1, v/v) using CNBr, in a one to one ratio (weight per weight) with protein. The reaction lasted 24 h.

2.4. *High performance liquid chromatography (HPLC)*

Reverse phase HPLC was performed with a Waters system (Waters, Milford, MA) equipped with two Model 6000A pumps, a variable ultraviolet detector and an automatic sampler (Wisp 710B), employing the following columns: Microbondapack C_{18} and Microbondapack RP-CN (Waters, MA) or Lichrosorb RP-8 (Merck, Darmstadt, FRG). Measurements were made at 230 nm in the sensitivity range of 0.05 to 0.2 absorption units full scale. Flow rates in all experiments were 2 ml/min at room temperature. The peptides were eluted with linear gradients made from 50 mM ammonium acetate (pH 6.5) (solvent A) and methanol (solvent B). The elution started with 100% solvent A and ended after 2 h with 80% solvent B.

2.5. *Amino acid sequence determinations*

Sequence analyses were performed by automated Edman degradation on 5–50 nmol protein and peptides with a Beckman 850C sequencer modified in the laboratory for automated conversion. Several programs were used including the 0.33 M Quadrol program with single or double cleavage [16]. Phenylthiohydantoins were iden-

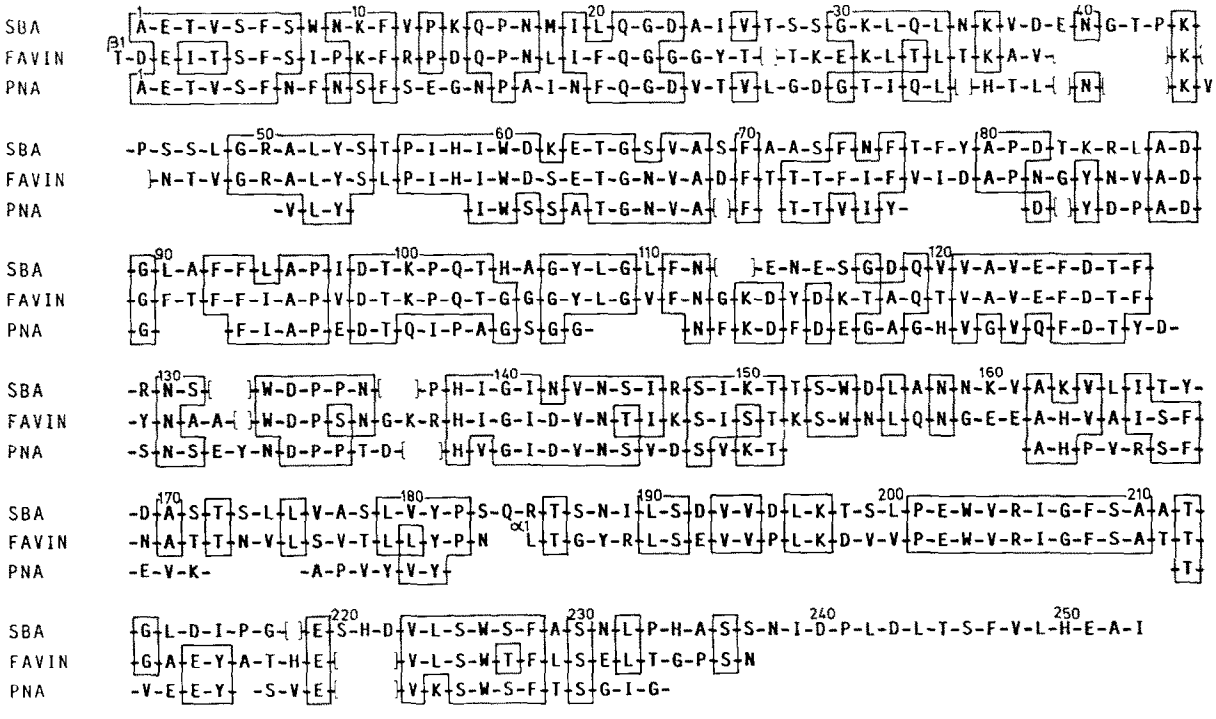

Fig.1. Comparison of the complete sequences of lectins from leguminous plants. The sequences of PNA peptides are placed for maximal homology.

tified by high pressure liquid chromatography using a Waters system and a Microbondapack C_{18} [17] or an Altex Ultrasphere ODS column [18].

3. RESULTS AND DISCUSSION

The 40 amino terminal positions of PNA were sequenced by automated Edman degradation and revealed extensive homology with SBA and favin the lectin of broad bean (fig.1). Additional sequence determinations carried out on peptides obtained mostly from tryptic digests of PNA under various conditions, and separated by HPLC (fig.2) permitted the establishment of nearly 60% of the primary structure (fig.1). The sequences were lined up according to their homology to portions of the lectins of soybean and broad bean (favin). By using this approach we were able to place 161 residues, 69 of which were identical to those of SBA, and 71 to those of favin. However, low solubility and resistance to any kind of hydrolysis have hampered the collection of additional data. In particular, it has been difficult to obtain overlapping sequences which would have made the placing of peptides unambiguous.

A number of tryptic peptides seemed to correspond to either duplicated portions of the protein or to alternate sequences probably present in isolectins. Thus 6 of the first 9 residues of peptide PT

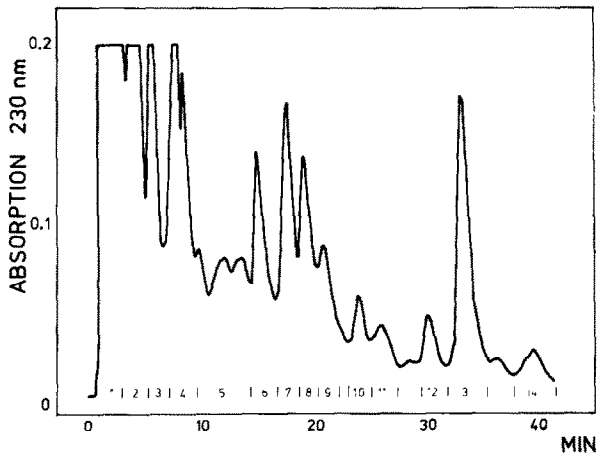

Fig.2. Separation of tryptic peptides of PNA. HPLC on Microbondapack RP-CN using the gradient system as described in section 2, monitored at 230 nm.

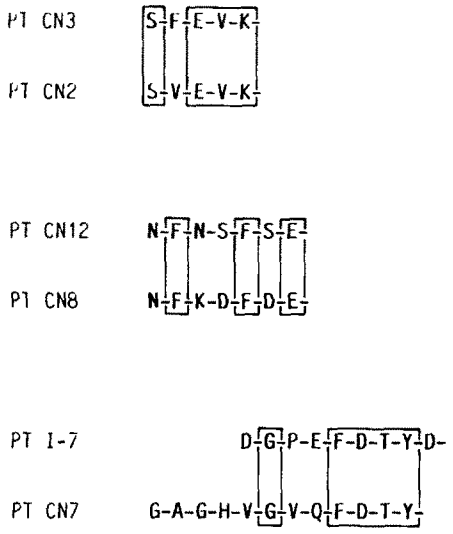

Fig.3. Evidence for homologous peptides present in a single preparation of PNA.

CN7 are identical to those between positions 114 and 125 of SBA and favin (fig.3). However, in a peptide with sequence D-G-P-E-F-D-T-Y-D-, which was also found during peptide mapping, but unpositioned until now, one finds four residues identical to those in positions 120–123 of favin. Three of these are also present in peptide PT CN7. While the maximum of homology favors the first peptide, the simultaneous presence of both sequences in the PNA digest suggest the existence of either internal duplications or isolectins.

REFERENCES

[1] Strosberg, A.D., Lauwereys, M. and Foriers, A. (1983) in: Chemical Taxonomy, Molecular Biology and Function of Plants (Goldstein, I.J. and Etzler, M. eds) vol.138, p.7, Alan R. Liss, New York.

[2] Edelman, G.M., Cunningham, B.A., Reeke, G.N. jr, Becker, J.W., Waxdal, M.J. and Wang, J.L. (1972) Proc. Natl. Acad. Sci. USA 69, 2580.

[3] Hemperly, J.J., Hopp, T.P., Becker, J.W. and Cunningham, B.A. (1979) J. Biol. Chem. 254, 6803.

[4] Foriers, A., Lebrun, E., Van Rappenbusch, R., De Neve, R. and Strosberg, A.D. (1981) J. Biol. Chem. 256, 5550.

[5] Young, N.M., Williams, R.E., Roy, C. and Yaguchi, M. (1982) Can. J. Biochem. 743, 212.

[6] Vodkin, L.O., Rhodes, P.R. and Goldberg, R. (1983) Cell 34, 1023.
[7] Higgins, T.J.V., Chandler, P.M., Zurawski, G., Button, S.C. and Spencer, D. (1983) J. Biol. Chem. 258, 9544.
[8] Hoffman, L.M., Ma, Y. and Barker, R.F. (1982) Nucl. Acids Res. 10, 7819.
[9] Foriers, A., Wuilmart, C., Sharon, N. and Strosberg, A.D. (1977) Biochem. Biophys. Res. Commun. 75, 980.
[10] Sharon, N. (1983) Adv. Immunol. 34, 213.
[11] Olsen, K.W. and Miller, R.L. (1982) FEBS Lett. 145, 303.
[12] Salunke, D.M., Khan, M.I., Surolia, A. and Vijayan, M. (1982) J. Mol. Biol. 154, 177.
[13] Pueppke, S.G. (1981) Arch. Biochem. Biophys. 212, 254.
[14] Lotan, R., Skutelsky, E., Danon, D. and Sharon, N. (1975) J. Biol. Chem. 250, 8515.
[15] Ozols, J. and Gerard, C. (1977) J. Biol. Chem. 252, 5986.
[16] Brauer, A., Margolies, M.N. and Haber, E. (1975) Biochemistry 14, 3029.
[17] Zeeuws, R. and Strosberg, A.D. (1978) FEBS Lett. 85, 68.
[18] Foriers, A., Lauwereys, M. and De Neve, R. (1984) J. Chromatogr. 297, 75.