The EMBO Journal vol.11 no.4 pp.1303-1307, 1992

Non-glycosylated recombinant pro-concanavalin A is active without polypeptide cleavage

W.Min, A.J.Dunn and D.H.Jones[1]

Molecular Biology Research Group, School of Biological Sciences, University College of Swansea, Swansea, SA2 8PP, Wales, UK
[1]Corresponding author

Communicated by G.Warren

The complex post-translational processing of concanavalin A (Con A) in maturing jackbeans is unique because the non-glycosylated mature active protein is circularly permuted in primary sequence relative to its own inactive precursor (glycosylated pro-Con A) and to other legume lectins. We show here that non-glycosylated pro-Con A expressed in bacteria from recombinant cDNA (rec-pro-Con A) folds *in vivo* and *in vitro* to a stable form which is active without further processing. *N*-glycosylation alone must therefore be sufficient to inactivate pro-Con A—a novel role for glycosylation in regulating activity during protein maturation.

Key words: concanavalin A/glycosylation/lectin/post-translational processing/recombinant protein

Introduction

Concanavalin A (Con A) 'the most celebrated of the plant lectins' (Goldstein and Poretz, 1986) is the only known natural protein which is a circular permutation of another (Goldenberg, 1989). The mature form of Con A is unique in being circularly permuted in sequence relative to favin (Cunningham *et al.*, 1979) and other legume lectins (Strosberg *et al.*, 1986). Carrington *et al.* (1985) showed that the amino acid sequence derived from a cDNA for Con A precursor has direct, not circular, homology with other legume lectins. It was suggested, in the face of some controversy (Gatehouse and Boulter, 1985), that a post-translational transposition and ligation (at mature residues 118 and 119) of two polypeptides must occur and that this processing was consistent with the three-dimensional structure of mature Con A (Figure 3). Bowles *et al.* (1986) demonstrated independently that such a transposition of fragments occurs by using pulse-chase experiments to follow processing *in vivo* in immature Jack Bean cotyledons. Immunoprecipitation, peptide mapping and N-terminal sequencing of abundant precursors revealed a complex series of events involving transient glycosylation and polypeptide cleavages to yield mature active lectin by religation at positions 118 and 119.

Our understanding of this novel form of protein maturation (Sharon and Lis, 1986) is complicated by glycosylation of the initial precursor (Figure 3), which we call pro-Con A, to give an apparent molecular weight of 33 500 (Bowles *et al.*, 1986; Bowles and Pappin, 1988). Whereas mature Con A is not a glycoprotein (Goldstein and Poretz, 1986), pro-Con A made by the plant carries one N-linked oligosaccharide of the oligomannose type (Herman *et al.*, 1985) and this glycosylated pro-Con A is unable to bind to a cross-linked dextran (Chrispeels *et al.*, 1986) or ovalbumin (Bowles *et al.*, 1986) affinity matrix. The oligosaccharide is removed *in vivo* before proteolytic cleavage since only the initial precursor (apparent M_r = 33 500) is found to be glycosylated (Marcus and Bowles, 1988) and the second species detected by pulse-chase analysis is only slightly smaller (apparent M_r = 30 600). Processing ultimately generates the active lectin, but activation could involve deglycosylation alone to leave a loop of peptide or could necessitate removal of a glycopeptide (Herman *et al.*, 1985) near the ligand-binding site (Bowles and Pappin, 1988) (Figure 3).

It was clearly necessary to produce non-glycosylated pro-Con A to resolve the question of the activation step and examine the properties of the protein. Glycosylated pro-Con A may be obtained from immature seeds for subsequent direct deglycosylation (Sheldon and Bowles, 1992), but biological constraints limit the quantity which may be prepared from its natural source. We have expressed pro-Con A in *Escherichia coli* to see whether the non-glycosylated recombinant protein (rec-pro-Con A) can fold successfully. This is also the initial stage of a wider programme on protein engineering various forms of this lectin.

Results

Expression was accomplished with the pIN-III-ompA vector (Ghrayeb *et al.*, 1984) which uses the *ompA* leader to direct a fusion protein into the periplasmic space where cleavage of the bacterial signal occurs. The coding sequence for pro-Con A was manipulated (Figure 1) so that *Escherichia coli* signal peptidase would generate the authentic N-terminus (corresponding to mature residue 119 = Ser). Constructs were checked by DNA sequencing, which also revealed three minor differences (Figure 1a) from the sequence originally reported (Carrington *et al.*, 1985): nucleotide 360 [•] is A not C which gives the sequence ...GCAAGCTT... a *Hin*dIII site; nucleotide 246 is T not G so that residue 33 (mature Con A 151) is Asp not Glu; nucleotides 256 and 257 are G and A not A and G so that residue 37 (mature Con A 155) is Glu not Arg. Both of these residues now agree with the directly determined protein sequence of mature Con A (Cunningham *et al.*, 1975).

Induction of expression (Figure 2) was generally followed by a reduction in growth rate, the appearance of soluble product and formation of protein aggregates (Ghrayeb *et al.*, 1984). The major portion of the product was recovered from the cell debris after sonication by solubilization with guanidine hydrochloride. Dilution (rather than dialysis which resulted in co-precipitation) was used for refolding before affinity purification on cross-linked dextran. SDS-PAGE showed that total cell extract (Figure 2, lane 3) treated by

the simple purification steps described yielded very pure rec-pro-Con A (Figure 2, lane 2). The affinity column facilitated the removal of all other proteins including any pre-pro-Con A. Approximately 200 μg rec-pro-Con A was purified from a 500 ml culture. N-terminal sequencing, performed on a MilliGen/Biosearch 6600 ProSequencer after dialysis of the purified protein against 0.1% SDS, showed exact correspondence for 19 residues with the sequence predicted from pIN-pro. (For cycles 2–18 the average repetitive sequence yield was 95% and initial sequence yield was 27 pmol). No other N-terminal sequences were detected. Glycosylated pro-Con A isolated from jackbeans has been sequenced for the first 12 residues (Bowles *et al.*, 1986; Chrispeels *et al.*, 1986) and our results match exactly. Our results thus provide further confirmation that residue 122 of mature Con A (4 of pro-Con A) is Glu not Gln

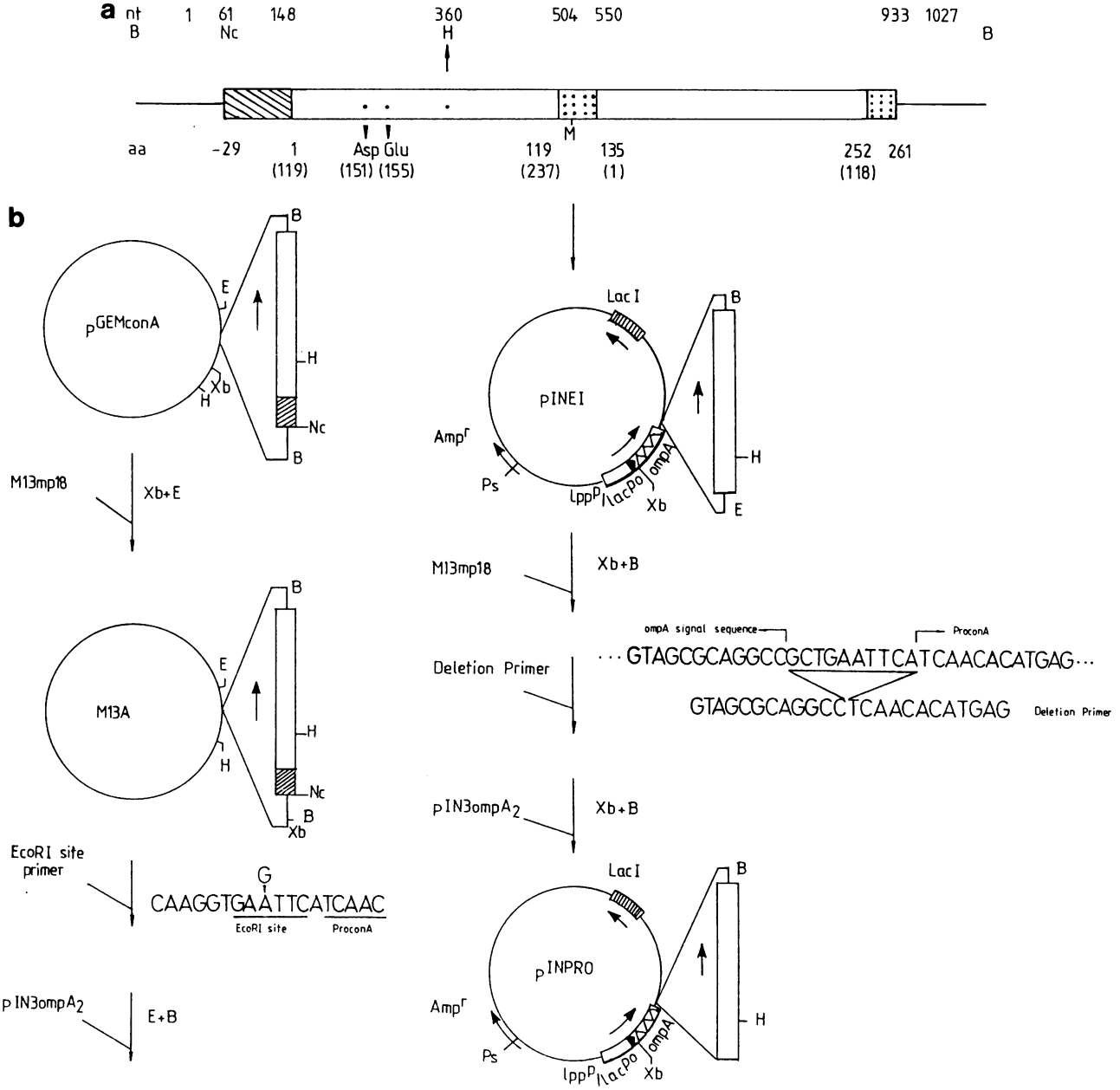

Fig. 1. **(a)** Map of pre(plant signal)-pro-Con A coding sequence (Carrington *et al.*, 1985) derived from cDNA: nt, nucleotide number; aa, amino acid residue numbers of pro-Con A (corresponding amino acid residue numbers of mature Con A in parenthesis); ▨, plant signal peptide residues −29 – −1; ▩, spacer peptide residues 120–134 and C-terminal extension residues 253–261; M , Asn123 glycosylation site (Bowles *et al.*, 1986; Chrispeels *et al.*, 1986). Other abbreviations as in (b). For explanation of sites marked (•) see Results.**(b)** Construction of rec-pro-Con A expression vector pIN-pro. The coding sequence was excised from the plasmid pGEM-Con A as an *Xba*I–*Eco*RI fragment. After cloning to M13mp18, an *Eco*RI primer (shown on right) was made to alter one base (G143→A) to generate a new *Eco*RI site just upstream of the first codon of the pro-Con A sequence. The fragment from this new *Eco*RI site to the 3′ *Bam*HI site was cloned to pIN-III-ompA2 (Ghrayeb *et al.*, 1984) to generate pIN-EI. However, pIN-EI carries the pro-Con A sequence in the wrong reading frame and has an intervening sequence between the *ompA* signal peptide and the N-terminus of pro-Con A. After cloning the *Xba*I–*Bam*HI fragment to M13mp18, the deletion primer (shown on right) was used to remove this intervening sequence and place the 21 amino acid *ompA* signal adjacent to the authentic N-terminus of pro-Con A. The fragment was then reinserted in the *Xba*I and *Bam*HI sites of pIN-III-ompA2 to create the final construct pIN–pro. This construct was checked by DNA sequencing and was then used to express rec-pro-Con A in the periplasm of *E.coli*. E, *Eco*RI; H, *Hin*dIII; B, *Bam*HI; Xb, *Xba*I; Nc, *Nco*I; Ps, *Pst*I; lpp^p, *lpp* promoter; lac^{po}, *lac* promoter-operator; ompA, *ompA* signal sequence.

(Cunningham *et al.*, 1975) and residue 124 of mature Con A (6 of pro-Con A) is Asn not Asp (Cunningham *et al.*, 1975).

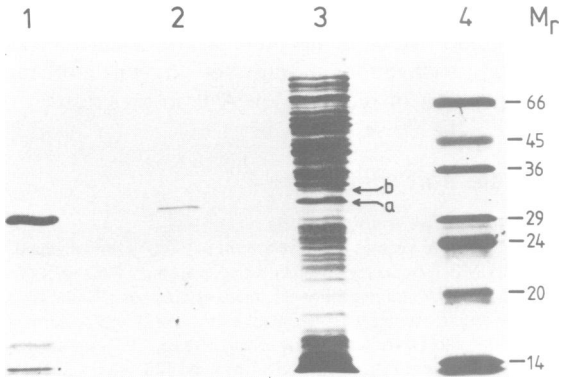

Fig. 2. Expression of rec-pro-Con A in *E.coli* and purification by affinity chromatography. SDS–polyacrylamide gel (15%) stained with Coomassie Blue: lane 1, standard Con A (Sigma, Type IV); lane 2, purified recombinant pro-Con A; lane 3, total cell protein from *E.coli* W620*recA*(pIN-pro); arrow a shows the position of rec-pro-Con A, arrow b the position of pre(ompA signal)-pro-Con A inferred from Western blots; lane 4, molecular weight markers. [Mature Con A (lane 1), M_r = 25 600 (calculated from amino acid sequence (Cunningham *et al.*, 1975) plus metal ions) to 26 500 [generally reported (Goldstein and Poretz, 1986)], migrates anomalously to a position corresponding to 30 000 on SDS–PAGE (Herman *et al.*, 1985; Bowles *et al.*, 1986). Rec-pro-Con A (lane 2) predicted M_r = 28 200 from translated DNA sequence migrates to approximate position 32 000].

Soluble product was always present after induction and could be recovered directly (i.e. without denaturation and refolding) by affinity adsorption from the supernatant after sonication. N-terminal sequence analysis (performed as above) confirmed that this soluble product was rec-pro-Con A. We have subsequently found that a further reduction in growth temperature to 23°C results in an increased amount of soluble rec-pro-Con A. Osmotic shock (Nossal and Heppel, 1966) also released soluble rec-pro-Con A which could adsorb to Sephadex G-75 and was identified by Western blotting. Pre(ompA)–pro-Con A was detected in the cell pellet after sonication (Figure 2) but could not be found in the periplasmic fraction. No bands from a control culture harbouring pIN-III-ompA2 correspond to rec-pro-Con A on Western blots.

Discussion

Some of the rec-pro-Con A produced in our system must therefore fold *de novo* to a soluble active form in *E.coli* periplasm. The major portion may be correctly folded and be bound to cell components or precipitate if its threshold of solubility is exceeded—this limit appears to be lower *in vitro* than for mature Con A. [Directly deglycosylated plant pro-Con A is also reduced in solubility (Sheldon and Bowles, 1992)]. Alternatively, in the heterologous environment the major portion of rec-pro-Con A may be incorrectly folded and therefore may aggregate (Mitraki and King, 1989) in the periplasm (Ghrayeb *et al.*, 1984).

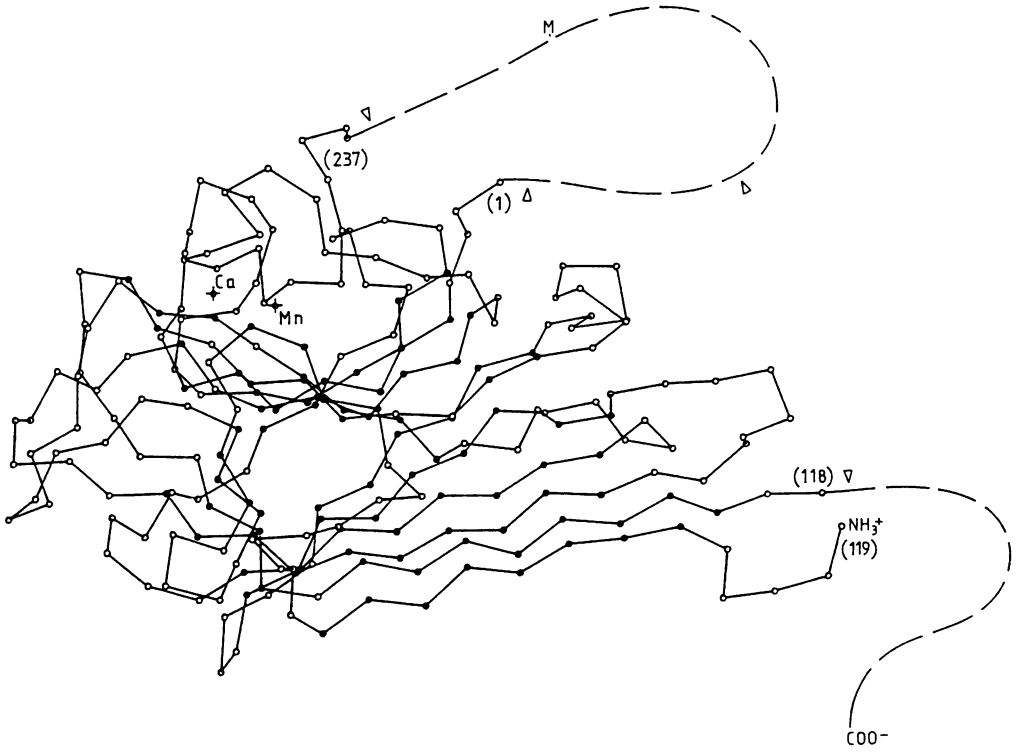

Fig. 3. Putative tertiary structure for pro-Con A (after Carrington *et al.*, 1985; Bowles *et al.*, 1986; Chrispeels *et al.*, 1986). Circles represent α-carbon atoms with coordinates taken from the 2.0Å resolution structure for the mature Con A protomer [2CNA, Brookhaven Protein Data Bank and Reeke *et al.* (1975)]. Infilled circles represent α-carbons in β-sheets; open circles, those not forming part of regular secondary structures. The saccharide-binding site is near the Ca^{2+} and Mn^{2+} ions. The 15 residue surface loop and 9 residue C-terminal extension are of undetermined conformation and are shown as broken lines (— — — —) sketched approximately to scale for an extended polypeptide chain. Residue numbers of mature Con A are shown in brackets as in Figure 1. The site of co-translational *N*-glycosylation in the plant is shown (M). During processing in jackbeans, pro-Con A is deglycosylated, an endopeptidase cleaves △ on the C-side of four asparagine residues (Bowles *et al.*, 1986) and residues 118 and 119 are ligated.

Pea lectin expressed from a cDNA sequence forms insoluble aggregates in the cytoplasm of *E.coli* (Stubbs *et al.*, 1986) which may be refolded from the denaturant *in vitro* to give an active protein. The product is a fusion protein which is not processed in the bacterial cell and thus contains extra amino acid residues in three regions (Stubbs *et al.*, 1986). After refolding it is active as a single chain even though the native pea lectin is cleaved to give one α- and one β-chain per protomer. Recombinant pea lectin is homologous to our recombinant pro-Con A, rather than to mature Con A. Our findings are therefore consistent with those of Stubbs *et al.* (1986). The only processing which appears to take place in our system is cleavage of the *ompA* signal (Figure 2) to generate the authentic N-terminus. It has long been known (Becker *et al.*, 1976)—since fragmented Con A monomers occur naturally (Figure 2, lane 1)—that the absence of a peptide bond between residues 118 and 119 (Figure 3) does not affect either the crystal structure or the biological activity of the mature molecule.

The cDNA used in our work (Figure 1) was originally obtained from *Canavalia ensiformis* (Carrington *et al.*, 1985). A cDNA for Con A has since been cloned from the closely related *Canavalia gladiata* (Yamauchi and Minamikawa, 1990) and on expression in *E.coli* no precursor forms [either pre(plant signal)-pro-Con A or pro-Con A] could be detected. Some evidence has been presented (Yamauchi and Minamikawa, 1990) which implies that complete processing occurs in order to generate mature Con A (of unstated cellular location). This is in contrast to our system where rec-pro-Con A is clearly stable and no evidence of *in vivo* processing akin to that in the plant (Bowles *et al.*, 1986) or consequent mature product was seen (Figure 2).

Most significantly, our results are completely consistent with those of Sheldon and Bowles (1992) for direct deglycosylation *in vitro* by *N*-glycanase of glycosylated pro-Con A isolated from jackbeans.

In this paper we have primarily addressed the question of the role of *N*-glycosylation of a precursor protein. Non-glycosylated recombinant pro-Con A is active in carbohydrate binding, as shown (Figure 2) by its adsorption to cross-linked dextran and elution by a monosaccharide derivative for which native Con A has high affinity (Goldstein and Poretz, 1986). N-terminal sequencing substantiates electrophoretic and blotting evidence that further (proteolytic) processing is not necessary for activation. We have thus shown that it must be glycosylation alone and not the circular permutation of primary structure (relative to mature Con A), and the consequent 15 residue surface loop (Figure 3)—which is responsible for inactivation of pro-Con A synthesized *de novo* in the plant system. Inhibition of glycosylation by tunicamycin treatment of maturing jackbeans inhibits transport of pro-Con A (Faye and Chrispeels, 1987). Inactivation of this initial lectin precursor by co-translational *N*-glycosylation appears to be essential for its transport from the rough endoplasmic reticulum and subsequent intracellular trafficking (Bowles and Pappin, 1988; Sheldon and Bowles, 1992). Glycosylation may also increase the solubility of pro-Con A. Since deglycosylation without peptide cleavage must be sufficient to activate carbohydrate binding this implies that subsequent steps in the plant (Bowles and Pappin, 1988) must process already active forms of this lectin. We are not aware of any precedents (Rademacher *et al.*, 1988; Paulson, 1990) for glycosylation as a mechanism of inactivation of protein precursors. The oligosaccharide chain might lie within the binding site of the same or an adjacent subunit thus self-neutralizing the lectin, or it might otherwise stabilize an inactive conformation of this protein. The detailed molecular basis for inactivation is not yet known and further characterization of rec-pro-Con A is in progress.

Materials and methods

Recombinant DNA manipulations

The pIN-III-ompA vectors (Ghrayeb *et al.*, 1984) were obtained from M.Inouye (UMD New Jersey) together with *E.coli* strain W620*recA* (Wurtzel *et al.*, 1981). DNA cloning followed standard methods (Sambrook *et al.*, 1989). Site-specific mutagenesis (Kunkel *et al.*, 1987) was performed with a Bio-Rad kit and DNA sequencing (Sanger *et al.*, 1977) used the M13 system. Primers were synthesized on an Applied Biosystems 391EP DNA synthesizer.

Growth, induction and lysis of cells

E.coli W620*recA*(pIN-pro) was grown in 500 ml M9 medium (Sambrook *et al.*, 1989) supplemented with 0.5% (w/v) cas-amino acids, glucose (4 mg/ml), thiamine (2 μg/ml), glutamic acid (300 μg/ml), uracil (40 μg/ml) and ampicillin (100 μg/ml) at 37°C, 250 r.p.m., to an A_{600} of 0.5–0.6. Isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 1 mM and the temperature reduced to 30°C (or below) for a further 5–6 h. Cells were then collected and resuspended in 50 ml MOPS–metals buffer [20 mM MOPS [3-(*N*-morpholino)propanesulphonic acid] pH 7.0, 1 mM $CaCl_2$, 1 mM $MnCl_2$, 1 M NaCl] and lysed by sonication (3×30 s) on ice. The lysate was placed on ice for 1 h followed by centrifugation at 27 000 *g* for 30 min at 4°C.

Detection of expressed product

Cell fractions were checked for the presence of recombinant products by SDS–PAGE (Laemmli, 1970) and Western blotting using a Bio-Rad Immuno-Blot assay kit (alkaline phosphatase) and either rabbit anti-Con A serum (Sigma) or rabbit anti-SDS-Con A serum prepared by ourselves. Antibody preparation and Western blotting procedures followed standard methods (Harlow and Lane, 1988).

Refolding aggregated product

The pelletted cell debris was resuspended in 2 ml denaturant (7 M guanidine hydrochloride in MOPS–metals buffer) and placed on a rotator overnight at 4°C. Following centrifugation (27 000 *g* for 30 min at 4°C) supernatant was mixed for 30 min with 0.5 ml DEAE–Sephacel (Pharmacia) equilibrated in the same denaturant (Stubbs *et al.*, 1986), then centrifuged as before. The DEAE–Sephacel pellet was discarded and the supernatant containing unbound protein diluted at least 30-fold with MOPS–metals buffer.

Affinity chromatography

Either cell lysate containing *de novo* soluble product or product solubilized from pelleted cell debris by denaturant and refolded (as above), was passed down a column (10×65 mm) of Sephadex G-75 (Pharmacia, dry particle size 40–120 μm). This acts as an affinity adsorbent for Con A (Goldstein and Poretz, 1986) and the concentration of methyl α-D-mannopyranoside necessary to elute native Con A (Sigma, Type IV) was found to be 1 mM. (For native Con A, 85.1% of the applied protein [A_{280}] bound to the column; 77.5% was specifically eluted by methyl α-D-mannopyranoside and 7.6% could not be recovered. A new column of affinity matrix was therefore used for each separation to avoid cross-contamination. The proportion of rec-pro-Con A binding to the matrix was estimated by densitometry on Western blots of crude and purified samples in a Joyce Loebl Chromoscan 3. From a crude sample, 83.2% of the detectable rec-pro-Con A bound to the column; 72.2% was specifically eluted by methyl α-D-mannopyranoside and 11.0% remained bound to the Sephadex G-75.)

The sample was loaded at a low flow rate (0.05 ml/min) and the column washed (1 ml/min) to base line A_{280} with MOPS–metals buffer, then eluted (1 ml/min) with 1 mM methyl α-D-mannopyranoside in the same buffer whence only one A_{280} peak emerged. This was collected and then concentrated by centrifugal ultrafiltration (Centricon 10, Amicon), recovered in 0.5 ml MOPS–metals buffer, and analysed by SDS–PAGE (Laemmli, 1970) and Western blotting.

Acknowledgements

We are indebted to D.M.Carrington and D.E.Hanke for generously providing pre-pro-Con A cDNA, M.Inouye for supplying pIN-III-ompA vectors, and J.N.Keen and J.B.C.Findlay for performing N-terminal sequencing at the University of Leeds and for discussions, N.A.Ratcliffe for help with immunisations and D.J.Bowles, P.S.Sheldon, J.M.Lord, L.Roberts and H.B.Dincturk for stimulating discussions. We thank R.Waters for the use of molecular graphics facilities provided by the Wolfson Foundation. This work was initiated with the support of the Nuffield Foundation, Research Corporation Trust and Society for General Microbiology and continued with a grant (GR/E 70641) from the SERC.

References

Becker,J.W., Cunningham,B.A., Reeke,G.N.,Jr, Wang,J.L. and Edelman,G.M. (1976) In Bittiger,H. and Schnebli,H.P. (eds), *Concanavalin A as a Tool*. John Wiley, London, pp. 33–54.

Bowles,D.J. and Pappin,D.J. (1988) *Trends Biochem. Sci.*, **13,** 60–64.

Bowles,D.J., Marcus,S.E., Pappin,D.J.C., Findlay,J.B.C., Eliopoulos,E., Maycox,P.R., and Burgess,J. (1986) *J. Cell Biol.*, **102,** 1284–97.

Carrington,D.M., Auffret,A. and Hanke,D.E. (1985) *Nature*, **313,** 64–67.

Chrispeels,M.J., Hartl,P.M., Sturm,A. and Faye,L. (1986) *J. Biol. Chem.*, **261,** 10021–10024.

Cunningham,B.A., Wang,J.L., Waxdal,M.J. and Edelman,G.M. (1975) *J. Biol. Chem.*, **250,** 1503–1512.

Cunningham,B.A., Hemperly,J., Hopp,T.P. and Edelman,G.M. (1979) *Proc. Natl. Acad. Sci. USA*, **76,** 3218–3222.

Faye,L. and Chrispeels,M.J. (1987) *Planta*, **170,** 217–224.

Gatehouse,J.A. and Boulter,D. (1985) *Nature*, **313,** 13.

Ghrayeb,J., Kimura,H., Takahara,M., Hsiung,H., Masui,Y. and Inouye,M. (1984) *EMBO J.*, **3,** 2437–2442.

Goldenberg,D.P. (1989) *Protein Engng.*, **2,** 493–495.

Goldstein,I.J. and Poretz,R.D. (1986) In Liener,I.E., Sharon,N. and Goldstein,I.J. (eds), *The Lectins: Properties, Functions, and Applications in Biology and Medicine*. Academic Press, London, pp. 33–247.

Harlow,E. and Lane,D. (1988) *Antibodies, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Herman,E.M., Shannon,L.M. and Chrispeels,M.J. (1985) *Planta*, **165,** 23–29.

Kunkel,T.A., Roberts,J.D. and Zabour,R.A. (1987) *Methods Enzymol.*, **154,** 367–382.

Laemmli,U.K. (1970) *Nature*, **227,** 680–684.

Marcus,S.E. and Bowles,D.J. (1988) *Protoplasma*, **147,** 113–116.

Mitraki,A. and King,J. (1989) *Bio/Technology*, **7,** 690–697.

Nossal,N.G. and Heppel,L.A. (1966) *J. Biol. Chem.*, **241,** 3055–3062.

Paulson, J.C. (1990) In Bradshaw,R.A. and Purton,M. (eds), *Proteins: Form and Function*. Elsevier, Cambridge, pp. 209–217.

Rademacher,T.W., Parekh,R.B. and Dwek,R.A. (1988) *Annu. Rev. Biochem.*, **57,** 785–838.

Reeke,G.N.,Jr, Becker,J.W. and Edelman,G.M. (1975) *J. Biol. Chem.*, **250,** 1525–1547.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74,** 5463–5467.

Sharon,N. and Lis,H. (1986) *Nature*, **323**, 203–204.

Sheldon,P.S. and Bowles,D.J. (1992) *EMBO J.*, **11**, 1297–1301.

Strosberg,A.D., Buffard,D., Lauwereys,M. and Foriers,A. (1986) In Liener,I.E., Sharon,N. and Goldstein,I.J. (eds), *The Lectins: Properties, Functions, and Applications in Biology and Medicine*. Academic Press, London, pp. 249–264.

Stubbs,M.E., Carver,J.P. and Dunn,R.J. (1986) *J. Biol. Chem.*, **261,** 6141–6144.

Wurtzel,E.T., Movva,N.R., Ras,F. and Inouye,M. (1981) *J. Mol. Appl. Gen.*, **1,** 61–69.

Yamauchi,D. and Minamikawa,T. (1990) *FEBS Lett.*, **260,** 127–130.

Received on November 19, 1991; revised on January 15, 1992