Production and comparison of mature single-domain ‘trefoil’ peptides pNR-2/pS2 Cys⁵⁸ and pNR-2/pS2 Ser⁵⁸

Mark P. CHADWICK, Felicity E. B. MAY and Bruce R. WESTLEY*

Department of Pathology, University of Newcastle upon Tyne, Royal Victoria Infirmary, Newcastle upon Tyne NE1 4LP, U.K.

The preparation and purification of recombinant mature pNR-2/pS2, a single-domain member of the ‘trefoil’ family of cysteine-rich secreted proteins, is described. Analysis of recombinant pNR-2/pS2 by ion-exchange chromatography showed that it was heterogeneous. The heterogeneity was reduced by treatment with thiol-group-containing reagents, suggesting that it is caused by the odd number of cysteine residues in mature pNR-2/pS2, and this view was reinforced by mutation of the extra-trefoil domain cysteine residue, Cys⁵⁸, to a serine residue. Electrophoresis of recombinant pNR-2/pS2 Cys⁵⁸ and pNR-2/pS2 Ser⁵⁸ proteins under non-denaturing conditions confirmed that the Ser⁵⁸ mutant is much more homogeneous, and showed that most of pNR-2/pS2 Ser⁵⁸ co-migrates as a single band with

pNR-2/pS2 secreted from breast-cancer cells in culture. Treatment of recombinant pNR-2/pS2 proteins with various thiol-group-reactive reagents indicated that cysteine is the most effective at producing recombinant pNR-2/pS2 that co-migrates with pNR-2/pS2 secreted by breast-cancer cells. Dithiothreitol appeared to denature the proteins, and GSH was relatively ineffective. pNR-2/pS2 Cys⁵⁸ treated with cysteine and untreated pNR-2/pS2 Ser⁵⁸ had the same apparent molecular mass, measured by gel filtration, as pNR-2/pS2 secreted from breast-cancer cells. This is the first report of the production of a recombinant mature single-domain trefoil peptide and should greatly facilitate elucidation of the structure and function of pNR-2/pS2.

INTRODUCTION

The mRNA encoding the pNR-2/pS2 protein was first identified in MCF-7 breast-cancer cells by virtue of its regulation by oestradiol [1,2]. The pNR-2/pS2 protein is synthesized as a preprotein of 84 amino acids, and is secreted as a mature protein of 60 amino acids with a molecular mass of ~ 6.5 kDa [3]. Transcription of the pNR-2/pS2 gene is highly inducible by oestradiol and is inhibited by anti-oestrogens in oestrogen-responsive breast-cancer cell lines [4,5]. The oestrogen-regulated transcription is controlled from an imperfect oestrogen response element [6].

Recently, a large number of studies have been published on the clinical significance of pNR-2/pS2 expression in breast cancer. Early reports suggested that pNR-2/pS2 expression might be a marker of good prognosis [7], but subsequent studies have not confirmed its prognostic value [8,9]. There is, however, a growing consensus that the pNR-2/pS2 protein is a predictive marker of response to endocrine therapy. Its expression is associated with that of the oestrogen receptor and is more predictive of a response to hormone therapy than expression of the oestrogen receptor [8,10]. Because the function of the pNR-2/pS2 protein is unknown, it is not established whether the effects of tamoxifen on tumour regression are mediated via an effect on pNR-2/pS2 expression or whether it simply provides a marker of oestrogen responsiveness.

pNR-2/pS2 is a member of a rapidly growing family of proteins now referred to as ‘trefoil’ peptides [11,12]. These peptides share sequence similarity within a 42–43-amino-acid domain, characterized by six conserved cysteine residues and a PWCF motif, that folds to produce a compact globular structure

(Fig. 1). Trefoil peptides contain either one copy of this domain or two copies separated by up to seven amino acids. The prototype trefoil peptide porcine pancreatic spasmolytic polypeptide (PSP) contains two domains. The pattern of disulphide bonds reported for domain I of PSP on the basis of Edman degradation of tryptic peptides was either 1–5, 2–4, 3–6 or 1–4, 2–5, 3–6 [12]. The structure of PSP has recently been solved by crystallographic [13] and NMR [14] techniques, and this has confirmed the former configuration. No structural information is available for a single-domain trefoil, although Thim [12] has speculated that the pattern of disulphide bonds is the same as in PSP (Figures 1a and 1b). Trefoil peptides usually contain a cysteine residue located three amino acids from the C-terminus. In PSP, this cysteine forms an intramolecular disulphide bond with a second non-trefoil domain cysteine located near to the N-terminus. The ability of this C-terminal cysteine residue to form intra- or inter-molecular disulphide bonds in single domain trefoil peptides such as pNR-2/pS2 is unknown.

The normal biological function of pNR-2/pS2 has yet to be identified. The protein is expressed at low levels in a proportion of normal breast epithelial cells [15] and at higher levels in normal stomach mucosa [16]. Two other human trefoil peptides have been described: a two-domain protein human spasmolytic polypeptide (hSP; [17]), and a single domain protein human intestinal trefoil factor (hITF; [18,19]). pNR-2/pS2, hITF and hSP are co-expressed in an ulcer-associated cell lineage found in patients with Crohn’s disease [18,20,21]. It has been suggested that these trefoil peptides might be involved in the repair of damaged gastrointestinal tissue. It has also been noted that many tissues that express trefoil peptides are associated with production of a protective mucous layer; in particular, expression of pNR-

Abbreviations used: PSP, pancreatic spasmolytic polypeptide; hITF, human intestinal trefoil factor; DTT, dithiothreitol; IGF-1, insulin-like growth factor I.

* To whom correspondence should be addressed.

Figure 1 Proposed trefoil structure of pNR-2/pS2

(a) Diagrammatic representation of disulphide-bond formation in pNR-2/pS2 by analogy with PSP. (b) The 42 amino acids of the conserved trefoil domain are indicated in black with the proposed intradomain disulphide bonds. The amino acids outside of the trefoil domain are shown stippled. The seventh cysteine residue, at position 58, is shown in white and in its reduced form.

2/pS2 and hSP has been linked with expression of the neutral mucin MUC-1 [22]. Circumstantial evidence that trefoils may be associated with mucin function is provided by the *Xenopus* protein FIMA1, which contains a mucin-like O-glycosylated core flanked by two pairs of trefoil domains [23].

The availability of significant quantities of recombinant pNR-2/pS2 would facilitate both structural analysis and identification of its biological function. We describe here the production of mature pNR-2/pS2 protein, which is similar to that secreted from MCF-7 breast-cancer cells. In addition, a mutated form of pNR-2/pS2 is described in which Cys⁵⁸ has been replaced by serine to facilitate examination of the influence of this cysteine residue on the characteristics of recombinant pNR-2/pS2 protein.

EXPERIMENTAL

Construction of pNR-2/pS2 expression plasmid

A 180 bp portion of the pNR-2/pS2 cDNA, sufficient to code for the mature protein, was amplified from the pNR-2 cDNA clone 80B12 [24] by PCR. The 5' PCR primer includes the recognition site for the *EcoRI* restriction endonuclease, adjacent to 12 bases that encode an in-frame consensus four-amino-acid recognition sequence for Factor Xa, followed by the 18 bases of the pNR-

Figure 2 Construction of recombinant pNR-2/pS2 expression plasmids

(a) The 5' and 3' PCR primers used to generate the DNA cassette for insertion into the expression vector are shown. The restriction-enzyme sites used in the cloning are indicated. The codons for the amino acids that constitute the Factor Xa recognition site and the N-terminus and C-terminus of the pNR-2/pS2 protein are delineated. Also shown is the mutagenic primer used to convert the codon for Cys⁵⁸ into a serine codon. (b) Diagrammatic representation of the essential features of the pEZZ18 vector, with insertion of the pNR-2/pS2 expression cassette. Abbreviations: ori, origin; Plac, lac promoter; P spa, spa promoter; S-Z-Z sequence, pre-sequence and two z domains; lac z', lac z gene; f1 ori, f1 origin of replication; Amp^R, ampicillin-resistance gene.

2/pS2 cDNA that encode the first six amino acids of the mature protein:

5'GCGAATTCGATCGAAGGGAGTGAGGCC-AGACAGAGACG-3' (Figure 2a)

Factor Xa is a specific proteinase that cleaves immediately 3' of its four-amino-acid recognition site. The 3' PCR primer contains

the antisense strand for the last seven amino acids and the stop codon of pNR-2/pS2 followed by the recognition site for the *Xba*I restriction endonuclease:

5'GCTCTAGATCAAAATTCACACTCCTCTTCTGG3'
(Figure 2a)

The PCR product was purified, ethanol-precipitated and digested with *Eco*RI and *Xba*I. The digested DNA was phenol-extracted, ethanol-precipitated and ligated into the multiple cloning site of the pEZZ18 vector, previously digested with *Eco*RI and *Xba*I (Figure 2b). The ligation mixture was electroporated into XL-1 cells (Stratagene) and the cells plated on to agar containing ampicillin, isopropyl thiogalactoside and 5-bromo-4-chloroindol-3-yl β -D-galactopyranoside ('X-Gal'). White colonies were grown in liquid culture overnight in the presence of ampicillin and the plasmid DNA prepared and analysed by restriction digestion and agarose-gel electrophoresis. The cloned DNA was sequenced in both directions to ensure that it contained the correct DNA sequence.

Site-directed mutagenesis

Cys⁵⁸ in the recombinant pNR-2/pS2 expression plasmid was mutated to the isosteric amino acid serine as described by Kunkel et al. [25]. As well as converting cysteine to serine, the mutagenic primer:

5'GCAGGTCGACTCGATATCAAAATT-
CAGACTCTTCTGG3' (Figure 2a)

introduces an *Eco*RV site 3' of the pNR-2/pS2 stop codon to aid identification of the mutated recombinants. The mutated pNR-2/pS2 DNA was sequenced to confirm that it contained the correct mutation.

Purification of recombinant pNR-2/pS2

The pNR-2/pS2 recombinant plasmid was electroporated into HB101 cells as this strain yields a higher proportion of the fusion protein in the periplasmic space. Cultures (500 ml) were grown overnight to stationary phase at 37 °C in 2-litre flasks. Bacteria were harvested by centrifugation, washed in 20% (w/v) sucrose/1 mM EDTA/0.3 M Tris/HCl, pH 8.0, and resuspended in 60 ml of 0.1 mM MgCl₂ per culture to release proteins contained within the periplasmic space [26]. After a 10 min incubation with agitation at room temperature, the extract was clarified by centrifugation at 10000 g for 10 min at 4 °C. When necessary the extract was re-clarified. The extract was filtered through a 0.45 μ m-mesh filter and then loaded on to a 10 ml IgG-Sepharose (Pharmacia) column in 150 mM NaCl/50 mM Tris/HCl, pH 7.6. The column was washed with 5 mM ammonium acetate, pH 5.0, and the bound material eluted with 0.5 M acetic acid, pH 3.4. The sample was transferred into 100 mM NaCl/1 mM CaCl₂/50 mM Tris/HCl, pH 8.0, by passage over Sephadex G-25 (Pharmacia). The protein concentration was measured by the BCA (bicinchoninic acid) assay (Pierce), adjusted to 1 mg/ml and digested for 24 h at 4 °C with Factor Xa (New England Biolabs) [1:150 (w/w) enzyme/substrate ratio]. Enzyme from this supplier performed consistently well, whereas we have failed to obtain digestion with enzyme obtained from Sigma. The cleaved pNR-2/pS2 was purified by passage of the cleavage mixture over IgG-Sepharose as described above. The mature protein was present in the flow through from the column. The first 20 amino acids of this preparation were sequenced using a Beckman gas-phase sequencer.

PAGE

Denaturing SDS/PAGE was performed essentially as described by Giulian [27] on a 20% separating gel, except that the sample buffer contained β -mercaptoethanol in place of dithiothreitol. Samples were boiled for 2 min prior to loading.

Non-denaturing PAGE was performed as described for the denaturing gels apart from the omission of SDS from all buffers and β -mercaptoethanol from the sample buffer. Samples were incubated overnight at 4 °C in 1 mM EDTA/10 mM Bistris propane, pH 6.5, with or without any additional chemical. They were not heated after addition of the non-denaturing sample buffer electrophoresis.

After electrophoresis, gels were stained in water/methanol/propionic acid [50:50:7 (by vol.)], containing 0.1% Coomassie Blue. They were destained in water/methanol/propionic acid [100:5:7, by vol.).

Western transfer

Proteins were transferred to 0.2 μ m-pore-size nitrocellulose with a semi-dry transfer apparatus (Hofer) according to the manufacturer's instructions for 10 min at 100 mA. The filter was air-dried overnight and then the proteins were fixed in 0.2% glutaraldehyde. The filter was blocked in 3% BSA/PBS and then incubated with the pNR-2/pS2 antisera at a 1:2000 dilution in 3% BSA/PBS for 2 h at 37 °C. It was then incubated with a secondary antibody conjugated to alkaline phosphatase and developed as described previously [15].

Ion-exchange chromatography

pNR-2/pS2 samples (100 μ g) were diluted at least 10-fold into 20 mM 1-methylpiperazine, pH 4.5, and loaded on to a 1 ml Mono Q column (Pharmacia) at 2 ml/min. Elution was with a 0–400 mM NaCl gradient at 1 ml/min. Fractions (1 ml each) were collected and 10–20 μ l aliquots analysed by PAGE and Western transfer.

Gel filtration

Protein (5–10 μ g) was loaded on to a Superdex 75 gel-filtration column (Pharmacia) via a 50 μ l loop. The running buffer was 150 mM NaCl/20 mM 1-methylpiperazine, pH 4.5. The flow rate was 0.5 ml/min. Protein standards were aprotinin (6.5 kDa), RNase A (13.7 kDa), ovalbumin (43 kDa), BSA (67 kDa), and the excluded volume was determined using Blue Dextran (~2000 kDa).

RESULTS AND DISCUSSION

Production of mature recombinant pNR-2/pS2

The N-terminus of mature pNR-2/pS2 was identified by N-terminal sequencing of the protein purified from human gastric juice and from breast-cancer-cell-conditioned medium [3]. A 24-amino-acid signal sequence is removed to leave Glu as the N-terminal amino acid. This is converted into pyroglutamic acid in gastric juice, but not in breast-cancer-cell-conditioned medium.

DNA encoding a cassette consisting of the proteinase-Factor-Xa-recognition signal upstream of cDNA encoding the mature pNR-2/pS2 protein was inserted into the expression vector pEZZ18 as described in the Experimental section (Figure 2). This vector initiates transcription from two strong promoters, *lac* and *spa*, and translation of the resulting RNA gives rise to a fusion protein with the 'S' sequence at its N-terminus to translocate the fusion protein to the periplasmic space. *Escherichia coli* (HB101) were electroporated with the recombinant pNR-2/pS2:pEZZ18

Table 1 Purification of recombinant pNR-2/pS2

Purification step	Protein ... pNR-2/pS2 Cys ⁵⁸		pNR-2/pS2 Ser ⁵⁸	
	Yield (mg)	Purity (%)	Yield (mg)	Purity (%)
IgG-Sepharose	16*	> 90	20	> 90
	~ 4.2†	~ 24	~ 5.3	~ 24
Factor Xa digestion	4.0	~ 22.5	5.0	~ 22.6
Second IgG-Sepharose	3.8	> 90	4.8	> 95
Anion-exchange	3.3	100	4.0	100
Buffer exchange	3.2	100	4.0	100

* Fusion protein.

† Estimated as a proportion of the yield of fusion protein.

Figure 3 Denaturing PAGE of recombinant pNR-2/pS2

A 2 µg portion of fusion protein (F.P.), cleaved fusion protein (Cleav.) and purified pNR-/pS2 (Pur.) were electrophoresed and stained as described in the experimental section, for the pNR-2/pS2 Cys⁵⁸ recombinant (a) and the pNR-2/pS2 Ser⁵⁸ recombinant (b). The positions of the molecular-mass (*M*) markers are shown on the left-hand side, and the positions of the fusion protein, cleaved ZZ protein (see the text) and mature pNR-2/pS2 are shown on the right-hand side.

expression plasmid. Analysis of periplasmic extracts and media established that the majority of the fusion protein (> 90%) was present in the periplasmic space. This provides substantial purification of recombinant protein from other bacterial proteins, but avoids processing large volumes of medium. In addition the recombinant protein is secreted into an oxidizing environment which is favourable for the folding of cysteine-rich secreted proteins.

The two Z domains are mutated variants of the E domain of Protein A (Figure 2) and allow purification of the fusion protein away from other bacterial proteins present in the periplasmic space by affinity chromatography on IgG-Sepharose [28]. Typically, 16 mg of fusion protein/litre of culture were purified on IgG-Sepharose (Table 1). Analysis of this material on SDS/polyacrylamide gels indicated that the major component was a protein of 24–28 kDa apparent molecular mass, which is the predicted size of the ZZ-pNR-2/pS2 fusion protein (Figure 3a). The eluate from the IgG-Sepharose column was transferred into Factor Xa cleavage buffer by gel filtration, and cleaved for 24 h at 4 °C at an enzyme/protein ratio of 1:150 (w/w). Electrophoresis indicated that cleavage was almost complete

yielding two major products; the ZZ protein of ~ 15 kDa and mature pNR-2/pS2 of 6.5 kDa (Figure 3a). The uncleaved fusion protein, the ZZ protein and the bacterial proteins which bind to IgG-Sepharose were then removed by passage over IgG-Sepharose. The flow-through contained pNR-2/pS2 protein, which was more than 90% pure with a yield of ~ 3.8 mg/litre of culture (Table 1). Protein sequencing confirmed glutamic acid as the N-terminal residue and that the first 20 amino acids of the recombinant protein were identical with those of the published sequence of mature pNR-2/pS2. There was no evidence of degradation of either the fusion protein or mature pNR-2/pS2 at any stage of the purification procedure (Figure 3).

Anion-exchange chromatography

Although electrophoresis of β-mercaptoethanol-treated protein on denaturing polyacrylamide gels indicated that the protein was very pure and the protein-sequencing data were consistent with this level of purity, the recombinant pNR-2/pS2 would be expected to be contaminated with Factor Xa and may comprise differently folded protein molecules. Before the recombinant protein was used for structural analysis, we attempted to purify it further by anion-exchange chromatography on Mono Q.

The pNR-2/pS2 protein secreted from MCF-7 breast-cancer cells is eluted from the Mono Q column between 130 and 150 mM NaCl at pH 4.5 (results not shown). This demonstrates that the protein has a low pI and is consistent with the high proportion of acidic residues in the pNR-2/pS2 protein. Recombinant pNR-2/pS2 was bound to Mono Q in *N*-methylpiperazine, pH 4.5, and eluted with a linear salt gradient (0–400 mM). The majority of the protein loaded on to the column bound to the matrix, and the elution profile is shown in Figure 4(a). Surprisingly the protein was eluted as multiple peaks over a wide range of salt concentrations. The different fractions from the ion-exchange column were electrophoresed on a denaturing polyacrylamide gel, transferred to nitrocellulose and allowed to react with a specific pNR-2/pS2 polyclonal antibody (inset to Figure 4a). This confirmed that all the peaks contained pNR-2/pS2 protein, but showed that it was present as multiple forms with varying affinities for the anion-exchange resin.

The observation that the multiple peaks all migrated as a single band of 6.5 kDa following reduction with β-mercaptoethanol suggested that the heterogeneity resulted from disulphide-bond formation. There could be several conformations of the protein differing in the arrangement of their disulphide bonds. In addition, intermolecular disulphide bonds between pNR-2/pS2 molecules, or other proteins, could give rise to a number of stable dimer conformations. To investigate these possibilities, recombinant pNR-2/pS2 was treated with thiol agents prior to chromatography on mono Q.

Treatment of recombinant protein with 0.1 mM dithiothreitol (DTT) in the presence of 1 mM EDTA reduced the heterogeneity, producing a major peak which was eluted at 140 mM NaCl (results not shown). Treatment with 1 mM DTT further reduced the heterogeneity and most of the protein was eluted in a single peak at 140 mM NaCl (Figure 4b). Recombinant pNR-2/pS2 was also treated with 50 mM cysteine in the presence of 1 mM EDTA and analysed as described above (Figure 4c). Cysteine also converted the recombinant pNR-2/pS2 protein into a form which was eluted as a single peak at about 140 mM NaCl.

Generation of pNR-2/pS2 Ser⁵⁸

The observation that thiol-group-containing reagents reduced the heterogeneity of forms of recombinant pNR-2/pS2 that were eluted from Mono Q demonstrated that this heterogeneity

Figure 4 Ion-exchange chromatography of pNR-2/pS2 Cys⁵⁸

A 100 μ g portion of pNR-2/pS2 Cys⁵⁸ was loaded on to a 1 ml mono Q column and eluted with a salt gradient as described in the Experimental section. Fractions (1 ml each) were collected and a 20 μ l sample was fractionated on a denaturing polyacrylamide gel, transferred to nitrocellulose and allowed to react with pNR-2/pS2 antisera as described in the Experimental section (insets). The sizes of the molecular-mass (M) markers are shown on the left-hand side. All samples of pNR-2/pS2 Cys⁵⁸ were incubated overnight in 1 mM EDTA/10 mM Bistris propane, pH 6.5; (a) with no further additions; (b) with 1 mM DTT; (c) with 50 mM cysteine.

Figure 5 Western transfer of pNR-2/pS2 proteins

Recombinant IGF-I, pNR-2/pS2 Cys⁵⁸ (Cys⁵⁸), pNR-2/pS2 Ser⁵⁸ (Ser⁵⁸) and pNR-2/pS2 from MCF-7 cells (MCF-7) were electrophoresed on a denaturing polyacrylamide gel, transferred to nitrocellulose and allowed to react with antisera against pNR-2/pS2 as described in the Experimental section. The sizes of the molecular-mass markers (M) are shown on the left-hand side.

resulted from the presence of a variety of molecular forms varying in their pattern of inter- or intra-molecular disulphide bonds. Mature pNR-2/pS2 contains an odd number of cysteine residues, and the heterogeneity could be generated by interactions of the unpaired cysteine residue (Figure 1).

pNR-2/pS2 contains six cysteine residues within the region of sequence identity with the prototype trefoil PSP. This region

forms a compact globular domain in PSP [14]. Although the pattern of disulphide bonds in pNR-2/pS2 is not known, the position of the seventh cysteine residue at the C-terminus of the protein suggests that it lies outside the globular domain and may be a good candidate for being the unpaired cysteine in correctly folded pNR-2/pS2 (Figure 1). To investigate this possibility and to assist characterization of the different forms of recombinant pNR-2/pS2, pNR-2/pS2 Ser⁵⁸ was made in which Cys⁵⁸ is mutated to the isosteric amino acid serine and the six cysteine residues lying within the trefoil domain are retained.

Cys⁵⁸ of the pNR-2/pS2 expression plasmid was altered to serine by site-directed mutagenesis as described in the Experimental section. Mutated recombinant protein was prepared as described for pNR-2/pS2 Cys⁵⁸. In general the yield and purity of pNR-2/pS2 Ser⁵⁸ was slightly higher than for the unmutated protein, with 4 mg of recombinant recovered from 1 litre of bacterial culture (Table 1). Figure 5 shows a Western transfer of a gel in which pNR-2/pS2 protein secreted by MCF-7 cells was co-electrophoresed with pNR-2/pS2 Cys⁵⁸ and pNR-2/pS2 Ser⁵⁸ and insulin-like growth factor I (IGF-I) recombinant proteins. The preimmune serum did not react with any of the four proteins (results not shown). The antisera raised against pNR-2/pS2 did not react with IGF-I, another small cysteine-rich protein of 7 kDa, but reacted strongly with the three pNR-2/pS2 proteins, all of which migrate with the same apparent molecular mass.

pNR-2/pS2 Ser⁵⁸ was then chromatographed on a Mono Q column as described above. The elution profile (Figure 6) was considerably more homogeneous than the profile obtained with pNR-2/pS2 Cys⁵⁸ and comprised a major protein peak ($\sim 80\%$) that was eluted at approx. 140 mM NaCl. This strongly suggests that the seventh cysteine residue is largely responsible for

Figure 6 Ion-exchange chromatography of pNR-2/pS2 Ser⁵⁸

A 100 μ g portion of pNR-2/pS2 Ser⁵⁸ was loaded on to a Mono Q column and eluted with a salt gradient as described in the Experimental section. Fractions (1 ml each) were collected, and 20 μ l samples were fractionated on a denaturing polyacrylamide gel, transferred to nitrocellulose and allowed to react with pNR-2/pS2 antisera as described in the Experimental section (inset). The sizes of the molecular-mass (*M*) markers are shown on the left-hand side.

generating the heterogeneity in the elution profile of pNR-2/pS2 Cys⁵⁸ from mono Q.

Comparison of the recombinant pNR-2/pS2 proteins under non-denaturing conditions

Recombinant pNR-2/pS2 Cys⁵⁸ and pNR-2/pS2 Ser⁵⁸ were compared by non-denaturing gel electrophoresis, which separates proteins on the basis of a combination of size and charge. The position of pNR-2/pS2 protein purified from conditioned medium of MCF-7 cells was determined by Western transfer and is indicated by an arrow in Figures 7(a) and 7(b). pNR-2/pS2 Cys⁵⁸ migrated as a variety of forms and did not contain a discrete band in the same position as pNR-2/pS2 produced by MCF-7 cells. In contrast, for pNR-2/pS2 Ser⁵⁸, the major band co-migrated with MCF-7 cell pNR-2/pS2. This reinforces the conclusion from the anion-exchange chromatography that Cys⁵⁸ is responsible for heterogeneity in the recombinant protein. When pNR-2/pS2 Cys⁵⁸ was treated with increasing concentrations of DTT (Figure 7a), two major forms were generated. The first co-migrated with MCF-7 cell pNR-2/pS2, while the second migrated somewhat more slowly and predominated after treatment with higher DTT concentrations. These data suggest that the single peak obtained by chromatography on Mono Q (Figure 4b) actually consists of more than one form. The observation that higher concentrations of DTT tend to generate the more slowly migrating band suggests that this represents a denatured form in which the disulphide bonds within the trefoil domain are

Figure 7 Non-denaturing gel electrophoresis of pNR-2/pS2 Cys⁵⁸ and pNR-2/pS2 Ser⁵⁸

After purification on the second IgG-Sepharose column, 10 μ g of recombinant protein was incubated overnight in 20 μ l of 1 mM EDTA, 10 mM Bis-tris propane, pH 6.5, with no further additions or with, in addition, 0.1 mM DTT, 1 mM DTT, 10 mM DTT, 50 mM cysteine ('Cys'), 10 mM GSH, 10 mM GSSG or 5 mM GSSG and 5 mM GSH. A 10 μ g portion of treated pNR-2/pS2 Cys⁵⁸ (a) and pNR-2/pS2 Ser⁵⁸ (b) were electrophoresed on non-denaturing polyacrylamide gels and stained as described in the Experimental section. The positions at which pNR-2/pS2 purified from MCF-7 cells migrates is indicated on the right-hand side by an arrow.

reduced. In contrast with the complex concentration-dependent effects of DTT, treatment of pNR-2/pS2 Cys⁵⁸ with cysteine produced one predominant form which co-migrated with MCF-7 pNR-2/pS2. Our interpretation of this data is that DTT (depending on the concentration used) reduces the intramolecular and the intermolecular disulphide bonds, and that the reduced cysteine residues are then unable to re-form disulphide bonds, producing a denatured protein. Cysteine probably also reduces the intermolecular disulphide bonds and inhibits them from re-forming. However, it may be that, after cysteine reduces intramolecular disulphide bridges, it favours the formation of the 'correct' intramolecular bonds that produce recombinant pNR-2/pS2 folded in a form that co-migrates with pNR-2/pS2 from MCF-7 cells.

Wrongly folded proteins are often treated with a mixture of GSH and GSSG glutathione to induce reshuffling of their disulphide bridges and therefore generation of the correctly folded form of the protein [29,30]. Unfortunately, treatment of pNR-2/pS2 Cys⁵⁸ with various combinations of GSH and GSSG (Figure 7a) did not produce a form that co-migrated with pNR-2/pS2 from MCF-7 cells. GSH gave rise to a more slowly migrating band, while GSSG alone had little effect, and a mixture of GSH and GSSG produced some material which migrated ahead of MCF-7 pNR-2/pS2 and some material which migrated more slowly. Importantly, these experiments demonstrate that cysteine could be used to produce a predominant form of pNR-2/pS2 Cys⁵⁸, which co-migrated with pNR-2/pS2 secreted by MCF-7 cells.

Figure 7(b) shows the effect of treating pNR-2/pS2 Ser⁵⁸ with the same thiol reagents as were used in the experiment shown in Figure 7(a). Unlike pNR-2/pS2 Cys⁵⁸, untreated pNR-2/pS2 Ser⁵⁸ shows a prominent band which co-migrates with pNR-2/pS2 secreted by MCF-7 cells. We suggest that this reflects in part the inability of pNR-2/pS2 Ser⁵⁸ to form intermolecular disulphide bonds. It is also possible that a higher proportion of recombinant pNR-2/pS2 Ser⁵⁸ molecules may be correctly folded, perhaps because the folding is not compromised by the presence of the extra trefoil-domain cysteine. Treatment of pNR-2/pS2 Ser⁵⁸ with increasing concentrations of DTT produced a more slowly migrating band in the same place as was found for DTT-treated pNR-2/pS2 Cys⁵⁸. This is consistent with the slower-migrating form representing a denatured form in which the intramolecular disulphide bonds are reduced. Interestingly, treatment of pNR-2/pS2 Ser⁵⁸ with cysteine appeared to increase the predominance of the major band co-migrating with MCF-7 pNR-2/pS2. As pNR-2/pS2 Ser⁵⁸ is unlikely to make intermolecular disulphide bonds, this suggests that a small proportion of the pNR-2/pS2 Ser⁵⁸ is incorrectly folded and that cysteine enables shuffling of intramolecular disulphide bonds to give the correctly folded form. Various combinations of GSH and GSSG had little effect on pNR-2/pS2 Ser⁵⁸. That even 10 mM GSH was ineffective indicates that pNR-2/pS2 Ser⁵⁸ has a remarkably stable structure. We suggest that, under these conditions, GSH is unable to reduce the intramolecular disulphide bonds of correctly folded pNR-2/pS2 protein and that recombinant pNR-2/pS2 Ser⁵⁸ protein is predominantly of this form.

In addition to analysing recombinant pNR-2/pS2 on non-denaturing polyacrylamide gels, the molecular masses of the recombinant proteins were estimated by gel filtration. Cysteine-treated pNR-2/pS2 Cys⁵⁸ and untreated pNR-2/pS2 Ser⁵⁸ were purified by ion-exchange chromatography at pH 4.5, and the fractions that were eluted between 130 and 150 mM NaCl were pooled and applied to a Superdex 75 column that separates globular proteins in the molecular-mass range 5–70 kDa (Pharmacia). Both cysteine-treated pNR-2/pS2 Cys⁵⁸ and pNR-2/pS2 Ser⁵⁸ were eluted as single symmetrical peaks at 14.4 ml (results not shown). Under identical conditions, pNR-2/pS2 from MCF-7-cell-conditioned medium was also eluted between 14 and 15 ml. After calibration of the column with globular proteins of known molecular mass, the apparent molecular masses of the pNR-2/pS2 proteins were found to be ~ 12.5 kDa. Neither alteration of the pH or salt concentration of the running buffer nor inclusion of 10% (v/v) ethylene glycol in the running buffer reduced the apparent molecular masses obtained. The anomalous molecular mass suggests that the pNR-2/pS2 protein either has a non-globular structure or is dimeric. The observation that the apparent molecular mass of pNR-2/pS2 Ser⁵⁸ (which would not be able to form intermolecular disulphide bonds) is the same as the Cys⁵⁸ form of the peptide suggests that the former explanation is correct. The conformation of the five-amino-acid N-terminal and the 13-amino-acid C-terminal extra-trefoil domains (Figure 1) could be responsible for a non-globular structure with an anomalous Stokes radius. However, it is possible that the peptide has other dimer-forming surfaces, and current studies are aimed at establishing the subunit structure of this peptide.

Conclusion

This is the first description of the production of a recombinant mature single-domain trefoil peptide. The methodology is rela-

tively straightforward and yields extremely pure recombinant pNR-2/pS2 in sufficient quantities for both structural and biological studies. As evaluated by their interaction with an ion-exchange matrix, mobility on non-denaturing polyacrylamide gels and fractionation on a gel-filtration matrix, cysteine-treated recombinant pNR-2/pS2 Cys⁵⁸ and mutated recombinant pNR-2/pS2 Ser⁵⁸ are indistinguishable from each other and from pNR-2/pS2 synthesized by, and secreted from, MCF-7 cells.

M.P.C. thanks the Medical Research Council for a training scholarship. We are grateful to the pupils of Newcastle upon Tyne Church High School for generously contributing towards the purchase of an FPLC system. We thank Dr. D. R. Hodgson for his advice and helpful discussions, and Mrs. R. Brown and Mrs. M. Earnshaw for technical support.

REFERENCES

- Masiakowski, P., Breathnach, R., Bloch, J., Gannon, F., Krust, A. and Chambon, P. (1982) *Nucleic Acids Res.* **24**, 7895–7903
- May, F. E. B. and Westley, B. R. (1986) *Cancer Res.* **46**, 6034–6040
- Rio, M.-C., Lepage, P., Diemunsch, P., Roitsch, C. and Chambon, P. (1988) *C. R. Acad. Sci. Ser. 3* **307**, 825–831
- Westley, B., May, F. E. B., Brown, A. M. C., Krust, A., Chambon, P., Lippman, M. E. and Rochefort, H. (1984) *J. Biol. Chem.* **259**, 10030–10035
- May, F. E. B. and Westley, B. R. (1987) *J. Biol. Chem.* **262**, 15894–15899
- Nunez, A.-M., Berry, M., Imler, J.-L. and Chambon, P. (1989) *EMBO J.* **8**, 823–829
- Foekens, J. A., Rio, M.-C., Sequin, P., Putten, W. L. J. van., Fauque, J., Nap, M., Klijn, J. G. M. and Chambon, P. (1990) *Cancer Res.* **50**, 3832–3837
- Henry, J. A., Piggott, N. H., Mallick, U. K., Nicholson, S., Farndon, J. R., Westley, B. R. and May, F. E. B. (1991) *Br. J. Cancer* **63**, 615–622
- Thor, A. D., Koerner, F. C., Edgerton, S. M., Wood, W. C., Stracher, M. A. and Schwartz, L. H. (1992) *Breast Cancer Res. Treat.* **21**, 11–119
- Schwartz, L. H., Koerner, F. C., Edgerton, S. M., Sawicka, J. M., Rio, M. C., Bellocq, J. P., Chambon, P. and Thor, A. D. (1991) *Cancer Res.* **51**, 624–628
- Baker, M. E. (1988) *Biochem. J.* **253**, 307–311
- Thim, L. (1989) *FEBS Lett.* **250**, 85–90
- De, A., Brown, D. G., Gorman, M. A., Carr, M., Sanderson, M. R. and Freemont, P. S. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 1084–1088
- Carr, M. D., Bauer, C. J., Greadwell, M. J. and Feeney, J. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 2206–2210
- Piggott, N. H., Henry, J. A., May, F. E. B. and Westley, B. R. (1991) *J. Pathol.* **163**, 95–104
- Rio, M. C., Bellocq, J.-P., Daniel, J. Y., Tamasetto, C., Lathe, R., Chenard, M. P., Batzenschlager, A. and Chambon, P. (1988) *Science* **241**, 705–707
- Tamasetto, C., Rio, M.-C., Gautier, C., Wolf, C., Hareuveni, M., Chambon, P. and Lathe, R. (1990) *EMBO J.* **9**, 407–414
- Podolsky, D. K., Lynch-Devaney, K., Stow, J. L., Oates, P., Murgue, B., DeBeaumont, M., Sands, B. E. and Mahida, Y. R. (1993) *J. Biol. Chem.* **268**, 6694–6702
- Hauser, F., Poulson, R., Chinery, R., Rogers, L. A., Hanby, A. M., Wright, N. A. and Hoffmann, W. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6961–6965
- Wright, N. A., Poulson, R., Stamp, G. W. H., Hall, P. A., Jeffrey, R. E., Longcroft, J. M., Rio, M. C., Tamasetto, C. and Chambon, P. (1990) *J. Pathol.* **162**, 279–284
- Wright, N. A., Poulson, R., Stamp, G., Van Noorden, S., Sarraf, C., Elia, G., Ahnen, D., Jeffrey, R., Longcroft, J., Pike, C., Rio, M. C. and Chambon, P. (1993) *Gastroenterology* **104**, 12–20
- Poulson, R. and Wright, N. A. (1993) *Am. J. Physiol.* **265**, 205–213
- Hoffman, W. (1988) *J. Biol. Chem.* **263**, 7686–7690
- May, F. E. B. and Westley, B. R. (1988) *J. Biol. Chem.* **263**, 12901–12908
- Kunkel, T. A., Roberts, J. D. and Zakour, R. A. (1987) *Methods Enzymol.* **154**, 367–382
- Nossal, N. G. and Heppel, L. A. (1966) *J. Biol. Chem.* **241**, 3055–3062
- Giulian, G. G., Shanahan, M. F., Graham, J. M. and Moss, R. L. (1985) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **44**, 686–699
- Nilsson, B., Forsberg, G. and Hartmanis, M. (1991) *Methods Enzymol.* **198**, 3–16
- Spear and Sliwkowski (1991) *Techniques in Protein Chemistry 2*, pp. 233–240, Academic Press, New York
- Gilbert, H. F. (1994) in *Mechanisms of Protein Folding* (Pain, R. H. ed.), pp. 104–136, Oxford University Press, Oxford