Biochem. J. (1992) 284, 795–802 (Printed in Great Britain)

Purification, characterization and cDNA cloning of human lung surfactant protein D

Jinhua LU, Antony C. WILLIS and Kenneth B. M. REID*
MRC Immunochemistry Unit, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, U.K.

Human pulmonary surfactant protein D (SP-D) was identified in lung lavage by its similarity to rat SP-D in both its molecular mass and its Ca^{2+}-dependent-binding affinity for maltose [Persson, Chang & Crouch (1990) J. Biol. Chem. **265**, 5755–5760]. For structural studies, human SP-D was isolated from amniotic fluid by affinity chromatography on maltose–Sepharose followed by f.p.l.c. on Superose 6, which showed it to have a molecular mass of approx. 620 kDa in non-dissociating conditions. On SDS/PAGE the human SP-D behaved as a single band of 150 kDa or 43 kDa in non-reducing or reducing conditions respectively. The presence of a high concentration of glycine (22%), hydroxyproline and hydroxylysine in the amino acid composition of human SP-D indicated that it contained collagen-like structure. Collagenase digestion yielded a 20 kDa collagenase-resistant globular fragment which retained affinity for maltose. Use of maltosyl-BSA as a neoglycoprotein ligand in a solid-phase binding assay showed that human SP-D has a similar carbohydrate-binding specificity to rat SP-D, but a clearly distinct specificity from that of other lectins, such as conglutinin, for a range of simple saccharides. Amino acid sequence analysis established the presence of collagen-like Gly-Xaa-Yaa triplets in human SP-D and also provided sequence data from the globular region of the molecule which was used in the synthesis of oligonucleotide probes. Screening of a human lung cDNA library with the oligonucleotide probes, and also with rabbit anti-(human SP-D), allowed the isolation of two cDNA clones which overlap to give the full coding sequence of human SP-D. The derived amino acid sequence indicates that the mature human SP-D polypeptide chain is 355 residues long, having a short non-collagen-like *N*-terminal section of 25 residues, followed by a collagen-like region of 177 residues and a *C*-terminal C-type lectin domain of 153 residues. Comparison of the human SP-D and bovine serum conglutinin amino acid sequences indicated that they showed 66% identity despite their marked differences in carbohydrate specificity.

INTRODUCTION

Pulmonary surfactant is a complex of lipids and proteins which forms the interface between the aqueous film on the surface of alveolar epithelium and the air in the alveolar space. The surfactant serves to lower the surface tension at the air/fluid interface in the peripheral air spaces after expiration of air from the lungs. At present, four proteins have been identified which are lung-specific and surfactant-associated. Three of these proteins have been well characterized; they are termed surfactant proteins A (SP-A) (King *et al.*, 1973), B (SP-B) (Glasser *et al.*, 1987; Hawgood *et al.*, 1987) and C (SP-C) (Warr *et al.*, 1987; Fisher *et al.*, 1989). SP-B and SP-C are small hydrophobic proteins of 14 kDa and 6 kDa respectively, which associate strongly with the surfactant phospholipids and play an important role in the mechanical structure of the surfactant in facilitating the fast spreading and stability of the surfactant film (Suzuki *et al.*, 1986; Takahashi & Fujiwara, 1986; Curstedt *et al.*, 1987). SP-A is a major surfactant protein and it is known to be composed of 18 polypeptide chains, each containing a seven-amino-acid-residue *N*-terminal section followed by a region of 23 Gly-Xaa-Yaa collagen-like repeating triplets and a *C*-terminal 148-residue globular domain (White *et al.*, 1985; Floros *et al.*, 1986). It binds to phospholipid and aggregates phospholipids in a Ca^{2+}-dependent manner which may promote the ordering of the tubular myelin within alveolar type II cells (King *et al.*, 1986; Hawgood *et al.*, 1987; Suzuki *et al.*, 1989). Studies *in vitro* have suggested that SP-A enhances the uptake and inhibits the secretion of surfactant phospholipids and therefore regulates the surfactant turnover (Dobbs *et al.*, 1987; Rice *et al.*, 1987; Kuroki *et al.*, 1988; Ryan *et al.*, 1989). Electron-microscopy studies have revealed that its overall structure is very similar to that of C1q of the complement system (Voss *et al.*, 1988). However, it differs from C1q in that its *C*-terminal globular domain contains a C-type lectin framework (White *et al.*, 1985; Floros *et al.*, 1986; Drickamer, 1988). Owing to the fact that SP-A in a purified form tends to aggregate under physiological conditions, it has been difficult to test for biological activity. It does not appear to show complement-dependent haemolytic activity (Tenner *et al.*, 1989) and its bivalent cation-dependent mannose specificity is only seen at low ionic strength (Haagsman *et al.*, 1987). However, it has been found to enhance FcR-mediated and CR1-mediated phagocytosis and to inhibit complement activation in a C1q-dependent haemolytic assay (Tenner *et al.*, 1989) and in a C4 consumption assay (J. Lu and K. B. M. Reid, unpublished work).

Pulmonary surfactant is synthesized mainly in type II alveolar cells and stored within the cell in the lamellar bodies (Macklin, 1954; Baritusso *et al.*, 1981; Young *et al.*, 1981), with its protein components detected in the clara cells of some species (Walker *et al.*, 1986; Phelps & Floros, 1988). Recently, a new protein has been identified and purified from primary cultures of rat type II alveolar cells and shown, like SP-A, to contain collagen-like structure (Persson *et al.*, 1988). This protein, initially named CP4 and later SP-D, was found in rat surfactant (Persson *et al.*, 1989; Ogasawara *et al.*, 1991). It was reported to have a molecular mass of approx. 200 kDa on SDS/PAGE in non-reducing conditions and appeared to be a disulphide-bond-linked trimer of the 43 kDa

Abbreviations used: SP-A, SP-B, SP-C and SP-D, surfactant proteins A, B, C and D; MBP, mannan-binding protein; 1 × SSC, 0.15 M-NaCl/0.015 M-sodium citrate; 1 × Denhardt's solution, 0.02% (w/v) Ficoll, 0.02% (w/v) polyvinylpyrrolidone and 0.02% (w/v) BSA.

* To whom correspondence should be addressed.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X65018.

polypeptide chain when examined on SDS/PAGE under reducing conditions. On gel filtration in non-dissociating conditions, rat SP-D appeared larger than the 450 kDa SP-A molecule (Persson *et al.*, 1989). Persson *et al.* (1990) demonstrated the high Ca^{2+}-dependent affinity of rat SP-D for maltose and showed that it also bound, to a lesser extent, a broad spectrum of other saccharide ligands. These characteristics make it a member of a family of collagenous lectins (Thiel & Reid, 1989).

Lung surfactant proteins were first detected in amniotic fluid by King *et al.* (1975). The appearance of surfactant proteins starts at approx. 30-week gestation to a maximum at about 37 weeks. In this paper, we have made use of amniotic fluid as a source of human SP-D as an alternative to the lung lavage. This has allowed the characterization of the human equivalent of rat SP-D at the protein level and the molecular cloning of the human SP-D, thus providing, for the first time, the entire derived amino acid sequence of the human SP-D molecule. An unexpected finding was the strong similarity in amino acid sequence between human SP-D and conglutinin which is a lectin containing collagen-like structure found in bovine serum (Davis & Lachmann, 1984). Although human SP-D was shown to be structurally related to conglutinin, its carbohydrate-binding specificity was found to be quite different from that of conglutinin.

MATERIALS AND METHODS

Purification of human SP-D from amniotic fluid

A pool of human amniotic fluid (approx. 1 litre) from normal full-term births, which had been stored at −20 °C, was thawed and clarified by passage through four layers of muslin. The filtrate was centrifuged at 400 ***g*** for 10 min, the pellet discarded and the supernatant then centrifuged at 50000 ***g*** for 30 min. The pellet from this high-speed spin of the crude surfactant contains most of the SP-A present in the amniotic fluid. The supernatant, which contains the SP-D, was made 2 mM with respect to $CaCl_2$ and mixed with maltose–Sepharose (15 ml) which had been equilibrated with TBS/NTCa^{2+} [50 mM-Tris/HCl, 150 mM-NaCl, 0.05% (w/v) NaN_3, 0.05% (v/v) Tween 20 and 2 mM-$CaCl_2$, pH 7.4]. After being mixed overnight at 4 °C, the gel was packed in a column (30 ml) and washed with TBS/NTCa^{2+} (approx. 400 ml). The column was then eluted with TBS/NTE [50 mM-Tris/HCl, 150 mM-NaCl, 0.05% (w/v) NaN_3, 0.05% (v/v) Tween 20 and 10 mM-EDTA, pH 7.4]. The fractions containing human SP-D were pooled and recalcified so that the pool was 2.5 mM with respect to $CaCl_2$, and then the pH was adjusted to 7.4 with 1 M-NaOH. The pool was mixed with equilibrated maltose–Sepharose (4 ml) overnight at 4 °C. The gel was packed in a column and washed as above with TBS/NTCa^{2+} and then eluted with TBS/NTCa^{2+} containing 100 mM-maltose. The protein fractions containing human SP-D were pooled and concentrated (to approx. 0.5 ml) with a Centriprep 30 concentrator (Amicon Corp., Danvers, MA, U.S.A.), and then further purified by f.p.l.c. on a Superose 6 column equilibrated and eluted with TBS/NTE.

Purification of bovine SP-D from lung lavage

Bovine SP-D was isolated from lung lavage by essentially the same procedure as was used to purify the human SP-D from amniotic fluid.

Electrophoresis

SDS/PAGE was carried out as described by Laemmli (1970) or by use of a gradient gel system. Samples were reduced by heating at 100 °C for 5 min in 53 mM-dithiothreitol/4 M-urea/1% (w/v) SDS/0.1 M-Tris (pH 8.0) and then alkylated by addition of iodoacetamide to a concentration of 130 mM. Non-reduced samples were heated at 100 °C for 2 min in 130 mM-iodoacetamide/4 M-urea/1% (w/v) SDS/0.1 M-Tris, pH 8.0. Protein bands were detected by a Ponceau-S dye (for amino acid analysis and sequencing studies), by Coomassie Blue or by the silver staining method.

Amino acid analysis and amino acid sequencing

Samples for amino acid analysis and amino acid sequencing were submitted to SDS/PAGE and then electroblotted on to poly(vinylidene difluoride) membranes (Immobilon P; Millipore, Watford, Herts, U.K.) before detection with Ponceau-S dye as described by Crawford *et al.* (1990). Amino acid analysis was performed using the Waters Pico-Tag system and protein sequencing was carried out on an Applied Biosystems 470A/120A sequencer as described by Crawford *et al.* (1990).

Preparation of anti-(human SP-D) serum

Rabbits were injected with human SP-D (10 μg in 0.75 ml of TBS) purified from amniotic fluid, or with human SP-D bound to maltose–Sepharose (10 μg on 30 μl packed volume) in TBS/Ca^{2+} (50 mM-Tris/HCl, 150 mM-NaCl, 2 mM-$CaCl_2$, pH 7.4), after being mixed with an equal volume of Freund's complete adjuvant. The rabbits were boosted with the same quantity of the respective antigens 4 weeks later. The antisera were adsorbed with possible contaminants in the human SP-D preparation by passing through a column made of immobilized human SP-D-depleted amniotic fluid on Sepharose (the amniotic fluid used has been passed through maltose–Sepharose in TBS/NTCa^{2+}). The antiserum prepared from human SP-D bound to maltose–Sepharose was further adsorbed with maltose–Sepharose. The specificity of the antisera was checked by Western blotting.

Collagenase digestion

Human SP-D (approx. 1 μg) was incubated with collagenase (0.25 unit) in 25 mM-Tris/10 mM-$CaCl_2$, pH 7.4, for 24 h at 37 °C. The sample was reduced with 53 mM-dithiothreitol and alkylated with 130 mM-iodoacetamide before being examined by SDS/PAGE on a 5–15% (w/v) gradient gel. The gel was monitored by silver staining. For binding studies, human SP-D (10 μg in 200 μl) was digested with collagenase (2.5 units) for 24 h at 37 °C and then mixed with maltose–Sepharose (20 μl packed volume), equilibrated with TBS/NTCa^{2+}, for 1 h at room temperature. The mixture was centrifuged and the beads were washed twice with TBS/NTCa^{2+} (1.0 ml per wash) and then eluted by mixing with TBS/NTE (200 μl) for 30 min, at room temperature. Each of the digested, adsorbed or eluted samples (20 μl) was examined by SDS/PAGE on 5–15% (w/v) gradient gels.

Solid-phase carbohydrate-binding assay

Maltosyl-BSA was prepared by dissolving BSA (68 mg), maltose (100 mg) and sodium cyanoborohydride (100 mg) in 0.2 M-potassium phosphate, pH 8.5 (5 ml) and incubating at 37 °C for 4 days (Schwarts & Gray, 1977). The mixture was dialysed against TBS/N [50 mM-Tris/HCl, 150 mM-NaCl, 0.05% (w/v) NaN_3, pH 7.4] and stored at 4 °C. For the carbohydrate-binding assay (Persson *et al.*, 1990), microtitre plates (Nunc-Immuno, Gibco, Grand Island, NY, U.S.A.) were coated with maltosyl-BSA or mannan (10 μg/well in 100 μl of 15 mM-Na_2CO_3/35 mM-$NaHCO_3$, pH 9.6, buffer) by incubation at room temperature overnight or at 37 °C for 1 h. All the following steps were carried out at room temperature unless otherwise stated. The microtitre plates were washed with TBS/NT [50 mM-Tris/HCl, 150 mM-NaCl, 0.05% (w/v) NaN_3, 0.05% (v/v) Tween 20, pH 7.4], and then non-specific sites blocked by BSA by incubation with TBS/NTBSA (1 mg of BSA/ml of TBS/NT) for

1–2 h. The plates were then washed with TBS/NTCa^{2+} and incubated with dilutions of the lectins [human SP-D, human mannan-binding protein (MBP) or bovine conglutinin] alone, or lectins together with dilutions of different saccharides. After 2 h, the plates were washed with TBS/NTCa^{2+} and incubated with 0.125% (w/v) glutaraldehyde in phosphate-buffered saline for 15 min and the free binding sites were blocked with 1 M-ethanolamine for 1 h. The plates were washed with TBS/NT and the bound human SP-D was detected by incubation for 2 h with rabbit antisera [1:200 dilution in TBS/NT of anti-(human SP-D), anti-(human MBP) or anti-(bovine conglutinin)] and then for 1 h with a 1:1000 dilution of goat anti-rabbit IgG–alkaline phosphatase conjugate (Sigma). The plates were developed with phosphatase substrate (*p*-nitrophenyl phosphate, disodium, hexahydrate, 1 mg/ml) in 50 mM-Tris/5 mM-$CaCl_2$/5 mM-$MgCl_2$/0.05% (w/v) NaN_3, pH 9.5, and read on a Titertek Multiskan Plus MKII plate reader (Flow Labs Ltd., Rickmansworth, U.K.).

Isolation of cDNA clones with anti-(human SP-D) antibody

The method used was essentially as described by Sambrook *et al.* (1989). Briefly, approx. 2×10^5 plaques of a human lung λgt11 cDNA library (Clontech Labs Inc., Palo Alto, CA, U.S.A.) was plated out with *Escherichia coli* Y1090. The plates were incubated for 4 h at 42 °C, then isopropylthio-β-D-galactoside-treated filters were applied to the plates and the plates were incubated at 37 °C for a further 4 h. After the first set of filters had been lifted, duplicate filters were applied and incubated overnight. The filters were then washed in TNT [10 mM-Tris, 150 mM-NaCl and 0.05% (v/v) Tween 20, pH 8.0] for 30 min. The free binding sites on the filters were blocked by 30 min incubation in a blocking buffer containing 20% (v/v) fetal calf serum in TNT. A 1:100 dilution of the rabbit anti-(human SP-D) serum in the blocking buffer was incubated with the filters for 3 h. After the filters had been washed, a 1:1000 dilution of a goat anti-rabbit IgG–alkaline phosphatase conjugate was used to detect the bound rabbit IgG. The filters were washed again before being developed with 50 μg of Nitro Blue Tetrazolium/ml and 10 μg of 5-bromo-4-chloro-3-indolyl phosphate/ml (*p*-toluidine salt; Sigma) in the substrate buffer composed of 100 mM-Tris, 100 mM-NaCl and 5 mM-$MgCl_2$, pH 9.5.

Screening of the library with oligonucleotides

A mixture of oligonucleotides (23 bases long) was synthesized on the basis of a portion of the protein sequence obtained for human SP-D (QQVEALQG). Inosine was introduced where four different bases were needed. This decreased the mixture from 2048 to 32 and still represented all the possible coding sequences. The oligonucleotide mixture was end-labelled using [γ-^{32}P]ATP and T4 polynucleotide kinase (Boehringer-Mannheim, Lewes, Sussex, U.K.).

The library was plated out as above and duplicate filters were taken and processed for hybridization (Sambrook *et al.*, 1989) with the radiolabelled oligonucleotide mixture in a solution containing 6 × SCC, 5 × Denhardt's solution, 0.1% (w/v) SDS and 100 μg of denatured salmon sperm DNA/ml, overnight at 42 °C, followed by washing at room temperature for 30 min and a further 30 min at 42 °C. The library was also screened with a 275 bp human SP-D-specific fragment, coding for the *N*-terminal region of the globular domain, obtained by screening of the library with the antiserum. In this case, the 275 bp human SP-D fragment was labelled using a random-primed DNA-labelling kit (Boehringer-Mannheim, Mannheim, Germany). The filters were prehybridized in a solution containing 50% (w/v) formamide, 10% (w/v) dextran sulphate, 5 × Denhardt's solution, 0.1% (w/v) sodium pyrophosphate, 0.1% (w/v) SDS, 1 M-NaCl and 50 mM-Tris/HCl, pH 7.4, for 2 h at 42 °C, and then hybridized overnight at 42 °C with the labelled probe in the same buffer. The filters were washed for 30 min in 2 × SSC/0.1% (w/v) SDS at room temperature followed by 30 min in 1 × SSC/0.1% (w/v) SDS and a further 30 min in 0.1 × SSC/0.1% SDS at 65 °C. The washed filters were autoradiographed overnight.

Subcloning and sequencing

The sizes of the inserts of positive clones were analysed by the PCR method using two primers synthesized on the basis of the sequence around the *Eco*RI cloning site in λgt11. The inserts were cut out with *Eco*RI or alternatively with *Kpn*I and/or *Sst*I. The inserts were then shot-gun-ligated into Bluescript pBS KS$^+$ and subjected to double-strand sequencing using T7 sequencing kits (Amersham International plc, Aylesbury, Bucks., U.K.).

RESULTS

Purification of human SP-D and bovine SP-D

Human SP-D was initially purified from human lung lavage (results not shown). However, in order to overcome the difficulty of the collection of significant volumes of human lung lavage, amniotic fluid was used as an alternative source of starting material. The purification procedure, which was based on the Ca^{2+}-dependent affinity of SP-D for maltose, yielded approx. 50 μg of human SP-D from 1 litre of the supernatant obtained after centrifugation at 50000 *g*. Greater than 80% purity can be achieved by use of the first maltose–Sepharose column (Fig. 1*a*, lane 2). The second affinity step helps to remove albumin, serum amyloid P component and IgG (Fig. 1*a*, lane 3). F.p.l.c. separates SP-D from a trace amount of IgG (Fig. 1*a*, lane 4). It appears that the majority of SP-D in the lung is only loosely associated with lung surfactant whereas SP-A is more strongly associated with the surfactant, since the EDTA treatment only elutes a small fraction of SP-A, but all the SP-D, from the surfactant. The SP-D was eluted in two fractions from the Superose 6 column, one immediately after the void volume and the other at about 620 kDa (results not shown). Both fractions showed the same behaviour on SDS/PAGE in reducing and non-reducing condi-

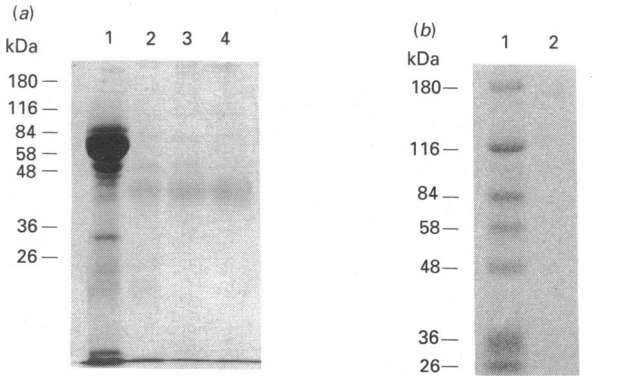

Fig. 1. SDS/PAGE under (*a*) reducing conditions on a 10% (w/v) gel or (*b*) non-reducing conditions on a 5–15% (w/v) gradient gel, of samples taken at different stages during the purification of human SP-D

(*a*) Lane 1, surfactant supernatant before application to maltose–Sepharose; lane 2, EDTA eluate from the first maltose–Sepharose column; lane 3, maltose eluate from the second maltose–Sepharose column; lane 4, 620 kDa peak from the f.p.l.c. (the void-volume peak from the f.p.l.c. column showed exactly the same behaviour as the 620 kDa peak). (*b*) Lane 1, standard proteins; lane 2, purified human SP-D 620 kDa peak from f.p.l.c.

Table 1. Amino acid compositions of human SP-D and rat SP-D

The amino acid composition of rat SP-D was taken from Persson *et al.* (1989). Cys was estimated as pyridylethylcysteine. ND, not determined.

	Residues per 100 residues	
Amino acid	Human SP-D	Rat SP-D
Cys	2.11	1.43
Hyp	3.10	1.43
Asp	6.99	7.73
Thr	2.42	3.55
Ser	5.86	6.56
Glu	13.04	14.18
Pro	6.20	4.82
Gly	22.97	21.17
Ala	9.41	11.28
Val	4.68	2.64
Met	1.63	0.83
Ile	1.46	1.91
Leu	4.28	6.01
Tyr	1.21	0.6
Phe	2.00	2.04
His	0.70	2.05
Hyl	1.23	0.43
Lys	3.90	4.91
Arg	4.14	4.70
Trp	ND	ND

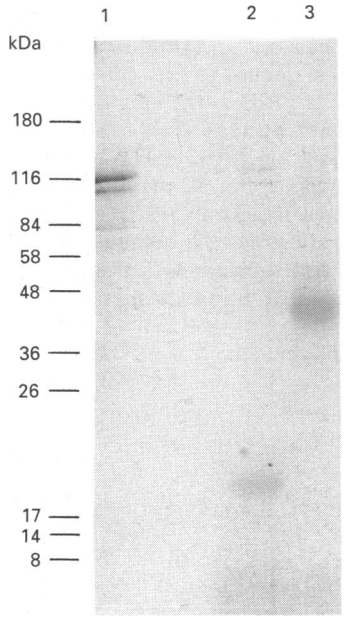

Fig. 2. SDS/PAGE under reducing conditions on a 5–15% (w/v) gradient gel

Lane 1, collagenase alone (at 10 times the concentration used in the digest); lane 2, human SP-D after digestion by collagenase; lane 3, human SP-D incubated with buffer alone.

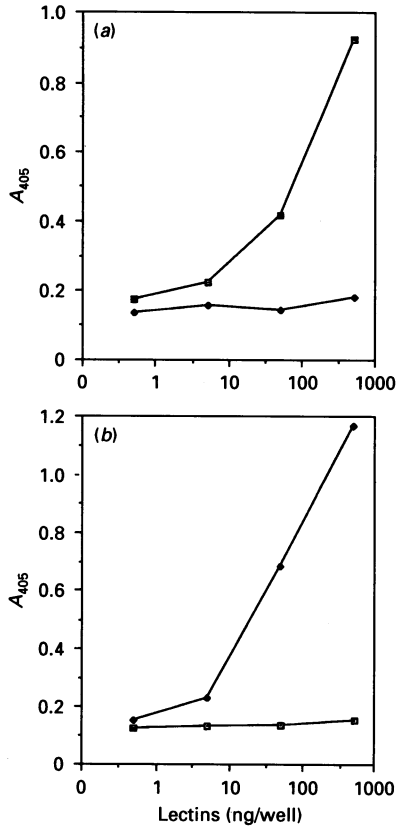

Fig. 3. Microtitre plates coated with (*a*) maltosyl-BSA or (*b*) mannan and incubated with dilutions of the lectins human SP-D (□) or human MBP (◆)

The degree of binding was estimated immunochemically as described in the text by measurement of the absorbance at 405 nm, which detected hydrolysis of the substrate for alkaline phosphatase.

Table 2. Saccharide specificities of human SP-D, rat SP-D and bovine conglutinin

Measurements were made for human SP-D and bovine conglutinin in this study. The values for rat SP-D are taken from Persson *et al.* (1990). ∞, No inhibition was observed by the sugar at 100 mM.

	I_{50} (mM)		
Sugar inhibitors	Human SP-D	Rat SP-D	Bovine conglutinin
Maltose	5.1	3.4	41.5
Fucose	6.5	31.0	41.5
Mannose	7.5	25.0	19.5
Methyl α-D-mannoside	9.3	21.0	42.0
Glucose	9.6	8.5	49.0
Glucosamine	18.6	> 100	53.5
Galactose	22.0	29.0	> 100
Lactose	23.5	31.0	∞
N-Acetyl-D-glucosamine	28.5	87.0	1.4
Galactosamine	∞	> 100	∞

tions (Fig. 1). The amino acid composition of human SP-D (Table 1) shows that, like rat SP-D (Persson *et al.*, 1989), it has a high content of glycine (22%) and that both hydroxyproline and hydroxylysine residues are present. These features are consistent with the view that human SP-D contains collagen-like structures and this was confirmed by its susceptibility to collagenase (Fig. 2). The collagenase digestion of human SP-D yielded a 20 kDa collagenase-resistant 'globular' fragment which retained affinity for maltose, as shown by its ability to bind to maltose–Sepharose and to be eluted from the maltose–Sepharose by EDTA (results not shown). Measurement of the affinity constant for the binding of the 20 kDa fragment to maltose–Sepharose was not made but it is likely to be considerably lower than that of the tetrameric 620 kDa intact SP-D molecule.

Bovine SP-D, isolated from lung lavage, showed exactly the

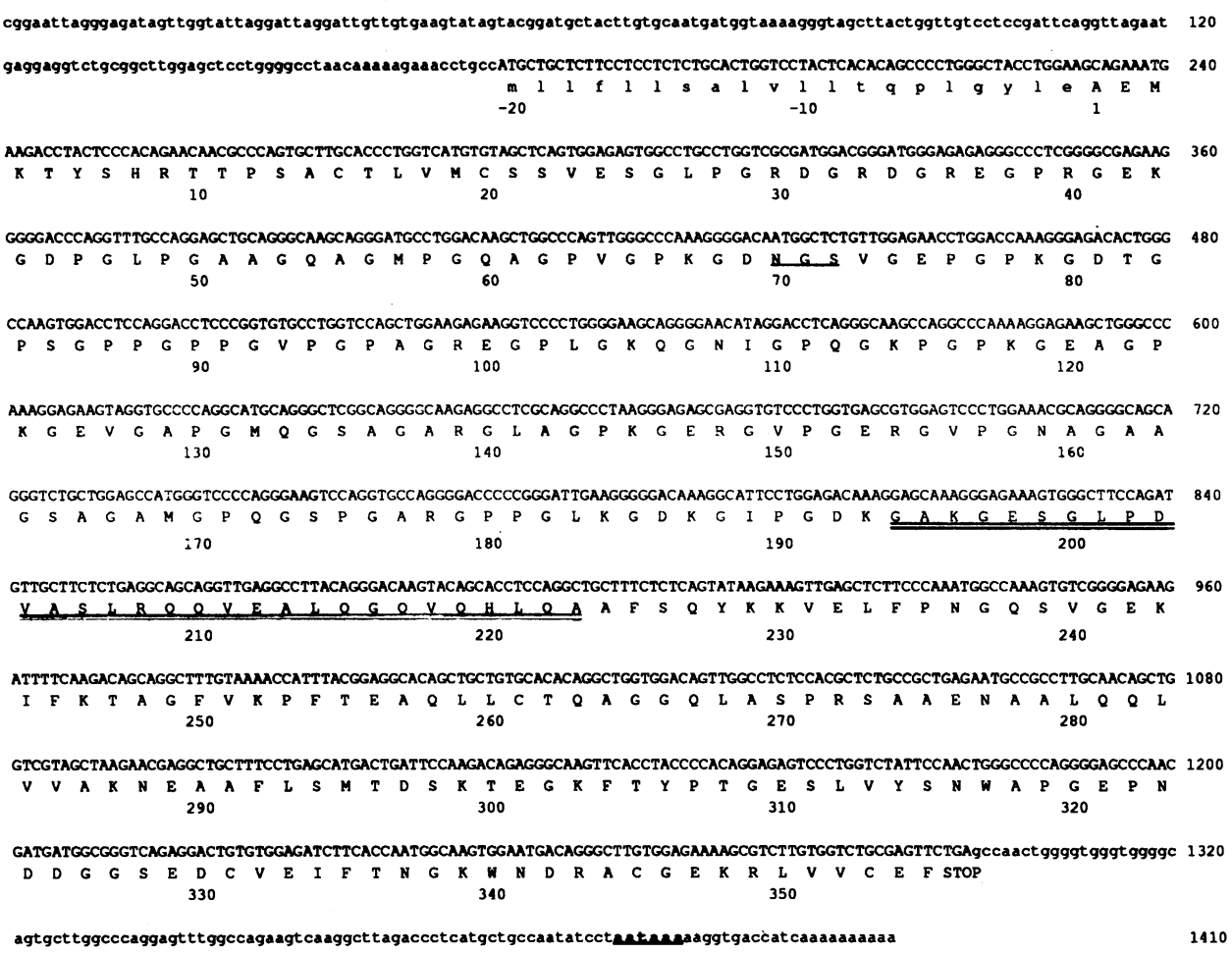

Fig. 4. cDNA and derived amino acid sequence of human SP-D

Amino acids are numbered under the lines with +1 taken to represent the *N*-terminus of the mature protein. Nucleotides considered to code for mature human SP-D are shown in upper case. The polyadenylation signal, AATAAA, in the 3′ untranslated region is shown in bold type underlined. The amino acid sequence (residues 194–223) identified in the 20 kDa collagenase-resistant fragment of human SP-D by automated sequencing is underlined (double line). The molecular mass of an unglycosylated single chain of human SP-D as calculated from the amino acid composition (assuming from the amino acid analysis the presence of six hydroxylysine residues and 12 hydroxyproline residues) of residues 1–355 is 39.26 kDa. If all these potential glycosylation sites are fully glycosylated, then the molecular mass of human SP-D would be approx. 43.6 kDa. The six hydroxylysine residues could be glycosylated by Glu-Gal disaccharide units and there is a single potential asparagine-linked glycosylation site at residue 70 which is underlined.

same behaviour on gel filtration in non-dissociating conditions, and also on SDS/PAGE, as human SP-D (results not shown).

Carbohydrate-binding properties of human SP-D

Human SP-D bound strongly to microtitre plates coated with maltosyl-BSA whereas it showed no binding to plates coated with mannan (Fig. 3). The specificity of binding is emphasized by the use of human MBP which binds strongly to mannan but shows no affinity for maltosyl-BSA (Fig. 3). The use of a variety of monosaccharides or disaccharides as inhibitors of the binding of human SP-D to maltosyl-BSA indicates that human SP-D has a similar carbohydrate-binding specificity to rat SP-D (Table 2) but that it has a clearly distinct specificity from that of other mammalian C-type lectins, such as bovine conglutinin (Young & Leon, 1987; Loveless *et al.*, 1989) and human MBP (Childs *et al.*, 1989).

Isolation of cDNA clones for human SP-D

Amino acid sequence analysis of the 20 kDa collagenase-resistant fragment of human SP-D yielded the following sequence: GAKGESGLPDVASLRQQVEALQGXVXXLQA (where X denotes an unidentified residue). This sequence information was used to construct oligonucleotides for use in screening the human lung cDNA library. The library was first screened, using the rabbit anti-(human SP-D) serum, and 13 potential positive clones were obtained after rescreening. One of these clones, identified as positive by screening with the oligonucleotides, as well as the antiserum, contained a 275 bp insert which encoded a sequence corresponding to that obtained at the protein level from the collagenase-resistant fragment. Twenty-four further potential human SP-D cDNA clones were identified from the library by use of the oligonucleotide and 275 bp cDNA probe. This allowed the identification and characterization of two clones which appear to cover the entire coding sequence of human SP-D (Fig. 4). The cDNA data contains 171 nucleotides of apparent 5′-non-translated sequence, followed by 60 nucleotides corresponding to a leader peptide of 20 amino acid residues, 1065 nucleotides corresponding to the coding sequence of the mature protein and 114 nucleotides of apparent 3′-non-translated sequence which includes an AATAAA polyadenylation signal

Fig. 5. Alignment of the human SP-D and bovine conglutinin (Lee *et al.*, 1991) sequences

Residues common to both sequences are boxed. There is an overall identity of approx. 66% between the two sequences. '-' indicates gap or insertion introduced to obtain maximum similarity. The limited amount of amino acid sequence available for rat SP-D (rSP-D) (Persson *et al.*, 1989) and for bovine SP-D (bSP-D) (J. Lu, A. C. Willis and K. B. M. Reid, unpublished work) has been aligned with the human SP-D (hSP-D) sequence. ? denotes an unidentified residue. * denote peptide fragments of rat SP-D (Persson *et al.*, 1989). Residues in the rat SP-D and bovine SP-D sequences which are common to those in human SP-D, and/or conglutinin, are boxed. The conserved residues found in the carbohydrate-recognition domain of C-type lectins are shaded.

and a short poly(A) tail. The derived amino acid sequence of the mature protein predicts that it is 355 residues long, having a short non-collagen-like region of 25 residues followed by a collagen-like region of 177 residues and a *C*-terminal 'globular' domain of 153 residues which contains the C-type lectin domain (Fig. 4). The *N*-terminal residue of human SP-D appeared to be blocked, as judged by *N*-terminal sequence analysis of the electroblotted intact 43 kDa chain. However, in view of the marked similarity between the human SP-D and bovine conglutinin sequences (Fig. 5), the alanine residue in human SP-D, which aligned with the *N*-terminal alanine residue (with a free amino group) in bovine conglutinin, was taken as the probable *N*-terminal residue of human SP-D. Comparison of the amino acid sequence of human SP-D with that of bovine conglutinin (Lee *et al.*, 1991) shows a remarkable degree of similarity with 66% identity even without taking conservative amino acid changes into account. If conservative changes are included, the identity rises to over 70%. The human SP-D sequence shows 71% and 68% identity with the limited data available for the rat SP-D and bovine SP-D sequences respectively (Fig. 5). The absorption coefficient, $A^{1\%}_{1cm,280}$, of human SP-D, as calculated from the derived amino acid sequence, is 4.01.

DISCUSSION

The molecule characterized in this study, at both the protein and cDNA levels, has been designated as human SP-D in view of: its strong Ca^{2+}-dependent affinity for maltose; its presence in human lung surfactant and amniotic fluid but not in serum; its overall similarity in structure to rat SP-D (with respect to size in non-dissociating conditions, behaviour on SDS/PAGE and presence of collagen-like sequences); the isolation of the cDNA clones from a lung cDNA library. Only a limited amount of amino acid sequence (from two CNBr peptides isolated from the collagen-like region of the molecule) has been reported for rat SP-D [Persson *et al.* (1989), as summarized in Fig. 5]. Thus the human SP-D cDNA data allow the derivation of the entire amino acid sequence of SP-D and show that the molecule is a member of the family of lectins having a C-type (Ca^{2+}-dependent) carbohydrate-recognition domain and containing collagen-like sequences (Thiel & Reid, 1989; Bezouska *et al.*, 1991). The C-type lectin carbohydrate-recognition domain contains approx. 130 amino acid residues and is characterized by 17–18 invariant residues (Fig. 5), four of which are cysteine residues considered to form interchain disulphide bonds in the fashion Cys-1–Cys-4 and Cys-2–Cys-3 (Drickamer, 1988). The precise maturation cleavage point in the derived SP-D sequence is not known, since the intact SP-D protein has a blocked *N*-terminal amino acid, but alignment of the derived human SP-D sequence with the bovine conglutinin sequence (Fig. 5) indicates that alanine is likely to be the *N*-terminal residue. Also, it can be calculated from the amino acid-analysis data (Table 1) that approx. 24% of the lysine residues and 33% of the proline residues in the derived sequence for human SP-D are hydroxylated, i.e. six hydroxylysine

residues and 12 hydroxyproline residues per 355 amino-acid-long polypeptide.

The human SP-D molecule, as isolated from amniotic fluid, appears to contain four identical subunits of 150 kDa, since the whole molecule has a molecular mass of approx. 620 kDa in non-dissociating conditions and yields a band of apparent molecular mass 150 kDa when examined, without reduction of disulphide bonds by SDS/PAGE (Fig. 1). Each subunit is composed of three identical disulphide-linked chains of approx. 43 kDa, as judged by SDS/PAGE in reducing conditions (Fig. 1) and protein sequence analysis. The chains are considered to form a collagen-like triple helix, via the Gly-Xaa-Yaa- repeating-triplet region within residues 26–193 (Fig. 4), as judged by the susceptibility of this region to collagenase (Fig. 2). Also the presence of hydroxyproline and hydroxylysine in human SP-D (Table 1) is consistent with the molecule containing triple-helical structure. The collagenase digestion of human SP-D yielded a 20 kDa fragment on SDS/PAGE in both non-reducing and reducing conditions (Fig. 2). Since this fragment corresponds to the *C*-terminal 161 amino acids of human SP-D (residues 194–355, Fig. 4), it is clear that the interchain disulphide bonds, joining the three chains in each 150 kDa subunit, are formed between the cysteine residues, at positions 15 and 20, in the short *N*-terminal non-collagen-like regions of each chain (Fig. 4), but it has not yet been established if this involves only Cys-15–Cys-20 linkages or both Cys-15–Cys-15 plus Cys-20–Cys-20 linkages between chains.

The derived amino acid sequence of human SP-D (Figs. 4 and 5) shows that it belongs to the 'Group III' mammalian lectins, i.e. those containing C-type carbohydrate-recognition domains and having collagenous domains (Bezouska *et al.*, 1991). This group includes the rat A, rat C and human mannose-binding proteins, bovine conglutinin and dog and human lung surfactant SP-A. The human SP-D sequence shows approx. 70% identity with regions of the sequence of rat SP-D and bovine SP-D (Fig. 5), but, surprisingly, also showed 66% identity on alignment with the entire chain of bovine conglutinin (Fig. 5), which is a serum lectin which has a carbohydrate specificity for terminal *N*-acetyl-D-glucosamine and mannose and shows little affinity for maltose (Table 2). The similarity between SP-D and conglutinin is emphasized by the 80% identity in sequence between a 40-residue stretch of bovine SP-D aligned with the bovine conglutinin sequence (Fig. 5). Thus in both overall structure and amino acid sequence, SP-D appears to be more closely related to conglutinin than to SP-A or mannose-binding protein. The saccharide specificity of human SP-D was found to be very similar to that of rat SP-D, except for the stronger affinity of human SP-D for glucosamine (Table 2). The derived amino acid sequence of the *C*-terminal half of human SP-D shows that it contains 17 out of the 18 highly conserved 'framework' residues found in the carbohydrate-recognition domains of the C-type lectins (Fig. 5). The lack of binding of human SP-D to mannan-coated microtitre plates (Fig. 3) emphasizes the difference in specificity between SP-D and MBP.

The physiological role of SP-D in the lung is unknown, but if it shows the capacity to interact via its collagen-like regions with the C1q receptor, which is found on the surface of a wide range of lymphoid cells (Ghebrehiwet, 1989), then it could be involved in the antibody-independent recognition and clearance of pathogens from lung fluids. It has been shown that C1q, MBP, SP-A and conglutinin can all bind to the C1q receptor (Malhotra *et al.*, 1990) and therefore it would not be surprising, especially in view of the overall structural similarity between conglutinin and SP-D, if SP-D also binds to this receptor. This would then provide an antibody-independent mechanism of recognition of certain pathogens, containing cell-surface carbohydrate residues, via the globular carbohydrate-recognition domains in SP-D and a presentation/clearance mechanism of the pathogen–SP-D complex via the collagen regions in SP-D interacting with the C1q receptor.

Note added in proof (received 10 March 1992)

The complete primary structure of rat SP-D has been derived by cDNA cloning (Shimizu *et al.*, 1992). A cDNA clone for human SP-D has also recently been characterized by Rust *et al.* (1991).

We thank Ms. J. Parsons for performing the electroblotting before protein sequencing and Dr. P. Phizackerly for arranging the collection of amniotic fluid. We also thank Dr. S. Thiel for providing purified bovine conglutinin.

REFERENCES

Baritussio, A. G., Magoon, M. W., Goerke, J. & Clements, J. A. (1981) Biochim. Biophys. Acta **666**, 382–393
Bezouska, K., Crichlow, G. V., Rose, J. M., Taylor, M. E. & Drickamer, K. (1991) J. Biol. Chem. **266**, 11604–11609
Childs, R. A., Drickamer, K., Kawasaki, T., Thiel, S., Mizuochi, T. & Feizi, T. (1989) Biochem. J. **262**, 131–138
Crawford, C., Brown, N. R. & Willis, A. C. (1990) Biochem. J. **265**, 575–579
Curstedt, J., Jornvall, H., Robertson, B., Bergman, T. & Berggren, P. (1987) Eur. J. Biochem. **168**, 255–262
Davis, III, A. E. & Lachmann, P. J. (1984) Biochemistry **23**, 2139–2144
Dobbs, L. G., Wright, J. R., Hawgood, S., Gonzalez, R., Venstrom, K. & Nellenbogen, J. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 1010–1014
Drickamer, K. (1988) J. Biol. Chem. **263**, 9557–9560
Fisher, J. H., Shannon, J. M., Hofmann, T. & Mason, R. J. (1989) Biochim. Biophys. Acta **995**, 225–230
Floros, J., Steinbrink, R., Jacobs, K., Phelps, D., Kriz, R., Recny, M., Suitzman, L., Jones, S., Taeusch, H. W., Frank, H. A. & Frisch, E. F. (1986) J. Biol. Chem. **261**, 9029–9033
Ghebrehiwet, B. (1989) Behring. Inst. Mitt. **84**, 20–30
Glasser, S. W., Korfhagen, T. R., Weaver, T., Pilot-Matias, T., Fox, J. L. & Whitsett, J. A. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 4007–4011
Haagsman, H. P., Hawgood, S., Sargeant, T., Buckley, D. & White, R. T. (1987) J. Biol. Chem. **262**, 13877–13880
Hawgood, S., Benson, B. J., Schilling, J., Damm, D., Clements, J. A. & White, R. T. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 66–70
King, R. J., Klass, D. J., Gikas, E. G. & Clements, J. A. (1973) Am. J. Physiol. **224**, 788–795
King, R. J., Ruch, J., Gikas, E. G., Platzker, A. C. G. & Creasy, R. K. (1975) J. Appl. Physiol. **39**, 735–741
King, R. J., Phillips, M. C., Horowitz, P. M. & Dang, S. C. (1986) Biochim. Biophys. Acta **879**, 1–13
Kuroki, Y., Mason, R. J. & Voelker, D. R. (1988) J. Biol. Chem. **263**, 3388–3394
Laemmli, U. K. (1970) Nature (London) **227**, 680–685
Lee, Y. M., Leiby, K. R., Allar, J., Paris, K., Lerch, B. & Okarma, T. B. (1991) J. Biol. Chem. **266**, 2715–2723
Loveless, R. W., Feizi, T., Childs, R. A., Mizuochi, T., Stoll, M. S., Oldroyd, R. G. & Lachmann, P. J. (1989) Biochem. J. **258**, 109–113
Macklin, C. C. (1954) Lancet **i**, 1099–1104
Malhotra, R., Thiel, S., Ried, K. B. M. & Sim, R. B. (1990) J. Exp. Med. **172**, 955–959
Ogasawara, Y., Kuroki, Y., Shiratori, M., Shimizu, H., Miyamura, K. & Akino, T. (1991) Biochim. Biophys. Acta **1083**, 252–256
Persson, A., Rust, K., Chang, D., Moxley, M., Longmore, W. & Crouch, E. (1988) Biochemistry **27**, 8576–8584
Persson, A., Chang, D., Rust, K., Moxley, M., Longmore, W. & Crouch, E. (1989) Biochemistry **28**, 6361–6367
Persson, A., Chang, D. & Crouch, E. (1990) J. Biol. Chem. **265**, 5755–5760
Phelps, D. S. & Floros, J. (1988) Am. Rev. Respir. Dis. **137**, 939–942
Rice, W. R., Ross, G. F., Singleton, F. M., Dingle, S. & Whitsett, J. A. (1987) J. Appl. Physiol. **63**, 692–698
Rust, K., Grosso, L., Zhang, V., Chang, D., Persson, A., Longmore, W., Cai, G. Z. & Crouch, E. (1991) Arch. Biochem. Biophys. **290**, 116–126
Ryan, R. M., Morris, R. E., Rice, W. R., Ciraolo, G. & Whitsett, J. A. (1989) J. Histochem. Cytochem. **37**, 429–440

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) Mol. Clon. **2**, 12.16–12.20
Schwarts, B. A. & Gray, G. R. (1977) Arch. Biochem. Biophys. **181**, 542–549
Shimizu, H., Fisher, J. H., Papst, P., Benson, B., Lau, K., Mason, R. J. & Voelker, D. R. (1992) J. Biol. Chem. **267**, 1853–1857
Suzuki, Y., Curstedt, T., Grossman, G., Kobayashi, T., Nilsson, R., Nohara, K. & Robertson, B. (1986) Eur. J. Respir. Dis. **69**, 336–345
Suzuki, Y., Fujita, Y. & Kogishi, K. (1989) Am. Rev. Respir. Dis. **140**, 75–81
Takahashi, A. & Fujiwara, T. (1986) Biochem. Biophys. Res. Commun. **135**, 527–532
Tenner, A. J., Robinson, S. L., Borchelt, J. & Wright, J. R. (1989) J. Biol. Chem. **264**, 13923–13928
Thiel, S. & Reid, K. B. M. (1989) FEBS Lett. **250**, 78–84

Received 8 August 1991/28 October 1991; accepted 5 November 1991

Voss, T., Eistetter, H., Schafer, K. P. & Engel, J. (1988) J. Mol. Biol. **201**, 219–227
Walker, S. R., Williams, M. C. & Benson, B. (1986) J. Histochem. Cytochem. **34**, 1137–1148
Warr, R. G., Hawgood, S., Buckley, D. I., Crisp, T. M., Schilling, J., Benson, B. J., Ballard, P. L., Clements, J. A. & White, R. T. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 7915–7919
White, R. T., Damm, D., Miller, J., Spratt, K., Schilling, J., Hawgood, S., Benson, B. & Cordell, B. (1985) Nature (London) **317**, 361–363
Young, N. M. & Leon, M. A. (1987) Biochem. Biophys. Res. Commun. **143**, 645–651
Young, S. L., Kremers, S. A., Apple, J. S., Crapo, J. D. & Brumley, G. W. (1981) J. Appl. Physiol. **51**, 248–253