Production and characterisation of recombinant forms of human pulmonary surfactant protein C (SP-C): Structure and surface activity[#]

Dunja Lukovic[1], Inés Plasencia[2,¥], Francisco J. Taberner[1], Jesús Salgado[1,3], Juan J. Calvete[4], Jesús Pérez-Gil[2,*] and Ismael Mingarro[1,*]

[1]Departament de Bioquímica i Biologia Molecular
Universitat de València
E-46 100 Burjassot (Valencia), Spain

[2]Departmento de Bioquímica y Biología Molecular I,
Universidad Complutense Madrid, Spain

[3]Instituto de Ciencia Molecular
Universitat de València
E-46 980 Paterna (Valencia), Spain

[4]Instituto de Biomedicina de Valencia, CSIC, Valencia, Spain

Running title: Structure and surface activity of recombinant SP-C
Text: Microsoft Word X for Mac (23 pages)
Tables included: 1   Figures included: 8

[*] Corresponding authors: I. Mingarro, Dept. Bioquímica i Biol. Molec., Universitat de València. Dr. Moliner, 50. E-46100 Burjassot, Spain. Phone: +34-963543796; Fax: +34-963544635. E-mail: Ismael.Mingarro@uv.es; and J. Pérez-Gil, Dept. Bioquímica y Biol. Molec. I, Facultad de Biología. Universidad Complutense. E-28040 Madrid, Spain. Phone: +34-913944994. E-mail: jpg@bbm1.ucm.es

[¥] Present address: Centre for membrane Physics, MEMPHYS, University of Southern Denmark, Odense.

Abstract

Surfactant protein C (SP-C) is an essential component for the surface tension-lowering activity of the pulmonary surfactant system. It contains a valine-rich α helix that spans the lipid bilayer, and is one of the most hydrophobic proteins known so far. SP-C is also an essential component of various surfactant preparations of animal origin currently used to treat neonatal respiratory distress syndrome (NRDS) in preterm infants. The limited supply of this material, and the risk of transmission of infectious agents and immunological reactions have prompted the development of synthetic SP-C-derived peptides or recombinant humanized SP-C for inclusion in new preparations for therapeutic use.

We describe herein the recombinant production in bacterial cultures of SP-C variants containing phenylalanines instead of the palmitoylated cysteines of the native protein, as fusions to the hydrophilic nuclease A (SN) from *Staphylococcus aureus.* The resulting chimerae were partially purified by affinity chromatography and subsequently subjected to protease digestion. The SP-C forms were recovered from the digestion mixtures by organic extraction and further purified by size exclusion chromatography. The two recombinant SP-C variants so obtained retained more than 50% α-helical content and showed surface activity comparable to the native protein, as measured by surface spreading of lipid/protein suspensions and from compression π-A isotherms of lipid/protein films. Compared to the protein purified from porcine lungs, the recombinant SP-C forms improved movement of phospholipid molecules into the interface (during adsorption), or out from the interfacial film (during compression), suggesting new possibilities to develop improved therapeutic preparations.

Abbreviations

DPPC: 1,2-dipalmitoyl phosphatidylcholine; LPC: lysophosphatidylcholine; LS: lung surfactant; NRDS: neonatal respiratory distress syndrome; POPG: palmitoyl oleyl phosphatidylglycerol; SDS-PAGE: sodium dodecylsulfate polyacryamide-gel electrophoresis; SP-C: surfactant protein C; rSP-C: recombinant SP-C; SN: nuclease A from *Staphylococcus aureus*; TM: transmembrane.

INTRODUCTION

Surfactant–associated protein C (SP-C) together with a second hydrophobic surfactant protein, SP-B, and phospholipids represents the active component of the lung surfactant (LS). LS is the surface active material present at the air-liquid interface in the alveoli of mammalian lungs. Its role is to lower the normal air-water surface tension, from 70 mN/m to near zero, upon expiration, thereby preventing alveolar collapse and minimizing the work of their expansion during inhalation [1]. Deficiency or inactivation of LS causes pulmonary dysfunctions, the most common being neonatal respiratory distress syndrome (NRDS), caused by the absence of surfactant in premature infants. Replacement surfactants currently applied in clinical therapy are obtained from animal lungs. However, this is an expensive and inexact source, and levels of SP-B and SP-C in commercial surfactant preparations vary from batch to batch and are typically lower that in native pulmonary surfactants [2, 3]. Additionally, the potential risks of transmission of animal-derived diseases or immunological reaction justify the need for a human-like standardised material as a suitable alternative [4]. The possibility to optimise artificial surfactant formulations to treat specific diseases should also open new perspectives for the treatment of respiratory pathologies.

Several efforts have been made in order to develop artificial mimics of the natural surfactant by combining synthetic lipids and peptides that correspond to segments of native SP-B and SP-C. Tests *in vitro* and in animal models show that their biophysical activity is usually superior to the one of phospholipids alone [5]. However, production of successful, entirely synthetic surfactant preparations has been so far hindered by the difficulties in obtaining both surfactant proteins, SP-B and SP-C, mainly due to their high hydrophobicity.

SP-C, one of the most hydrophobic proteins in the proteome, is a 35 amino acid polypeptide that constitutes approximately 1% of the surfactant mass [6]. It is composed of a short N-terminal region, palmitoylated at residues Cys-5 and Cys-6, and followed by an α-helical hydrophobic transmembrane (TM) C-terminal stretch made of aliphatic residues. The sequence of SP-C is highly conserved through evolution, with predominance of valine residues in the TM domain and only small differences between species (Fig. 1A). It is proposed that SP-C is involved in the formation of multilayer lipid stacks

attached to the air-liquid interface, which act as a source of reservoir of lipids that can be reinserted into the monolayer during alveolar expansion [7]. However, detailed molecular mechanistic studies are still missing.

Apart from the above mentioned extreme hydrophobicity, the difficulty for obtaining artificial SP-C analogues derives from its strong tendency to misfold and aggregate in the absence of phospholipids [8, 9]. The TM α helical fragment, represented by 10-12 valines with intermittent Ile and Leu residues, is capable of populating irreversibly into a fibrilar β-structure aggregate. In aqueous solutions β-branched valines and isoleucines display a preference toward β-strands and are usually underrepresented in α helices. However, in lipid environments Val and Ile residues promote helix formation [10]. In agreement with this general tendency, the α helical polyVal-rich region of SP-C was found to convert spontaneously into β sheets *in vitro*, with an enhanced rate of α→β transition in the presence of polar solvents and deacylation [11]. Current data emphasize a strong correlation between the loss of helical structure of SP-C and its malfunctioning [12]. Moreover, SP-C amyloid fibrils are found in patients with pulmonary alveolar proteinosis [8].

The highly hydrophobic membrane proteins are generally difficult to over-express and purify [13]. They are usually produced in heterologous systems, as fusions with soluble portions of bacterial proteins. The presence of detergents during the purification is often required in order to prevent formation of aggregates, especially at the point where the TM portion is cleaved from its fusion. Following enzymatic cleavage the TM proteins are generally purified by reversed phase (RP) HPLC [14]. Membrane proteins that do not have fibril forming propensities can, once unfolded and precipitated, be recovered in the presence of detergents or lipids. However, this three-step method is not readily applicable to SP-C, since purifying the protein by RP-HPLC bears the risk of its denaturation leading to irreversible aggregation and complete inactivation. A method has been published for expression and purification of recombinant SP-C (rSPC) in *Escherichia coli* as a fusion to the N-terminal portion of bacterial chloramphenicol acetyl transferase [15]. Surfactant preparations containing this version of rSP-C as the unique protein component have shown to be effective for the treatment of respiratory distress [16, 17].

In this report, we describe a novel approach for expression and purification of human SP-C, based on the combination of a successful TM protein expression strategy in bacteria with the established organic extraction protocol typically used for separation of native SP-C from LS mixtures. This new procedure is very effective for preserving the native helical structure of SP-C throughout the purification protocol and for producing protein forms with high surface activity in reasonable amounts. We have modified the SP-C sequence by Cys5Phe and Cys6Phe substitutions. The rationale behind these changes is that the naturally occurring palmitoylation of the two Cys residues cannot be achieved through bacterial protein expression. The modified protein remains very close to the native sequence present in some animal surfactants (Fig. 1). Moreover, the same Cys→Phe replacements are present in the SP-C analogue used in a successful synthetic surfactant preparation [18]. Detailed analysis of the highly pure rSP-C forms produced in this work reveals a slightly improved surface activity when compared to the native SP-C obtained from animal lungs.

MATERIALS AND METHODS

Protein expression and purification

Construction of plasmids encoding the His-tagged chimerical protein (SN/GpA) has been described [19]. The glycophorin A (GpA) sequence in the original plasmid was replaced by a PCR-amplified sequence of human SP-C (generous gift from Dr. Joanna Floros, University of Pennsylvania at Hershey), using *ApaI/BamHI* restriction sites. The N-terminal His-tag, the optimised thrombin digestion site, and all point mutations in the wild type sequence of SP-C were introduced by site-directed mutagenesis using the QuikChange kit from Stratagene (La Jolla, California) and verified by DNA sequencing.

The expression of the fusion protein was performed as described earlier [20]. Briefly, a 10 mL overnight culture of BL21 (DE3) pLys (Novagen, Madison, WI) cells transformed with the corresponding vector was inoculated into 1 L of LB medium and grown at 37 °C until the OD_{600} reached 0.6. IPTG (1 mM) was added for induction and the culture was grown for additional 3 hours. The cells were harvested by centrifugation at 3800 g for 10 min in a Sorval centrifuge and resuspended in resuspension buffer (10 mM Tris HCl pH 7.9, 1 mM EDTA, 0.01 M PMSF). Pellets were frozen until use. In order to break the cellular membranes, the pellets were freeze-thawed three times, sonicated and added to the same volume of TBS 1% lauroyl sarcosine detergent. This mixture was centrifuged at 18000 g for 10 minutes. The fusion protein was purified from the supernatant via immobilized metal affinity chromatography. An ÄKTA purifier liquid chromatography system (Amersham Bioscience) was used with a His-Trap column filled with nickel containing resin. The buffers for chimerical protein purifications were the following: equilibration buffer, TBS 0.5% lauroyl sarcosine; wash buffer, TBS 0.2% lauroyl sarcosine 10mM imidazol; and elution buffer, TBS 0.2% lauroyl sarcosine 500 mM imidazol. The eluted protein was dialysed to remove imidazol and finally concentrated by using an Amicon concentrator with YM10 ultra-filtration membrane (Milipore, Bedford MA) 5-6 times. Generally, 4 mg of the fusion protein were obtained per 1 litre of LB medium.

Thrombin digestion

A thrombin cleavage site (LVPR↓GP), introduced between SN and SP-C, allowed the proteolysis of the fusion protein. A glycine-rich 'kinker' next to the thrombin cleavage site was introduced to improve the cleavage efficiency [21]. Thrombin (Novagen) digestion trials were performed before each large-scale digestion. Usually, 10 U of thrombin were required for digestion of 1 mg rSP-C and 1 U of thrombin was needed for 1 mg of rGP/SP-C. Thrombin digestions were performed at 21 °C in the buffer recommended by the supplier, with shaking at 550rpm in eppendorf thermomixer. The reactions were concentrated again using a cell concentrator with membranes of 1 kDa molecular cut off. Aliquots taken from the digestion reaction mixture were applied to tris-tricine gels in order to monitor the appearance of the bands corresponding to the *Staphylococcal nuclease* (18.5 kDa) and rSP-C (3.9kDa) (Fig. 2B).

Organic extraction

An organic extraction of the protein solution was performed according to the protocol of Bligh and Dyer [22] for lipid extraction. Shortly, two volumes of methanol and one volume of chloroform were added to the protein mixture, which was vortexed and incubated at 37°C. After addition of one more volume of chloroform and one of water, the mixture was vortexed and centrifuged at 550 g/5min. The lower, organic, phase was collected and two volumes of chloroform were added two more times to the aqueous phase in order to increase protein recovery. All organic phases were stored in glass bottles at -20°C. The organic extract was concentrated to approximately 2 mL prior to application to the lipophilic column. Protein concentration was performed in the presence of egg yolk PC (Avanti Polar Lipids, Alabaster, AL) at a 1:5 (w:w) ratio in a rota-vaporizer apparatus.

Sephadex LH-20 lipophilic size exclusion chromatography

In order to eliminate components other than SP-C that could have co-purified in the organic phase, the organic extraction was applied to a Sephadex LH-20 column (Amersham Biosciences). This column is used for lipid separation and is adapted to the use with organic solvents. Proteins and lipids are eluted from the column by a chlorofom:metanol (2:1, v:v) eluting system. Fractions were collected and absorbance at 240 and 280 nm was recorded on a

spectrophotometer. Absorbance peaks were pooled and loaded on SDS-PAGE and subsequently analysed for their amino acid content and by mass spectrometry.

Mass spectrometry

For determination of the molecular masses, 0.85 µL of the of peptide dissolved in chloroform:methanol (2:1, v:v) were spotted onto a MALDI-TOF sample holder and allowed to dry at room temperature. An equal volume of a saturated solution of sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid) (Sigma) saturated in 70% acetonitrile containing 0.1% TFA was added, let dry, and analyzed with an Applied Biosystems Voyager-DE Pro MALDI-TOF mass spectrometer, operated in delayed extraction and linear or reflector modes. The mass calibration standard in linear mode consisted of a mixture of the following proteins, whose calculated isotope-averaged molecular masses in Daltons are given between brackets: bovine insulin (5734.6), *E. coli* thioredoxin (11674.5), horse apomyoglobin (16952.6). For reflector calibration, a tryptic peptide mixture of *Cratylia floribunda* seed lectin (SwissProt accession code P81517) prepared and previously characterised in our laboratory was used.

Circular dichroism

To prepare the samples for CD measurements, the purified protein (50 µg) in chloroform:methanol solutions was mixed with lysophosphatidylcholine (LPC) or DPPC:POPG (7:3) at a protein: lipid ratio of 1:5 (w:w), dried under a stream of N_2 and then under vacuum for 2h. The resulting dried films were hydrated by addition of 0.5 mL of 5mM Tris buffer (pH 7) containing 150 mM NaCl with intermittent vortexing at room temperature for LPC micelles, or at 55°C for DPPC:POPG vesicles. The suspensions were then sonicated during 1 min. CD spectra were obtained on a Jasco J-810 CD spectropolarimeter. Quartz cells of 1 mm optical path were used to record spectra at a scanning speed of 50 nm/min. Generally, 15 scans were accumulated and averaged, subjected to noise reduction and corresponding blanks subtracted. Data analysis was performed with the help of the CDPro software package, which contains three commonly

used programs: SELCON3, CONTIN/LL and CDSSTR [23, 24]. This software allows the use of different reference sets of proteins, including 13 membrane proteins (SMP50), to increase the reliability of the analysis.

SDS-PAGE

Fusion proteins and protein samples obtained after organic extraction or eluted from the Sephadex column were analyzed onto 16% SDS-PAGE. The SP-C peptide samples were first evaporated under Nitrogen gas, resuspended in the loading buffer, boiled 5 min at 95ºC and applied to the gel. After running, gels were fixed, stained with Comassie blue stain and distained. In order to resolve small molecular mass peptides generated after thrombin digestion, the samples were loaded on tricine 15% acrylamide gel. Tricine gels were processed identically as SDS-PAGE gels.

Interfacial adsorption

Adsorption π-t kinetics for the formation of phospholipids films as promoted by native and recombinant forms of SP-C, were monitored on a Wilhemy surface balance (Nima, Coventry, UK) equipped with a minitrough (15 cm^2 of surface, 5 mL volume of sub-phase) and using a Whatman Nº 1 paper dipping plate attached to the pressure transducer. Hundred μg of a mixture of DPPC:POPG (7:3, w:w) containing 0, 2, 5 or 10% protein (w/w) were reconstituted in 100 μL of 5 mM Tris buffer pH 7.0 containing 150 mM NaCl. To start acquisition of the π-t adsorption isotherms, 10 μL of this suspension were deposited directly and allowed to spread on top of the surface of the balance and film formation was monitored following the changes in surface pressure with time. Experiments were performed at 25 ºC and repeated at least three times with two different samples for each protein.

π-A compression isotherms

Monolayers of lipid (DPPC:DPPG, 7:3, w/w) or lipid/peptide binary systems were made by spreading a small volume of a concentrated lipid or lipid/peptide solution in chloroform:methanol (3:1, v/v) at the surface of a 5 mM Tris buffered sub-phase, pH 7, containing 150 mM NaCl. These monolayers were prepared as previously described [25], in a thermostatic

(25°C) Langmuir-Blodgett trough (NIMA Technologies, Coventry, United Kingdom) equipped with a ribbon barrier to minimize film leakage during compression. Sub-phases were prepared with double distilled water (the second distillation performed in the presence of potassium permanganate). After spreading the sample on top of the sub-phase, the organic solvent was allowed to evaporate for 10 minutes before starting compression. The total area of the interface was 225 cm^2 and the monolayer was compressed at 65 cm^2/min, while changes in surface pressure were recorded and plotted against the area occupied per phospholipid molecule.

RESULTS

Over-expression and digestion of chimeric proteins

rSP-C is based on the primary sequence of human SP-C (Fig. 1). Two critical issues had to be considered to define the expression and purification strategy of this protein: the hydrophobicity conferred by the polyVal stretch and the palmitoylation at Cys-5 and Cys-6, which represents an insuperable challenge for the bacterial expression machinery. The high hydrophobicity was overcome by making a fusion with the hydrophilic nuclease A (SN) from *Staphylococcus aureus* (SN/SP-C, see Materials and Methods). An optimised thrombin cleavage site [21] is introduced between the SN and the SP-C sequence to allow removal of the nuclease moiety (Fig. 1C). In order to prevent possible dimer formation via disulphide bonds in the absence of palmitic chains we introduced a double $C_5C_6 \rightarrow F_5F_6$ mutation [26]. Phenylalanine residues were chosen because: i) they represent a native surrogate to the missing Cys-6 in the native SP-C of some animals (Fig.1A), ii) they have high propensity to partition into the membrane interface [27], and iii) the Cys→Phe substitutions were shown not to affect the activity of the SP-C protein [18]. The fusion is preceded by a His-tag to enable its purification through nickel-agarose affinity chromatography. The gels in Fig. 2 show that the nuclease fusion proteins can be expressed to high levels and partially purified on a Ni-agarose resin (Fig. 2A), following which the nuclease moiety is removed by specific digestion with thrombin. The two major protein bands emerging after cleavage correspond to the nuclease and the rSP-C moieties, respectively (Fig. 2B). The lower intensity of the rSP-C (Fig. 2B, lane 3, bottom band) compared to the nuclease moiety (top band) results from the six times lower molecular weight of the former. Importantly, only less than 10% of the total SN/SP-C fusion proteins are expressed as inclusion bodies, as estimated from Coomassie-stained gels (not shown). The digestion with thrombin yields a variant of SP-C with two additional amino acids at the N-terminus, namely glycine and proline (rGP/SP-C, Fig. 1, panels B and C). A deletion of these two residues rendered a chimera that was still susceptible to thrombin digestion, although with a significantly decreased efficiency. The two variants were analysed in parallel. During both, the initial purification step and the enzymatic cleavage of the nuclease moiety, the presence of detergents is essential in order to dissolve the cellular membrane, where a large part of the

fusion protein is localised, and, most important, to prevent aggregation of the rSP-C. Therefore, to maintain the chimera in an aqueous-detergent environment, we optimised the initial purification step and thrombin digestion. Experimental parameters investigated included: growth media, concentration of inducer (IPTG), size of culture batch, detergent type and concentration, number of freeze-thaw cycles and sonication, time and temperature (data not shown). The best results were obtained with an initial membrane pelleting, three freeze-thaw cycles followed by sonication and subsequent addition of the same volume of a buffered solution containing 1% lauroyl sarcosine detergent. This mixture was centrifuged and the fusion protein was purified from the supernatant via immobilized-metal affinity chromatography. Thrombin digestion was performed in the presence of the same detergent.

Extraction and purification

The highly hydrophobic nature of SP-C was used to separate rSP-C from the rest of the thrombin reaction mixture. Recombinant SP-Cs are expected to partition into the organic phase once the thrombin reaction mixture is subjected to organic extraction, whereas the hydrophilic nuclease moiety and thrombin would remain in the water-soluble phase. In this way, passing from a detergent-rich environment to an organic solution, the recombinant protein is likely to maintain its helicity, a crucial feature for its activity. Together with the rSP-C separation, through the organic extraction procedure we eliminate detergent molecules that partition into the organic solvent-water interface. In order to study the electrophoretic mobility of the recombinant polypeptides and roughly estimate the amount of protein, an SDS-PAGE analysis was performed. Fig. 2C shows aliquots of the organic phase bearing rSP-C (lane 2) and rGP/SP-C (lane 3) applied to the SDS-PAGE and compared with a known amount of the native protein, isolated from porcine lungs (lane 1). In agreement with previous SDS-PAGE analysis of SP-C [28, 29], native and recombinant peptides run as smeared bands with identical electrophoretic mobility, corresponding to an apparent molecular weight of ~6 kDa.

The final purification step involves separation of rSP-C from components that could have co-isolated in the organic phase, by using a lipophilic Sephadex LH-20 resin. This size exclusion chromatography resin is resistant to organic solvents and has been widely used for lipid-lipid and protein–lipid separations

[29, 30]. Organic extractions of the natural lung surfactants are subjected to this chromatographic procedure in order to separate the lipid fraction from surfactant proteins (SP-B and SP-C). Large-scale extractions yielded large volumes of extract, therefore a concentration step was necessary to minimise the sample volume loaded onto the chromatography column. In order to prevent protein self-aggregation and precipitation, we concentrated the rSP-C organic-extracted solution, in the presence of egg yolk phosphatidylcholine (PC) added at 1:5 protein:lipid ratio (w:w), in a rota-vaporizer. The concentrated extract was loaded onto a chloroform:methanol (2:1, v:v)-equilibrated LH-20 column and eluted using the same solvent system. The eluted fractions were measured for their absorbance at 240 and 280 nm (Fig. 3A). The rSP-C peak was later followed by the presence of a band at the expected migration position on SDS-PAGE of the pooled peak fractions (Fig. 3B).

The mass spectra of the purified peptides were in agreement with the expected molecular masses of 3783 and 3938 Daltons, for rSP-C and rGP/SP-C, respectively (Fig 4). The identity of these recombinant surfactant peptides was further confirmed by amino acid sequencing (data not shown) and their precise concentration was established by amino acid analysis. The final yield obtained for purified rSP-C was up to 0.5 mg per litre of bacterial culture.

Secondary structure determination

In order to test whether the rSP-Cs retained their helical conformation after the purification process, their secondary structure was evaluated by circular dichroism (CD) spectroscopy (Fig. 5). CD spectra in the far UV region, between 250 and 190nm, were used to estimate the contributions of different secondary structures to the conformation of the proteins. We studied the structure of recombinant and native SP-Cs in LPC micelles (Fig. 5A) and lipid (DPPC:POPG, 7:3 w/w) vesicles (Fig. 5B), to mimic the natural environment of SP-C. As seen in Fig. 5, all proteins exhibited CD spectral features consistent with a predominantly α-helical conformation, as judged by the presence of two peaks at 208 and 222 nm. The spectra of rGP/SP-C and the native SP-C were quantitatively comparable and similar to those reported in the literature [29, 31], while rSP-C exhibited a somewhat higher content of α-helix. To obtain a more detailed structural description, the CD spectra were analyzed by using the

CDPro suite of programs (SELCON3, CONTIN/LL and CDSSTR) [23, 24]. Similar results were obtained with the three programs using a reference set of proteins that includes several membrane proteins, which supports the reliability of the structure estimation. Table 1 summarises the relative proportions of different types of secondary structure obtained, although the values should be taken with caution because for small proteins flexibility may reduce the dichroic properties of the C=O bond. The increased α-helical content of rSP-C relative to the native SP-C is consistent with an N-terminal extension of the helical region of the protein, as observed in recent NMR structures of a recombinant SP-C mutant and a synthetic SP-C analogue, both bearing phenylalanine instead of cysteine residues [26, 32]. Thus, NOE-data supported an extended α-helix between residues 11 and 34 in the native protein [12], between residues 5 and 34 in the recombinant variant [26], or between residues 7 and 34 in the synthetic SP-C analogue [32].

Interfacial adsorption activity

Figure 6 illustrates the ability of recombinant SP-Cs to promote formation of surface-active interfacial films from lipid/protein suspensions, in comparison with the activity of native SP-C purified from porcine lungs. We choose the mixture of DPPC and POPG in proportion 7:3 (w:w) as a model system because it is a simple lipid mixture widely used to mimic roughly the lipid composition of LS. Increasing amounts of protein were added to the same amount of lipids in order to monitor the protein effect. As shown in Fig. 6, in the absence of proteins the lipids hardly adsorbed into the interface. In contrast, in the presence of native SP-C, film formation occurs rapidly depending on the protein concentration, to reach maximal equilibrium pressures that were slightly lower than the expected equilibrium pressure for an interface fully saturated with phospholipids during the time assayed. The amount and concentration of lipid or lipid/protein material used in these assays was chosen to detect potential quantitative differences in the surface activity of the different protein preparations. When amounts higher than 10 μg, or concentrations higher 1 mg/mL, of phospholipids were used, preparations containing native SP-C were able to reach equilibrium surface pressures close to 45-48 mN/m in a few minutes (not shown). Under the limiting amounts of material used in the present adsorption experiments, rSP-C and rGP/SP-C exhibited slightly, but

consistently, better activity, with the mixtures containing the highest amount of protein reaching equilibrium pressures of 48 mN/m in a few minutes. This suggests an optimal interaction of the recombinant peptides with the phospholipids. Optimal adsorption activities have been previously reported for other rSP-C forms expressed in bacteria using different strategies [15, 33, 34], but a quantitative comparison of their activity with respect to native SP-C under limiting conditions was not been performed.

Compression isotherms

To get a deeper insight into the apparently higher surface activity of the recombinant forms of SP-C produced here, we have analyzed in detail the π-A compression isotherms of DPPC/dipalmitoyl-phosphatidylglycerol (DPPG) (7:3, w/w) films in the absence and in the presence of different proportions of the native porcine SP-C or the recombinant rSP-C or rGP/SP-C forms. For these experiments we have used only disaturated lipids to preserve the main features of the isotherm, including the conspicuous liquid-expanded (LE)-to-liquid-condensed (LC) transition plateau, which occurs at around 10 mN/m in either DPPC or DPPG films at 25°C. Figure 7 shows that progressive proportions of the three proteins induce a shift of the isotherms towards increasing molecular area, as a result of the space taken by the protein molecules. Overall, recombinant proteins produce larger expansions of the isotherm than the native palmitoylated protein. This observation indicates that the structural changes introduced at the N-terminal segment of the recombinant forms with respect to native SP-C have significant consequences on the disposition and/or orientation of the protein at the interface. Lipid-protein films containing the three analyzed proteins produce isotherms that converge upon compression with the isotherms of the pure lipid monolayers at similar pressures of around 50 mN/m. This feature is consistent with a similar stability of the three proteins at the interface, from where they are squeezed-out at similar compression states. It is remarkable that the compression-driven exclusion of the three proteins is accompanied by the squeeze-out of some lipid molecules, as deduced from the

progressive shift of the isotherm to smaller areas at pressures higher than 50 mN/m [35]. The extent of such a shift increases progressively with the amount of protein present. Again, the recombinant SP-C forms produce larger shifts of the isotherms than native SP-C when lipid-protein films are compressed beyond 50 mN/m. In Figure 8 we have plotted the percentage of area removed from the interface in lipid/protein films compressed to 60 mN/m –a pressure well above the squeeze-out plateau of the isotherms- with respect to the area occupied by a pure lipid film at the same pressure, as a function of the molar proportion of protein in the films. Assuming that all protein molecules were squeezed-out from the interface upon compression up to 60 mN/m, the slope of these plots gives an estimate of the molar lipid/protein ratio excluded. In this case, each native SP-C molecule could be squeezed-out accompanied by 17±2 lipid molecules, while rGP/SP-C and rSP-C expel 25±2 and 53±3 phospholipids per protein molecule, respectively. Previous calculations had already suggested that the squeeze-out of native SP-C from compressed DPPC/DPPG films is accompanied by around 10 molecules of phospholipid per mol of protein [35]. Analysis by electron spin resonance of the stoichiometry of the interaction of SP-C with phospholipids in bilayers allows estimating that around 18 phospholipid molecules interacted with each SP-C molecule at saturation [36]. Although the numbers obtained from our isotherms may not be entirely correct without considering the protein segments that could still be inserted into the films at high pressures, the differences observed clearly indicate that the recombinant proteins bearing phenylalanines remove more lipid molecules per mol of protein than the native acylated protein. This feature can be correlated with a potentially higher affinity of the N-terminal segment of the recombinant proteins to maintain association with phospholipids, compared with the natural SP-C, which could be also responsible for the observed higher ability to promote interfacial adsorption of phospholipids.

DISCUSSION

The production of a recombinant SP-C has been a challenging task since the revealing of its importance in the late 80s. The extreme hydrophobicity and its tendency to aggregate make this protein very difficult to manipulate and many different approaches have been taken in order to circumvent these features. Synthetic peptides [12, 28, 37] and peptoids [38] mimicking the essential structural features of SP-C have been tested to enhance the surface-active properties of surfactant formulations composed of synthetic phospholipids, being in all cases less effective than the native (animal-extracted) SP-C. Concerning the recombinant expression of the protein, efforts have been made both for a eukaryotic baculovirus expression system [39] and using bacterial cultures [15, 33]. We decided to employ a prokaryotic system in order to obtain SP-C in large quantities, since bacteria are to date the best large-scale protein-producing organism. Compared to the previously published rSP-C expressed in *Escherichia coli* as a fusion to the N-terminal portion of bacterial chloramphenicol acetyl transferase, our strategy results more protective of the α-helical conformation necessary for optimal surface activity. This is probably because our method does not involve resolubilisation of the protein from inclusion bodies and exposure to polar solvents during RP-HPLC [15]. In addition, the use of the nuclease A (SN) as a fusion partner provides several advantages (reviewed in [40]) for high-level heterologous expression of integral membrane proteins at full-length. For instance, the three-dimensional structure of SN shows numerous basic residues exposed to the solvent [41], therefore precluding aggregation of fusion proteins and becoming a versatile system for membrane protein over-expression.

We expressed a modified sequence of human SP-C, which bears two phenylalanines instead of palmitoylated cysteines, inspired by the native sequence of SP-C from dog and mink surfactants. The role of palmitoyl chains in SP-C is still under debate. It is evident that SP-C maintains its function in the presence of one phenylalanine instead of a palmitic chain, as in the case of dog and mink SP-Cs, and the possibility remains that this particular amino acid can mimic at least partially the role of palmitic chains. The fact that this mutant variant has been successfully used both *in vitro* and in animal experiments supports the idea that it could be a functional equivalent to the natural SP-C

[18]. Our interfacial adsorption measurements together with compression isotherms suggest that, at least for the bilayer-monolayer conversions, these modified SP-C versions could have superior properties than the native SP-C purified from animal lungs. The compression isotherms of phospholipid films containing increasing proportions of protein indicate that the recombinant versions of SP-C expel more phospholipid molecules from the interface than the native SP-C, when the lipid-peptide films are compressed beyond the collapse pressure of the protein. These results suggest that the aromatic residues included into the N-terminal region of the recombinant forms may either increase the affinity of the protein to associate with phospholipids or promote a more appropriate conformation of this region to accommodate a higher number of lipid molecules. Additionally, our production procedure, which maintains the structural integrity of the recombinant SP-C forms, may preserve better its activity in quantitative terms, compared to the delipidation and purification procedure of native SP-C from animal lungs. The role of SP-C *in vivo* has been mostly discussed in connection to the protein-mediated stability of the interfacial surface active film at end-expiration, and not much as promoting interfacial adsorption, an activity which is better performed by SP-B [1]. However, there are strong indications suggesting that a Phe-containing rSP-C version has enough surface activity to be used as the only protein component of a therapeutic surfactant preparation [15, 17]. The N-terminal segment of SP-C has been revealed as a very dynamic structural motif [42], which is able to interact and perturb the phospholipid packing of phospholipid bilayers and monolayers [43]. Acylation of this segment may be especially important to maintain association of the lipid/protein structures squeezed-out from the compressed films at end-expiration with the interface [44]. In native surfactant complexes, efficient bilayer-monolayer (during adsorption) and monolayer-bilayer (during compression) conversions are ensured by the presence of SP-B. A high affinity of the protein to associate with the interface has been recently shown to be a crucial determinant for the SP-B activity, this association being critically dependent on the presence of aromatic residues at the N-terminal segment of the protein [45]. It is thus reasonable to expect that in the absence of SP-B, SP-C variants with enhanced affinity for the interface may show enough activity to promote bilayer-monolayer transitions and serve as the only additive needed for efficient therapeutic surfactant formulations. Further biophysical

studies of the surface behaviour of these and other variants are ongoing and will allow the establishment of their potential as SP-C (or SP-B) mimics, as well as the determination of critical structure-function motifs.

The protocol presented here is a combination of a common strategy for over-expression of integral membrane proteins and the organic extraction used for native SP-C isolation from animal pulmonary lavages. This purification scheme may also be used for obtaining other highly hydrophobic proteins. The flexibility for creating mutants of the SP-C, aiming to specific structural or functional studies, will greatly improve our understanding of the role of this protein in the LS, as well as optimizing current replacement preparations for respiratory diseases. Further biophysical studies are in progress in order to explore the full potential of these and new SP-C variants for surfactant therapy.

FIGURE LEGENDS

Figure 1. Comparison between sequences of SP- C from different origins and sequence and schematic organisation of the rSP-C variants. (**A**) Native SP-C sequence alignment highlighting the Cys residues in blue and Phe residues in red. In the porcine sequence the helical region from the NMR structure [46] is boxed. (**B**) Amino acid sequences of the rSP-Cs produced in this work. (**C**) Schematic organisation of the chimerical protein expressed in bacteria.

Figure 2. SDS-PAGE analyses of over-expressed chimeric proteins, thrombin digestion and organically extracted native and rSP-C. (**A**) Representative Coomassie blue-stained 12% SDS-PAGE of aliquots taken during the purification of chimeric proteins. Lane 1, molecular markers; lane 2, whole cell extract; lane 3, supernatant of bacterial lysate; lane 4, nickel column flow through; lane 5, eluate form nickel column. The SN/SP-C chimera is marked with an arrowhead to the right. (**B**) Coomassie blue-stained 15% acrylamide tricine gel loaded with an aliquot of the SN/SP-C chimera after (lane 2) and before thrombin digestion (lane 3). (**C**) Native (lane 1) and rSP-Cs (rGP/SP-C, lane 2 and rSP-C, lane 3) organically extracted were resolved on a 16% acrylamide SDS-PAGE stained with Coomassie blue.

Figure 3. Purification of rSP-Cs by Sephadex LH20. (**A**) LH 20 chromatographic analysis of organic extracts. Absorbance at 240 (black line) and 280 (grey line) nm was monitored. (**B**) Comassie blue stained SDS-PAGE analysis of increasing aliquots of the initial peak (lanes 1- 3) and two aliquots of the distal peak (lanes 4 and 5). Lane 6 was loaded with 15 μg of native SP-C. Size markers are indicated on the left.

Figure 4. Mass spectra of rSP-Cs purified on Sephadex LH20.
The content of the first peak in the chromatograms of Figure 3 were analysed by MALDI-TOF-MS as described in Materials and Methods. rGP/SP-C and rSP-C spectra are shown in panels A and B, respectively.

Figure 5. Circular dichroism spectra of native and recombinant SP-Cs reconstituted in surfactant-mimetic media. Native SP-C (solid line), rGP/SP-C (dotted line) and rSP-C (dashed line) spectra were recorded in the presence of LPC micelles (panel A) and DPPC:POPG (7:3) vesicles (panel B). In all samples 25 μM of proteins were analysed (see Materials and Methods).

Figure 6. Interfacial adsorption activity. The panels show the π-t adsorption isotherms for interfacial film formation after deposition of 10 μL of 10 mg/mL DPPC:POPG (7:3) suspension in the absence or presence of 2, 5 or 10% native SP-C (panel A), rGP/SP-C (panel B) or rSP-C (panel C).

Figure 7. Compression π-A isotherms of lipid and lipid/protein films. Films of DPPC/DPPG (7:3, w/w) in the absence (bold lines) or presence of 2% (circles), 5% (triangles) or 10% (squares) of the SP-C variants were formed by spreading lipid or lipid-protein mixtures in chloroform/methanol 3:1 (Volvo) on top of a buffered sub-phase Tris 5mM pH 7, containing 150 mM NaCl. The compression rate was 65 cm²/min and the temperature was 25ºC.

Figure 8. Effect of native and recombinant forms of SP-C on the removal of lipid/protein material from highly compressed interfacial films. Percent of DPPC/DPPG film area at 60 mN/m reduced by the presence of native porcine SP-C (closed circles), rGP/SP-C (open circles) or rSP-C (open squares), has been plotted versus the proportion of protein in the films (mol/mol).

REFERENCES

[1] J. Perez-Gil, Molecular interactions in pulmonary surfactant films, Biol Neonate 81 Suppl 1 (2002) 6-15.

[2] M. van Eijk, C.G. De Haas and H.P. Haagsman, Quantitative analysis of pulmonary surfactant proteins B and C, Anal Biochem 232 (1995) 231-7.

[3] W. Seeger, C. Grube, A. Gunther and R. Schmidt, Surfactant inhibition by plasma proteins: differential sensitivity of various surfactant preparations, Eur Respir J 6 (1993) 971-7.

[4] R.H. Pfister and R.F. Soll, New synthetic surfactants: the next generation?, Biol Neonate 87 (2005) 338-44.

[5] J. Johansson, M. Gustafsson, M. Palmblad, S. Zaltash, B. Robertson and T. Curstedt, Synthetic surfactant protein analogues, Biol Neonate 74 Suppl 1 (1998) 9-14.

[6] T.E. Weaver and J.J. Conkright, Function of surfactant proteins B and C, Annu Rev Physiol 63 (2001) 555-78.

[7] S. Schurch, R. Qanbar, H. Bachofen and F. Possmayer, The surface-associated surfactant reservoir in the alveolar lining, Biol Neonate 67 Suppl 1 (1995) 61-76.

[8] M. Gustafsson, J. Thyberg, J. Naslund, E. Eliasson and J. Johansson, Amyloid fibril formation by pulmonary surfactant protein C, FEBS Lett 464 (1999) 138-42.

[9] M.F. Beers and S. Mulugeta, Surfactant protein C biosynthesis and its emerging role in conformational lung disease, Annu Rev Physiol 67 (2005) 663-96.

[10] S.-C. Li and C.M. Deber, A measure of helical propensity for amino acids in membrane environments, Nat Struct Biol 1 (1994) 368-373.

[11] M. Gustafsson, W.J. Griffiths, E. Furusjo and J. Johansson, The palmitoyl groups of lung surfactant protein C reduce unfolding into a fibrillogenic intermediate, J Mol Biol 310 (2001) 937-50.

[12] J. Johansson, G. Nilsson, R. Stromberg, B. Robertson, H. Jornvall and T. Curstedt, Secondary structure and biophysical activity of synthetic analogues of the pulmonary surfactant polypeptide SP-C, Biochem J 307 (1995) 535-41.

[13] H. Kiefer, In vitro folding of alpha-helical membrane proteins, Biochim Biophys Acta 1610 (2003) 57-62.

[14] A.G. Therien, M. Glibowicka and C.M. Deber, Expression and purification of two hydrophobic double-spanning membrane proteins derived from the cystic fibrosis transmembrane conductance regulator, Protein Expr Purif 25 (2002) 81-6.

[15] S. Hawgood, A. Ogawa, K. Yukitake, M. Schlueter, C. Brown, T. White, D. Buckley, D. Lesikar and B. Benson, Lung function in premature rabbits treated with recombinant human surfactant protein-C, Am J Respir Crit Care Med 154 (1996) 484-90.

[16] R.G. Spragg, J.F. Lewis, W. Wurst, D. Hafner, R.P. Baughman, M.D. Wewers and J.J. Marsh, Treatment of acute respiratory distress syndrome with recombinant surfactant protein C surfactant, Am J Respir Crit Care Med 167 (2003) 1562-6.

[17] R.G. Spragg, J.F. Lewis, H.D. Walmrath, J. Johannigman, G. Bellingan, P.F. Laterre, M.C. Witte, G.A. Richards, G. Rippin, F. Rathgeb, D. Hafner, F.J. Taut and W. Seeger, Effect of recombinant surfactant protein C-based surfactant on the acute respiratory distress syndrome, N Engl J Med 351 (2004) 884-92.

[18] A.J. Davis, A.H. Jobe, D. Hafner and M. Ikegami, Lung function in premature lambs and rabbits treated with a recombinant SP-C surfactant, Am J Respir Crit Care Med 157 (1998) 553-9.

[19] I. Mingarro, P. Whitley, M.A. Lemmon and G. von Heijne, Ala-insertion scanning mutagenesis of the glycophorin A transmembrane helix. A rapid way to map helix-helix interactions in integral membrane proteins, Protein Sci. 5 (1996) 1339-1341.

[20] M. Orzaez, E. Perez-Paya and I. Mingarro, Influence of the C-terminus of the glycophorin A transmembrane fragment on the dimerization process, Protein Sci 9 (2000) 1246-53.

[21] D.J. Hakes and J.E. Dixon, New vectors for high level expression of recombinant proteins in bacteria, Anal Biochem 202 (1992) 293-8.

[22] E.G. Bligh and W.J. Dyer, A rapid method of total lipid extraction and purification, Can J Biochem Physiol 37 (1959) 911-7.

[23] N. Sreerama and R.W. Woody, On the analysis of membrane protein circular dichroism spectra, Protein Sci 13 (2004) 100-12.

[24] N. Sreerama and R.W. Woody, Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set, Anal Biochem 287 (2000) 252-60.

[25] A.G. Serrano, A. Cruz, K. Rodriguez-Capote, F. Possmayer and J. Perez-Gil, Intrinsic structural and functional determinants within the amino acid sequence of mature pulmonary surfactant protein SP-B, Biochemistry 44 (2005) 417-30.

[26] B. Luy, A. Diener, R.P. Hummel, E. Sturm, W.R. Ulrich and C. Griesinger, Structure and potential C-terminal dimerization of a recombinant mutant of surfactant-associated protein C in chloroform/methanol, Eur J Biochem 271 (2004) 2076-85.

[27] W.C. Wimley and S.H. White, Experimentally determined hydrophobicity scale for proteins at membrane interfaces, Nature Struct Biology 3 (1996) 842-848.

[28] M. Palmblad, J. Johansson, B. Robertson and T. Curstedt, Biophysical activity of an artificial surfactant containing an analogue of surfactant protein (SP)-C and native SP-B, Biochem J 339 (Pt 2) (1999) 381-6.

[29] J. Perez-Gil, A. Cruz and C. Casals, Solubility of hydrophobic surfactant proteins in organic solvent/water mixtures. Structural studies on SP-B and SP-C in aqueous organic solvents and lipids, Biochim Biophys Acta 1168 (1993) 261-70.

[30] T. Curstedt, H. Jornvall, B. Robertson, T. Bergman and P. Berggren, Two hydrophobic low-molecular-mass protein fractions of pulmonary surfactant. Characterization and biophysical activity, Eur J Biochem 168 (1987) 255-62.

[31] A. Cruz, C. Casals and J. Perez-Gil, Conformational flexibility of pulmonary surfactant proteins SP-B and SP-C, studied in aqueous organic solvents, Biochim Biophys Acta 1255 (1995) 68-76.

[32] J. Li, E. Liepinsh, A. Almlen, J. Thyberg, T. Curstedt, H. Jornvall and J. Johansson, Structure and influence on stability and activity of the N-terminal propeptide part of lung surfactant protein C, FEBS Journal 273 (2006) 926-935.

[33] J.T. Stults, P.R. Griffin, D.D. Lesikar, A. Naidu, B. Moffat and B.J. Benson, Lung surfactant protein SP-C from human, bovine, and canine sources contains palmityl cysteine thioester linkages, Am J Physiol 261 (1991) L118-25.

[34] L.A. Creuwels, R.A. Demel, L.M. van Golde, B.J. Benson and H.P. Haagsman, Effect of acylation on structure and function of surfactant protein C at the air-liquid interface, J Biol Chem 268 (1993) 26752-8.

[35] S. Taneva and K.M. Keough, Pulmonary surfactant proteins SP-B and SP-C in spread monolayers at the air-water interface: II. Monolayers of pulmonary surfactant protein SP-C and phospholipids, Biophys J 66 (1994) 1149-57.

[36] J. Perez-Gil, C. Casals and D. Marsh, Interactions of hydrophobic lung surfactant proteins SP-B and SP-C with dipalmitoylphosphatidylcholine and dipalmitoylphosphatidylglycerol bilayers studied by electron spin resonance spectroscopy, Biochemistry 34 (1995) 3964-71.

[37] G. Nilsson, M. Gustafsson, G. Vandenbussche, E. Veldhuizen, W.J. Griffiths, J. Sjovall, H.P. Haagsman, J.M. Ruysschaert, B. Robertson, T. Curstedt and J. Johansson, Synthetic peptide-containing surfactants--evaluation of transmembrane versus amphipathic helices and surfactant protein C poly-valyl to poly-leucyl substitution, Eur J Biochem 255 (1998) 116-24.

[38] C.W. Wu, S.L. Seurynck, K.Y. Lee and A.E. Barron, Helical peptoid mimics of lung surfactant protein C, Chem Biol 10 (2003) 1057-63.

[39] E.J. Veldhuizen, J.J. Batenburg, G. Vandenbussche, G. Putz, L.M. van Golde and H.P. Haagsman, Production of surfactant protein C in the baculovirus expression system: the information required for correct folding and palmitoylation of SP-C is contained within the mature sequence, Biochim Biophys Acta 1416 (1999) 295-308.

[40] R. Laage and D. Langosch, Strategies for prokaryotic expression of eukaryotic membrane proteins, Traffic 2 (2001) 99-104.

[41] T.R. Hynes and R.O. Fox, The crystal structure of staphylococcal nuclease refined at 1.7 A resolution, Proteins 10 (1991) 92-105.

[42] I. Plasencia, L. Rivas, K.M. Keough, D. Marsh and J. Perez-Gil, The N-terminal segment of pulmonary surfactant lipopeptide SP-C has intrinsic propensity to interact with and perturb phospholipid bilayers, Biochem J 377 (2004) 183-93.

[43] I. Plasencia, K.M. Keough and J. Perez-Gil, Interaction of the N-terminal segment of pulmonary surfactant protein SP-C with interfacial phospholipid films, Biochim Biophys Acta 1713 (2005) 118-28.

[44] X. Bi, C.R. Flach, J. Perez-Gil, I. Plasencia, D. Andreu, E. Oliveira and R. Mendelsohn, Secondary structure and lipid interactions of the N-terminal segment of pulmonary surfactant SP-C in Langmuir films: IR reflection-absorption spectroscopy and surface pressure studies, Biochemistry 41 (2002) 8385-95.

[45] A.G. Serrano, M. Ryan, T.E. Weaver and J. Perez-Gil, Critical Structure-Function Determinants within the N-Terminal Region of Pulmonary Surfactant Protein SP-B, Biophys J 90 (2006) 238-49.

[46] J. Johansson, T. Szyperski, T. Curstedt and K. Wuthrich, The NMR structure of the pulmonary surfactant-associated polypeptide SP-C in an apolar solvent contains a valyl-rich alpha-helix, Biochemistry 33 (1994) 6015-23.

A

```
     1          10           20           30
    FGIPCCPVHL KRLLIVVVVV VLIVVVIVGA LLMGL human
    LRIPCCPVNL KRLLVVVVVV VLVVVVIVGA LLMGL porcine
    FGIPCCPVHL KRLLIVVVVV VLVVVVIVGA LLMGL rabbit
    FGIPCCPVHL KRLLIVVVVV VLVVVVIVGA LLMGL monkey
    FRIPCCPVHL KRLLIVVVVV VLVVVVIVGA LLMGL rat
    FRIPCCPVHL KRLLIVVVVV VLVVVVIVGA LLMGL mouse
    LRIPCCPVNI KRLLIVVVVV VLVVVVIVGA LLMGL sheep
    LGIPCFPSSL KRLLIIVVVI VLVVVVIVGA LLMGL dog
    LGIPCFPSSL KRLLIIVVVI VLVVVVIVGA LLMGL mink
```
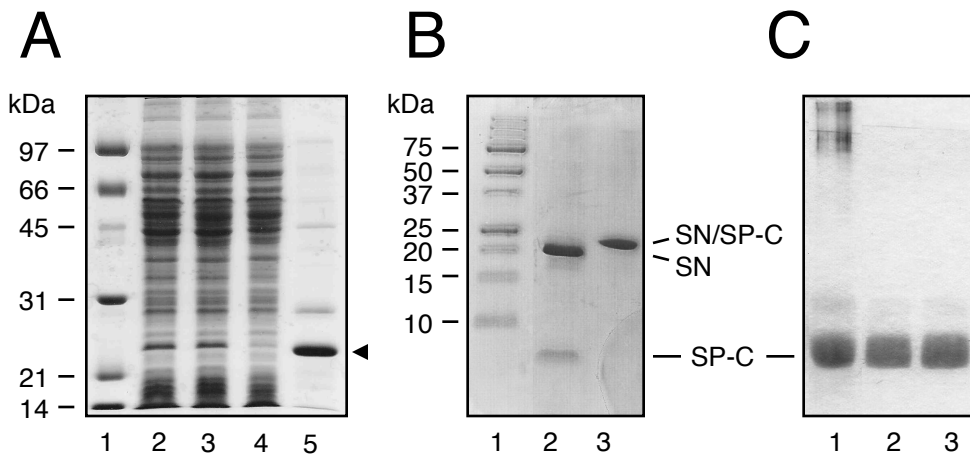
B

```
    FGIPFFPVHL KRLLIVVVVV VLIVVVIVGA LLMGL rSP-C
  GPFGIPFFPVHL KRLLIVVVVV VLIVVVIVGA LLMGL rGP/SP-C
```

C

H-tag	SN	GGGGG	LVPR↓GP	rSP-C
		Gly-kinker	Thrombin site	

A

kDa					
97 —					
66 —					
45 —					
31 —					
21 —					
14 —					

◄

1 2 3 4 5

B

kDa			
75 —			
50 —			
37 —			
25 —			— SN/SP-C
20 —			— SN
15 —			
10 —			
			— SP-C —

1 2 3

C

1 2 3

A

B