

RESEARCH ARTICLE

# Conformational Stability of the NH<sub>2</sub>-Terminal Propeptide of the Precursor of Pulmonary Surfactant Protein SP-B

Ángeles Bañares-Hidalgo, Jesús Pérez-Gil\*, Pilar Estrada

Departamento de Bioquímica y Biología Molecular I, Facultad de Biología, Universidad Complutense, Ciudad Universitaria, 28040, Madrid, Spain

\* [jperezgil@bio.ucm.es](mailto:jperezgil@bio.ucm.es)



**OPEN ACCESS**

**Citation:** Bañares-Hidalgo Á, Pérez-Gil J, Estrada P (2016) Conformational Stability of the NH<sub>2</sub>-Terminal Propeptide of the Precursor of Pulmonary Surfactant Protein SP-B. PLoS ONE 11(7): e0158430. doi:10.1371/journal.pone.0158430

**Editor:** Sabato D'Auria, CNR, ITALY

**Received:** April 10, 2016

**Accepted:** June 15, 2016

**Published:** July 5, 2016

**Copyright:** © 2016 Bañares-Hidalgo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This research has been supported by grants from the Spanish Ministry of Economy and Competition (BIO2012-30733, BIO2015-67930-R), and the Regional Government of Madrid (P2013/MIT-2807).

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

Assembly of pulmonary surfactant lipid-protein complexes depends on conformational changes coupled with proteolytic maturation of proSP-B, the precursor of pulmonary surfactant protein B (SP-B), along the surfactant biogenesis pathway in pneumocytes. Conformational destabilization of the N-terminal propeptide of proSP-B (SP-B<sub>N</sub>) triggers exposure of the mature SP-B domain for insertion into surfactant lipids. We have studied the conformational stability during GdmCl- or urea-promoted unfolding of SP-B<sub>N</sub> with trp fluorescence and circular dichroism spectroscopies. Binding of the intermediate states to bis-ANS suggests their molten globule-like character.  $\Delta G^0_{H_2O}$  was  $\sim 12.7 \text{ kJ}\cdot\text{mol}^{-1}$  either with urea or GdmCl. None of the thermal transitions of SP-B<sub>N</sub> detected by CD correspond to protein unfolding. Differential scanning calorimetry of SP-B<sub>N</sub> evidenced two endothermic peaks involved in oligomer dissociation as confirmed with 2 M urea. Ionic strength was relevant since at 150 mM NaCl, the process originating the endotherm at the highest temperature was irreversible ( $T_{m2} = 108.5^\circ\text{C}$ ) with an activation energy of  $703.8 \text{ kJ}\cdot\text{mol}^{-1}$ . At 500 mM NaCl the process became reversible ( $T_{m2} = 114.4^\circ\text{C}$ ) and data were fitted to the Non-two States model with two subpeaks. No free thiols in the propeptide could be titrated by DTNB with or without 5.7 M GdmCl, indicating disulfide bonds establishment.

## Introduction

The function of a protein depends on its ability to adopt a specific structure. Proteins can form partially folded, collapsed states resembling the intermediate states along the protein folding pathway, and this is important in understanding the mechanisms of protein folding [1]. Those states have been often considered as molten globule structures, that is, partially folded proteins with native-like secondary structure, but lacking the extensive, specific side-chain packing interactions of the native structure [2]. In addition to the molten globule state, two kinds of intermediate states called the pre-molten globule state and the highly-ordered molten globule state may exist during protein folding and refolding [3]. The protein in the pre-molten globule state is half-way between the molten globule and the unfolded state, it is less compact than the

molten globule but is still more compact than the random coil [4] whereas the highly-ordered molten globule has more structure and is much more similar to the native state [5]. Intermediate states in the folding-unfolding processes may be revealed by mild denaturing conditions, such as changes in pressure, temperature, pH or upon addition of chaotropes (i.e. urea or guanidinium chloride) [6].

We are interested in the study of proteins involved in pulmonary surfactant biogenesis and function from a structural and functional point of view. The stability of the respiratory surface at the lung is maintained by lipids and proteins synthesized and secreted by type II pneumocytes into the alveolar spaces, to form a film at the air-water interface known as pulmonary surfactant. The surfactant protein B (SP-B) is strictly necessary to promote the formation of this film in order to prevent collapse during deflation [7] and is produced from a precursor (proSP-B) containing two propeptides flanking the mature protein at its NH<sub>2</sub>- and COOH-termini respectively (S1 Fig). The 177 amino acid propeptide flanking the NH<sub>2</sub> terminus of the mature protein is necessary and sufficient for targeting processing and assembly of the hydrophobic mature SP-B into surfactant complexes *in vivo* [8]. Since deletion of this propeptide results in accumulation of SP-B within the endoplasmic reticulum, a role as intramolecular chaperone for SP-B has been suggested [9]. Moreover, proSP-B is an homologous protein to prosaposin, the precursor protein of four saposins (*sphingolipid activator proteins*) generated through cleavage of the precursor prosaposin [10]. Prosaposin contains four Saposin B (SAPB) domains and two Saposin A (SAPA) modules, one at its NH<sub>2</sub> and another at its COOH terminus whereas proSP-B contains three SAPB modules (in mature SP-B and in both flanking propeptides) plus an additional SAPA module in the NH<sub>2</sub>-terminal propeptide (S1 Fig). Other members of the saposin-like family proteins (SAPLIPs), such as NK-lysin or the amoebapore have antipathogenic activities [11]. As occurs with prosaposin, it has been proposed that the propeptides of proSP-B could have additional functions in the lung once released from the mature protein. This has been confirmed after isolation of the saposin module from the NH<sub>2</sub>-terminal propeptide of proSP-B from bronchoalveolar lavage fluid and its characterization as participating in the host defence in rat lung [12]. The structural features of the NH<sub>2</sub>-terminal propeptide of proSP-B (SP-B<sub>N</sub>) are therefore important not only to understand SP-B and surfactant biogenesis but also to unravel the intrinsic structure-function determinants of a protein that may have its own role to protect the pulmonary surface from the entrance of pathogens.

In this context, we have produced a recombinant form of human SP-B<sub>N</sub>, soluble and containing the microbicide module [13] and the conditions that favoured or hindered its aggregation have been studied [14]. Also we have observed that acidic pH triggers, in the propeptide, the formation of a coiled coil structure [15]. In the present work, we extend the structural dissection of the protein by analyzing the thermodynamical stability of the propeptide, and the detection of their folding intermediates, if any, in the presence of chaotropes. Results obtained employing far-UV circular dichroism, intrinsic and extrinsic fluorescence spectroscopy and differential scanning calorimetry are discussed regarding the conformational stability of the protein and its oligomeric state determined through sedimentation velocity. The lack of free cysteines in the recombinant protein indicates that disulfide bonds have been formed in the propeptide. This is relevant since proteins of the SAPLIP family contain conserved cysteines, which are involved in intramolecular disulfide bonds responsible for their high thermal stability and resistance to proteolytic degradation [16, 17]. Still, our results reveal a structural complexity in SP-B<sub>N</sub> that should be taken into account to understand the mechanisms triggering the liberation and assembly of surface active and antipathogenic modules.

## Materials and Methods

### Production and Purification of SP-B<sub>N</sub>

The human wild-type propeptide was expressed in *E. coli* as a fusion protein (MBP-SP-B<sub>N</sub>) with the Maltose Binding Protein (MBP) and purified after cleaving the fusion with factor Xa as described [13]. The fractions containing the purified SP-B<sub>N</sub> in 20 mM Tris-HCl buffer pH 7, 500 mM NaCl were dialyzed towards the buffer needed in subsequent experiments. The mean net charge of the propeptide is defined as the net charge at pH 7 [18] divided by the total number of residues. The hydrophobicity at each sequence position was calculated by the Kite and Doolittle scale [19] using a window size of 5 amino acids and normalized to a scale of 0–1. The mean hydrophobicity is defined as the sum of the normalized hydrophobicities of all residues divided by the number of residues of the propeptide.

### Far-UV Circular Dichroism Experiments

Circular dichroism spectra of SP-B<sub>N</sub> were recorded at 25°C in a Jasco J-715 spectropolarimeter using thermostated quartz cells of 0.1-cm path length, at 50 nm·min<sup>-1</sup> (1 s response time) for the far-UV (250–195 nm) spectral range, each spectrum being the accumulation of 5 scans. The spectra were obtained in 200 μL of 5mM acetate, 5 mM MES, 5 mM Tris-HCl buffer (AMT buffer) 150 mM NaCl pH 7 at 0.115 mg·mL<sup>-1</sup> protein. Mean residue molar ellipticities [θ] were calculated from the measured ellipticity taking into account the protein concentration, the molecular weight of SP-B<sub>N</sub> (19,902 Da, DNASTAR program) and the number of amino acids per molecule (177). Estimations of the secondary structure content from the CD spectra were performed by using the CDPro suite program and the α-helix and β-sheet contents were calculated using three different methods, CONTIN/LL, SELCON3 and CDSSTR employing their mean value [20]. CD spectra in the presence of chaotropes were obtained after preincubating the samples with the denaturant for 1 h at 25°C. Samples at the corresponding chaotrope concentration, 0.8–8 M GdmCl and 0.5–7.5 M urea, were prepared from stock solutions of 9.75 M GdmCl (Sigma) and 9.87 M urea (Sigma) respectively. Control samples in the absence of protein were used to subtract a baseline from samples with protein. The reversibility of chemical unfolding was analyzed by dialyzing extensively samples with the highest chaotrope concentration towards 5 mM AMT buffer, 150 mM NaCl pH 7 at 4°C and by recording the spectra thereafter.

Temperature studies were carried out with 0.115 mg·mL<sup>-1</sup> protein in 5 mM AMT buffer pH 7 with 150 or 500 mM NaCl. The temperature dependence of the CD signal was determined by heating the samples from 25 to 85°C at 30°C·h<sup>-1</sup> and collecting the ellipticity at 208 nm every 0.2°C. After the T ramp the sample was cooled back to 25°C and the CD spectrum was recorded again.

### Intrinsic and Extrinsic Fluorescence Studies

The intrinsic fluorescence emission spectra of the propeptide were recorded at 25°C in a SLM-Aminco AB2 spectrofluorimeter using a 1-cm quartz cell with excitation (290 nm) and emission slits set at 4 nm and scan speed of 2 nm·s<sup>-1</sup>. Samples contained 0.09 mg·mL<sup>-1</sup> protein in 5 mM AMT buffer, 150 mM NaCl pH 7. The spectra of the protein in the presence of chaotropes, 0.8–6.6 M GdmCl or 0.5–7.5 M urea were recorded after preincubating the samples for 4 h at 25°C. Spectra were also recorded of samples dialyzed as described above to check the reversibility of protein unfolding.

Extrinsic fluorescence of 11 μM bis-ANS probe (4,4'-bis-1-phenylamine-8-naphthalene sulfonate from Thermo Fisher Scientific) was determined as follows. Samples containing chaotropes

were prepared in a final volume of 200  $\mu\text{L}$  by adding 11  $\mu\text{L}$  of the probe from a 200  $\mu\text{M}$  solution in methanol to 29  $\mu\text{g}$  protein in 5 mM AMT buffer, 150 mM NaCl pH 7 with the corresponding amount of GdmCl or urea. After incubation at 37°C for 5 min, the emission spectra (400–625 nm) were recorded in the spectrofluorimeter connected to a water-bath thermostated at 37°C. The excitation wavelength was 395 nm and samples without protein were used as blank. Scan speed and slit widths were as described above.

## Analysis of CD and Fluorescence Emission Data

The transition curves obtained by CD and fluorescence spectroscopy from the propeptide unfolding experiments were analyzed according to the following equations:

$$Y_{obs} = \frac{Y_N + Y_U \exp\left(\frac{-\Delta G_{H_2O}^0 + mD}{RT}\right)}{1 + \exp\left(\frac{-\Delta G_{H_2O}^0 + mD}{RT}\right)} \quad (1)$$

$$Y_{obs} = \frac{Y_N + a}{1 + \exp\left(\frac{D_{1/2} - D}{b}\right)} \quad (2)$$

$$\Delta G^0 = -RT \ln\left(\frac{Y_N - Y_{obs}}{Y_{obs} - Y_U}\right) \quad (3)$$

$$\Delta G^0 = \Delta H^0 - T\Delta S^0 \quad (4)$$

In the unfolding process by chaotropes, in Eqs (1) and (2):  $Y_{obs}$  is the observed parameter ( $[\theta]^{220}$  or  $FI_{350}$ ) at each denaturant concentration,  $Y_N$  and  $Y_U$  are the parameter values in native and unfolded conditions respectively,  $D$  is the denaturant concentration ( $\text{mol}\cdot\text{L}^{-1}$ ),  $\Delta G_{H_2O}^0$  is the change of free energy in the absence of denaturant ( $\text{kJ}\cdot\text{mol}^{-1}$ ),  $m$  is a measure of the dependence of the free energy on the denaturant concentration ( $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{M}^{-1}$ ),  $R$  is the gas constant ( $8.314 \text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ ) and  $T$  is the absolute temperature (K). The midpoint of the unfolding curve ( $D_{1/2}$ ) is the denaturant concentration at which half of the protein is unfolded and can be calculated as  $D_{1/2} = \Delta G_{H_2O}^0/m$  since  $\Delta G_{H_2O}^0$  and  $m$  are obtained from Eq (1). Alternatively,  $D_{1/2}$  can be determined directly from fitting the same data to Eq (2) being  $b$  and  $a$  constants. In temperature analysis,  $D$  and  $D_{1/2}$  in Eq (2) were substituted by  $T$  and the transition temperatures respectively (all of them in°C for easy recognition) and  $Y_{obs}$  was  $[\theta]^{208}$  in Eqs (2) and (3). The change of free energy,  $\Delta G^0$  was calculated for each temperature (K) according to Eq (3). The enthalpy change,  $\Delta H^0$  ( $\text{kJ}\cdot\text{mol}^{-1}$ ) and the entropy change,  $\Delta S^0$  ( $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ ) were the intercept and the slope, respectively, in Eq (4).

## Analytical Ultracentrifugation

Hydrodynamic studies of 0.15  $\text{mg}\cdot\text{mL}^{-1}$  SP-B<sub>N</sub> in 5 mM Tris-HCl pH 7.0 containing either, 0, 150 or 500 mM NaCl were performed at 20°C and 48,000 rpm in an Optima XL-1 (Beckman-Coulter Inc) analytical ultracentrifuge equipped with UV-visible optics. The partial specific volume  $\bar{V}$  of SP-B<sub>N</sub> was 0.73  $\text{mL}\cdot\text{g}^{-1}$  estimated from its amino acid composition with the program SEDNTERP, version 1.09 (retrieved from RASMB server) [21]; the solvent density  $\rho$  was 1.000  $\text{g}\cdot\text{mL}^{-1}$  and the solvent viscosity  $\eta$  was 1.002 cpoise in the absence of salt, 1.005  $\text{g}\cdot\text{mL}^{-1}$  and 1.017 cpoise with 150 mM NaCl and 1.019  $\text{g}\cdot\text{mL}^{-1}$  and 1.049 cpoise with 500 mM NaCl respectively, estimated with the same program.

## Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) experiments were performed on a VP-DSC (Micro-Cal) differential scanning microcalorimeter with a cell volume of 514.9  $\mu\text{L}$ . The propeptide ( $0.2\text{--}0.6\text{ mg}\cdot\text{mL}^{-1}$ ) in 5 mM Tris-HCl buffer pH 7, 150 or 500 mM NaCl and the corresponding buffer were degassed for 3 min at room temperature in a chamber under vacuum and gentle stirring and then loaded into the sample and reference cells where overpressure was kept to prevent degassing. The measurements were taken every  $0.1^\circ\text{C}$  and the scan rate was  $60^\circ\text{C}\cdot\text{h}^{-1}$ . DSC scans began at  $20^\circ\text{C}$  and were over at the highest temperature possible in the VP-DSC ( $\sim 120^\circ\text{C}$ ). Second scans were obtained by reheating samples after cooling them for 20 min upon the completion of the first scan. The apparent  $C_p$  profiles were obtained by subtracting the instrumental baseline (obtained with buffer in both cells) from the experimental thermograms. Then, the thermograms were normalized for protein concentration based on a monomer of 19,902 Da and the pre- and post-transition baselines were subtracted.

The reversible transitions were subjected to thermodynamic deconvolution analysis and the thermogram peak of each successive scan was fitted to the Non Two-State model with two peaks according to Eq (5) of the DSC tutorial guide with the Origin Microcal software, as this model accounted for the best fitting of the data (the lowest  $\chi^2 / \text{DoF}$  (Degree of Freedom)).

$$C_p(T) = \frac{\exp\left\{\frac{-\Delta H_{VH1}}{RT}\left(1 - \frac{T}{T_{m1}}\right)\right\} \Delta H_{VH1} \Delta H_{cal1}}{\left[1 + \exp\left\{\frac{-\Delta H_{VH1}}{RT}\left(1 - \frac{T}{T_{m1}}\right)\right\}\right]^2 RT^2} + \frac{\exp\left\{\frac{-\Delta H_{VH2}}{RT}\left(1 - \frac{T}{T_{m2}}\right)\right\} \Delta H_{VH2} \Delta H_{cal2}}{\left[1 + \exp\left\{\frac{-\Delta H_{VH2}}{RT}\left(1 - \frac{T}{T_{m2}}\right)\right\}\right]^2 RT^2} \quad (5)$$

where  $T_{m1}$  (melting temperature),  $\Delta H_{VH1}$  (van 't Hoff enthalpy change) and  $\Delta H_{cal1}$  (calorific enthalpy change) account for the thermal transition of domain 1 whereas  $T_{m2}$ ,  $\Delta H_{VH2}$  and  $\Delta H_{cal2}$  account for the thermal transition of domain 2.  $C_p$  is the excess heat capacity ( $\text{kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ ).

The activation energy for irreversible transitions (EA), considered as one-step processes from native to irreversibly inactivated state of the protein [22], was determined by Eq (6) where  $C_p^{\text{ex}}_{\text{max}}$  is the maximum excess of heat capacity obtained at the  $T_m$ .

$$E_A = \frac{eRT_m^2 C_p^{\text{ex}}_{\text{max}}}{\Delta H_{cal}} \quad (6)$$

Samples of  $0.6\text{ mg}\cdot\text{mL}^{-1}$  protein in 5 mM Tris-ClH buffer 150 mM NaCl pH 7 containing 2 M urea in the same buffer were subjected to DSC in the temperature range  $20\text{--}120^\circ\text{C}$  at  $60^\circ\text{C}\cdot\text{h}^{-1}$  scan rate. The chaotrope was added to the reference cell at the same concentration than was added to the sample cell.

## Analytical Procedures

Along the propeptide production and purification processes, protein concentration was routinely determined with colorimetric methods as described [13] but once purified, the propeptide concentration was calculated from its absorbance at 280 nm using  $20,790\text{ M}^{-1}\cdot\text{cm}^{-1}$  as molar extinction coefficient at 280 nm [14].

Free thiols of purified SP-B<sub>N</sub> were titrated by adding 40  $\mu\text{L}$  of 5 mM 5,5'-dithio-bis-(2-nitrobenzoic acid) (DTNB, Sigma) to 38.4  $\mu\text{g}$  of protein in 960  $\mu\text{L}$  of 5 mM Tris-HCl pH 8.2, 150 mM NaCl (assay buffer). Protein was previously dialyzed towards the same buffer at  $4^\circ\text{C}$ . The absorbance at 412 nm was recorded after 2 h at  $22^\circ\text{C}$  to check the release of 5-thiobis-(2-nitrobenzoic acid) (TNB) [23]. L-Cysteine (Fluka) was the standard and  $\epsilon_{412} = 13,800\text{ M}^{-1}\cdot\text{cm}^{-1}$  was used to quantify cysteines. For titration under denaturing conditions the same

protocol was applied except that the protein was preincubated with the assay buffer containing 5.7 M GdmCl, which was added 30 min prior to DTNB addition. Free thiols were also analyzed in the fusion MBP-SP-B<sub>N</sub> (33.5 μg) and in MBP (75.2 μg; released from the fusion by factor Xa) after removing the 2-mercaptoethanol contained in the purification buffer by extensive dialysis towards the assay buffer.

## Results

### Effect of Chaotropes on the Secondary Structure of SP-B<sub>N</sub>

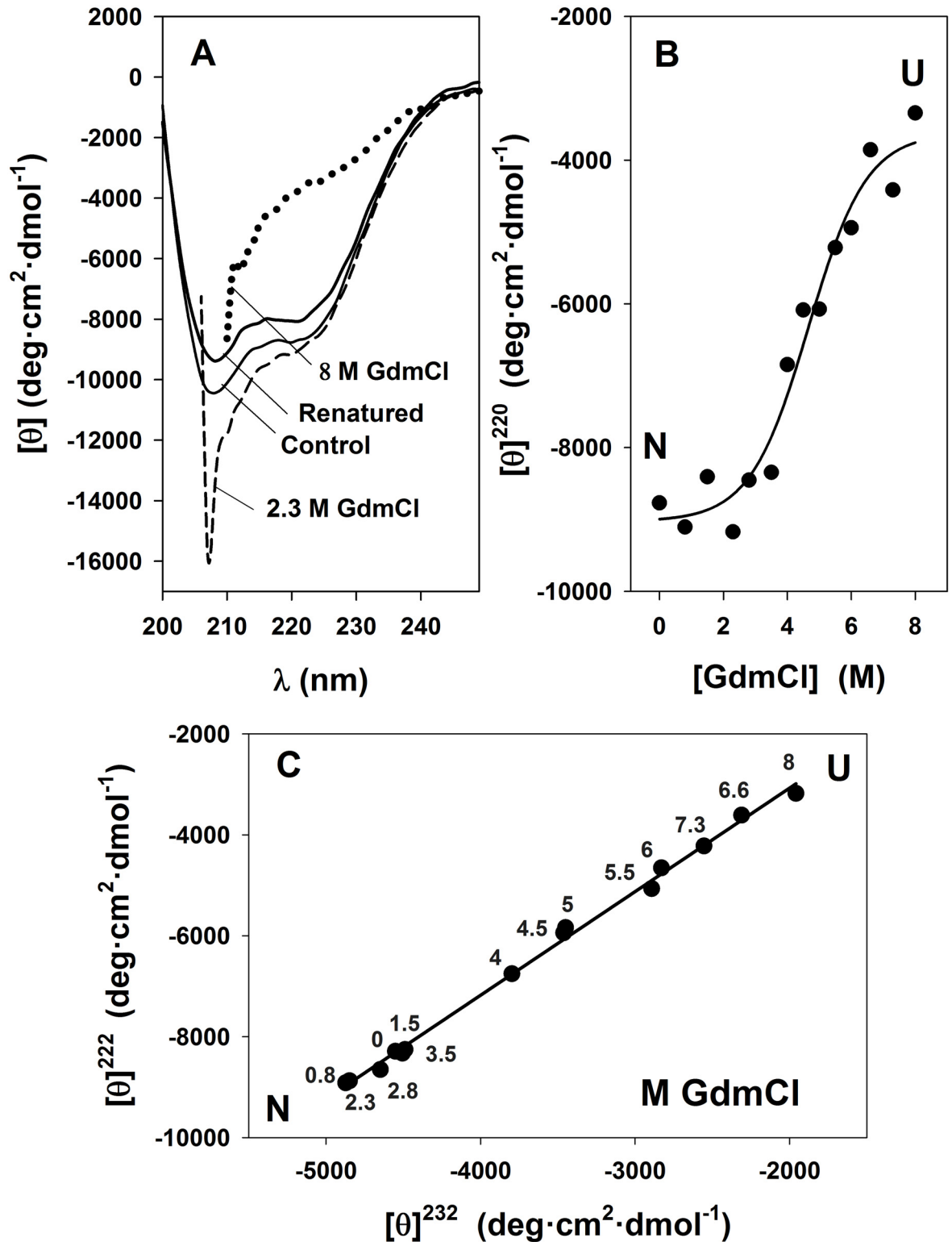
The chemical unfolding of the propeptide was achieved with GdmCl and urea. The far-UV CD spectra of the propeptide in the presence of 0–8 M GdmCl are depicted in Fig 1.

GdmCl is an electrolyte, and thus it is able to establish undesirable protein-denaturant electrostatic interactions [24] but the ionic strength provided by the salt (150 mM NaCl) in the sample must screen the charges on the surface of the protein, avoiding those interactions. Fig 1A shows the propeptide spectrum recorded in the absence of denaturant (control, solid line) and the spectra obtained in the presence of 2.3 M GdmCl (dashed line) and 8 M GdmCl (dotted line). The shape of the spectrum in the presence of 8 M GdmCl indicates that the protein is unfolded. To check the reversibility of the unfolding process, the chaotrope was eliminated from the sample containing 8 M GdmCl through exhaustive dialysis at 4°C towards the buffer without denaturant, and the spectrum was recorded again (renatured, solid line). As the spectra of control and renatured sample are quite close (92% of the control molar ellipticity was recovered), we can assume that renaturation is achieved, and this allows the thermodynamic treatment of the data. The unfolding transition in Fig 1B shows the variation of the molar ellipticity at 220 nm ( $[\theta]^{220}$ ) as [GdmCl] is increased. The data were fitted to Eq (1) and the sigmoid indicates that the backbone of the propeptide unfolds cooperatively, following a two-state transition process from the native protein (N) to (U), the assumed unfolded state, since no clear plateau was reached. No accumulation of intermediate states was detected in the unfolding process. The thermodynamic parameters obtained through the fitting,  $\Delta G^0_{\text{H}_2\text{O}} = 12.7 \pm 3.2$  kJ·mol<sup>-1</sup> and  $m = 2.7 \pm 0.7$  kJ·mol<sup>-1</sup>·M<sup>-1</sup>, are summarized in Table 1.

The midpoint transition,  $D_{1/2}$ , calculated as  $\Delta G^0_{\text{H}_2\text{O}}/m$  was 4.7 M, but the calculated standard error was higher than 70%, according to the error theory. Thus,  $D_{1/2}$  was determined by fitting the same set of data to Eq (2), an equation describing the same sigmoid but containing  $D_{1/2}$  as parameter instead of  $m$  and  $\Delta G^0_{\text{H}_2\text{O}}$  (see Materials and Methods). The midpoint transition was  $D_{1/2} = 4.7 \pm 0.2$  M. To detect hidden intermediate states in the GdmCl unfolding process, if any, we applied the phase diagram method, which deals with the parametric dependence of independent parameters under different experimental conditions for a protein undergoing structural transformations [25]. Fig 1C depicts the molar ellipticity of SP-B<sub>N</sub> at 232 nm ( $[\theta]^{232}$ ) vs the obtained at 222 nm ( $[\theta]^{222}$ ) for 0–8 M GdmCl. All data in plot were fitted by linear regression confirming that, regarding secondary structure, only the states N and U are detected along the propeptide unfolding process. However, some caution must be considered since the election of some parameters instead of others to create a phase diagram, may allow either to detect or to miss intermediate states as has been reported for other proteins [25].

The secondary structure of SP-B<sub>N</sub> was also studied in the presence of 0–7.5 M urea by far-UV CD spectroscopy (Fig 2).

The CD spectra of SP-B<sub>N</sub> are plotted in Fig 2A for control sample (solid line), with 2 M urea (dashed line) and with 7.5 M urea (dotted line). The spectrum at 2 M urea exhibits more negative  $[\theta]^{220}$  than control with the concomitant increase of the  $\alpha$ -helix content, which suggests the existence of a stabilized intermediate state in the protein unfolding process regarding secondary structure. The CD spectrum recorded after dialysis of the denaturant at 4°C of the



**Fig 1. Effect of GdmCl on the secondary structure of SP-B<sub>N</sub> measured by far-UV circular dichroism spectroscopy.** (A) Spectra of the propeptide recorded in 5 mM AMT buffer, 150 mM NaCl pH 7 (control, solid line) and with 2.3 M GdmCl (dashed line), 8 M GdmCl (dotted line) or 8 M GdmCl and ulterior dialysis (renatured, solid line). Spectra recorded at other GdmCl concentrations are not depicted for clarity. (B) Variation of the mean residue ellipticity at 220 nm with 0–8 M GdmCl. The line is a fit of the experimental data to Eq (1) ( $r = 0.983$ ). N and U are native and unfolded forms of the propeptide. (C) Phase diagram of

0–8 M GdmCl based on far-UV CD (mean residue molar ellipticity at 222 nm vs 232 nm). The line is the lineal regression fit of the data ( $r = 0.99$ ). Numbers beside the points are denaturant concentration.

doi:10.1371/journal.pone.0158430.g001

sample containing 7.5 M urea (renatured sample, solid line) is also shown in Fig 2A. Since the unfolding was practically reversible, the thermodynamic analysis of the CD data was possible. The dependence of the  $[\theta]^{220}$  on urea concentration is depicted in Fig 2B. We see that the ellipticity becomes more negative from control to ~ 1.5–2 M urea whereas above ~ 3.5 M urea, the  $[\theta]^{220}$  increases cooperatively up to 7.5 M urea. As moderate urea concentrations stabilize the propeptide, it excludes the existence of a two-state transition process regarding its secondary structure and suggests that the unfolding is reached through an intermediate state. When more than two states exist in protein unfolding, the treatment of the data is usually carried out considering a sequence of independent separate two-state transitions [26]. Therefore, only a part of the experimental data (1.5–7.5 M urea) in Fig 2B were fitted to Eq (2), being substituted  $[\theta]^{220}$  in the N state ( $Y_N$ ) by  $[\theta]^{220}$  in the intermediate state ( $Y_I$ ). The midpoint transition was  $5.2 \pm 0.3$  M urea. These results point to the existence of, at least, an intermediate state (I) at ~ 2 M urea as the propeptide goes from its native form (N) to its assumed unfolded state (U, as a clear plateau is not observed) with the increase in urea concentration. The experimental data in the 1.5–7.5 M urea range were also fitted to Eq (1) obtaining  $\Delta G^0_{H_2O} = 12.8 \pm 3.1$  kJ·mol<sup>-1</sup> and  $m = 2.5 \pm 0.7$  kJ·mol<sup>-1</sup>·M<sup>-1</sup> (Table 1). The phase diagram in Fig 2C confirms the existence of the I state regarding the secondary structure since two sets of data were fitted by linear regression: 0–1.5 M urea (transition N-I) and 1.5–7.5 M urea (transition I-U)

### Effect of Chaotropes on the Tertiary Structure of SP-B<sub>N</sub>

The effect of chaotropes in the tertiary structure of SP-B<sub>N</sub> was analyzed studying their effect on the trp environment by fluorescence spectroscopy. Although further analysis of the trp fluorescence emission change would require time-resolved measurements, we will simply use it as a reporter of the environment of the residues. The effect of GdmCl on the propeptide fluorescence intensity emission (FI) is shown in Fig 3.

The spectra of the propeptide without GdmCl (control, solid line), with 2.3 M GdmCl (dashed line) and with 6 M GdmCl (dotted line) are depicted in Fig 3A. As the protein becomes unfolded, the emission increases since the chemical unfolding modifies the environment of trp residues, which become less quenched than in the folded protein. In parallel, the maximum wavelength of emission is shifted from 338 nm in the absence of GdmCl to 345 nm with 6.6 M GdmCl (not shown), indicating a transition of the trp residues, from a hydrophobic environment inner in the protein to that of the polar aqueous medium. The spectrum of the dialyzed 6.6 M GdmCl sample (renatured, solid line) shows the shape of the control spectrum but twice

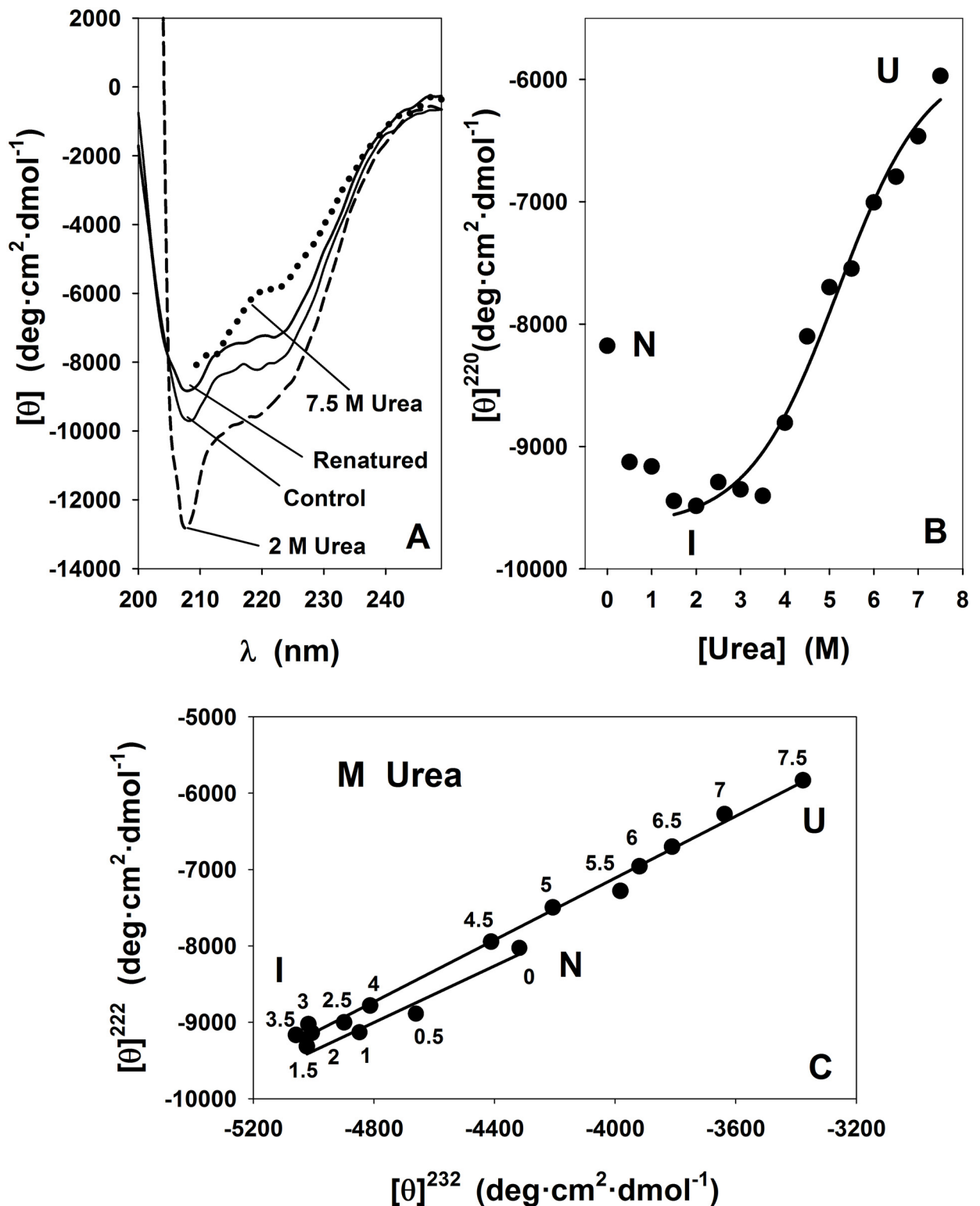
**Table 1. Thermodynamic parameters of SP-B<sub>N</sub> unfolding by GdmCl and urea followed by tryptophan fluorescence and far-UV circular dichroism.**

	$[\theta]^{220}$			$FI_{350}$		
	$D_{1/2}$ (M)	$\Delta G^0_{H_2O}$ (kJ·mol <sup>-1</sup> )	$m$ (kJ·mol <sup>-1</sup> ·M <sup>-1</sup> )	$D_{1/2}$ (M)	$\Delta G^0_{H_2O}$ (kJ·mol <sup>-1</sup> )	$m$ (kJ·mol <sup>-1</sup> ·M <sup>-1</sup> )
<b>GdmCl</b>	4.7±0.2 (N-U)	12.7±3.2 (N-U)	2.7±0.7 (N-U)	3.8±0.2 (I <sub>1</sub> -U)	12.8±3.3 (I <sub>1</sub> -U)	3.4±0.8 (I <sub>1</sub> -U)
<b>Urea</b>	5.2±0.3 (I-U)	12.8±3.1 (I-U)	2.5±0.7 (I-U)	ND	ND	ND

$D_{1/2}$  is the denaturant concentration at the midpoint transition in the  $[\theta]^{220}$  (deg·cm<sup>2</sup>·dmol<sup>-1</sup>) or  $FI_{350}$  (nm) dependence on denaturant.  $D_{1/2}$  values were obtained by fitting the data of Figs 1B, 2B and 3B to Eq (3).  $\Delta G^0_{H_2O}$  and  $m$  values were obtained by fitting the same data to Eq (2). N, U and I mean native, unfolded and intermediate state forms of the propeptide respectively and their appearance in parentheses indicate the protein forms holding the actual transition. I<sub>1</sub> mean intermediate state of the protein when more than one is observed. ND is not determined.

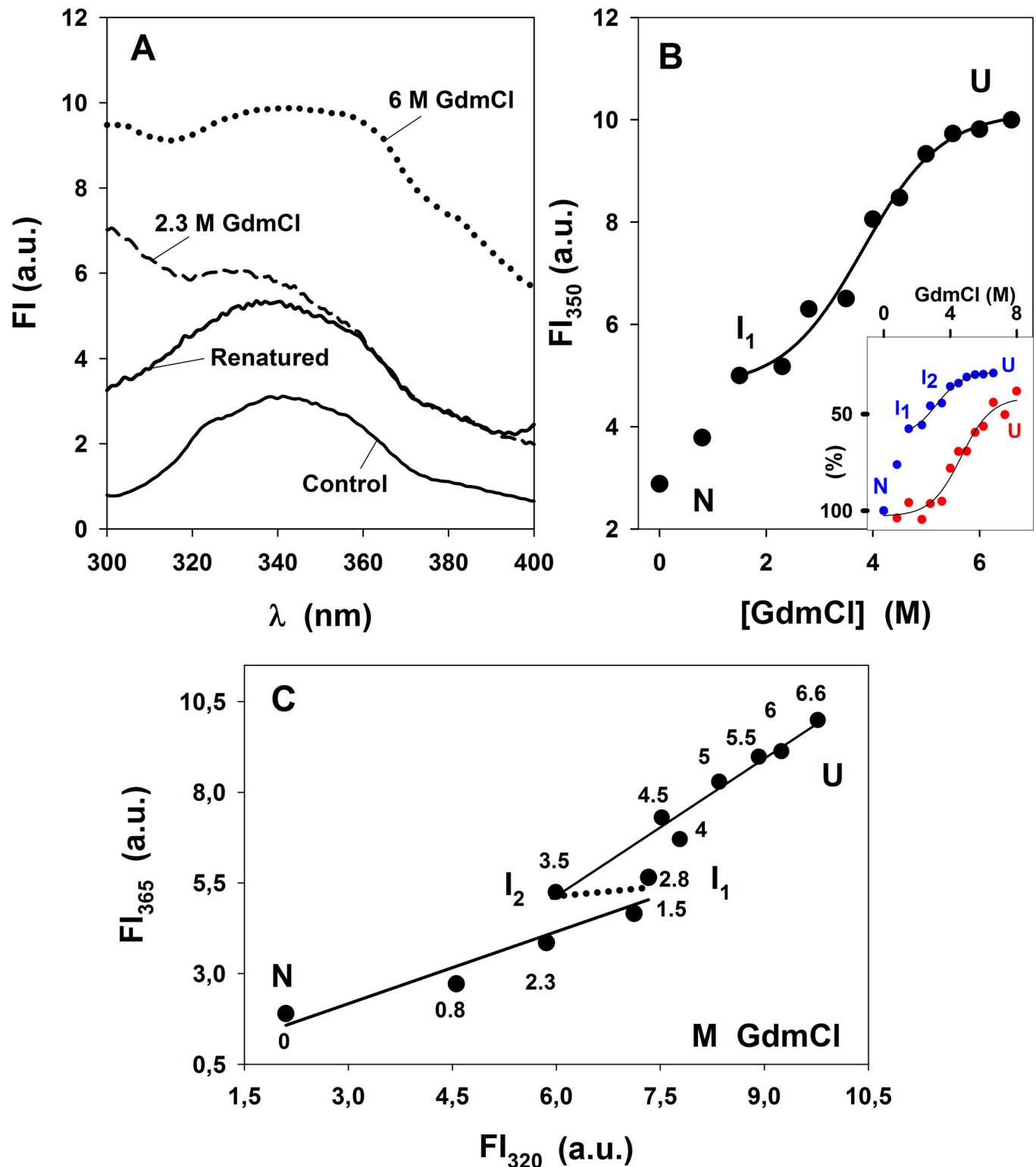
doi:10.1371/journal.pone.0158430.t001





**Fig 2. Effect of Urea on the secondary structure of SP-B<sub>N</sub> measured by far-UV circular dichroism spectroscopy.** (A) Spectra of the propeptide recorded in 5 mM AMT buffer, 150 mM NaCl pH 7 (control, solid line), with 2 M urea (dashed line), with 7.5 M urea (dotted line) and with 7.5 M urea and ulterior dialysis (renatured, solid line). (B) Mean residue molar ellipticity at 220 nm vs urea concentration. The line is a fit of part of the data (1.5–7.5 M urea;  $r = 0.99$ ) to Eq (1). (C) Phase diagram analysis of urea-induced equilibrium unfolding of SP-B<sub>N</sub>. The lines (0–1.5 M;  $r = 0.98$ ) and (1.5–7.5 M;  $r = 0.99$ ) were obtained by linear regression. I is an intermediate state in the propeptide unfolding process; N and U as in Fig 1. Numbers beside the points are denaturant concentration.

doi:10.1371/journal.pone.0158430.g002



**Fig 3. Effect of GdmCl on the fluorescence emission of SP-B<sub>N</sub>.** (A) Fluorescence intensity emission (FI) of the propeptide in 5 mM AMT buffer, 150 mM NaCl pH 7 with 0–6.6 M GdmCl after excitation at 290 nm vs the wavelength. Spectrum without GdmCl (control, solid line), with 2.3 M GdmCl (dashed line), 6 M GdmCl (dotted line) or 6 M GdmCl and ulterior dialysis towards the same buffer (renatured, solid line) were recorded at 25°C. Arbitrary units (a.u.) were employed. (B) Variation of the fluorescence emission at 350 nm with 0–6.6 M GdmCl. The line is a fit of data in the 1.5–6.6 M interval to Eq (1) ( $r = 0.99$ ). Inset: Comparison of secondary structure (%  $[\theta]^{220}$ , red circle, data and fitting from Fig 1B) and tertiary

structure ( $FI_{350}N \cdot (FI_{350})^{-1} \cdot 10^2$ ), blue symbol, data and fitting from Fig 3B, N is the native form). (C) Phase diagram of 0–6.6 M GdmCl based on the fluorescence emission of the propeptide at two wavelengths. Solid lines are the fit of experimental data by linear regression: 0–2.8 M ( $r = 0.96$ ) and 3.5 to 6.6 M ( $r = 0.98$ ). Dotted line (2.8–3.5 M) is not a fit.  $I_1$  and  $I_2$  are intermediate states in the protein GdmCl-induced unfolding; N and U as in Fig 1. Numbers beside the points are denaturant concentration.

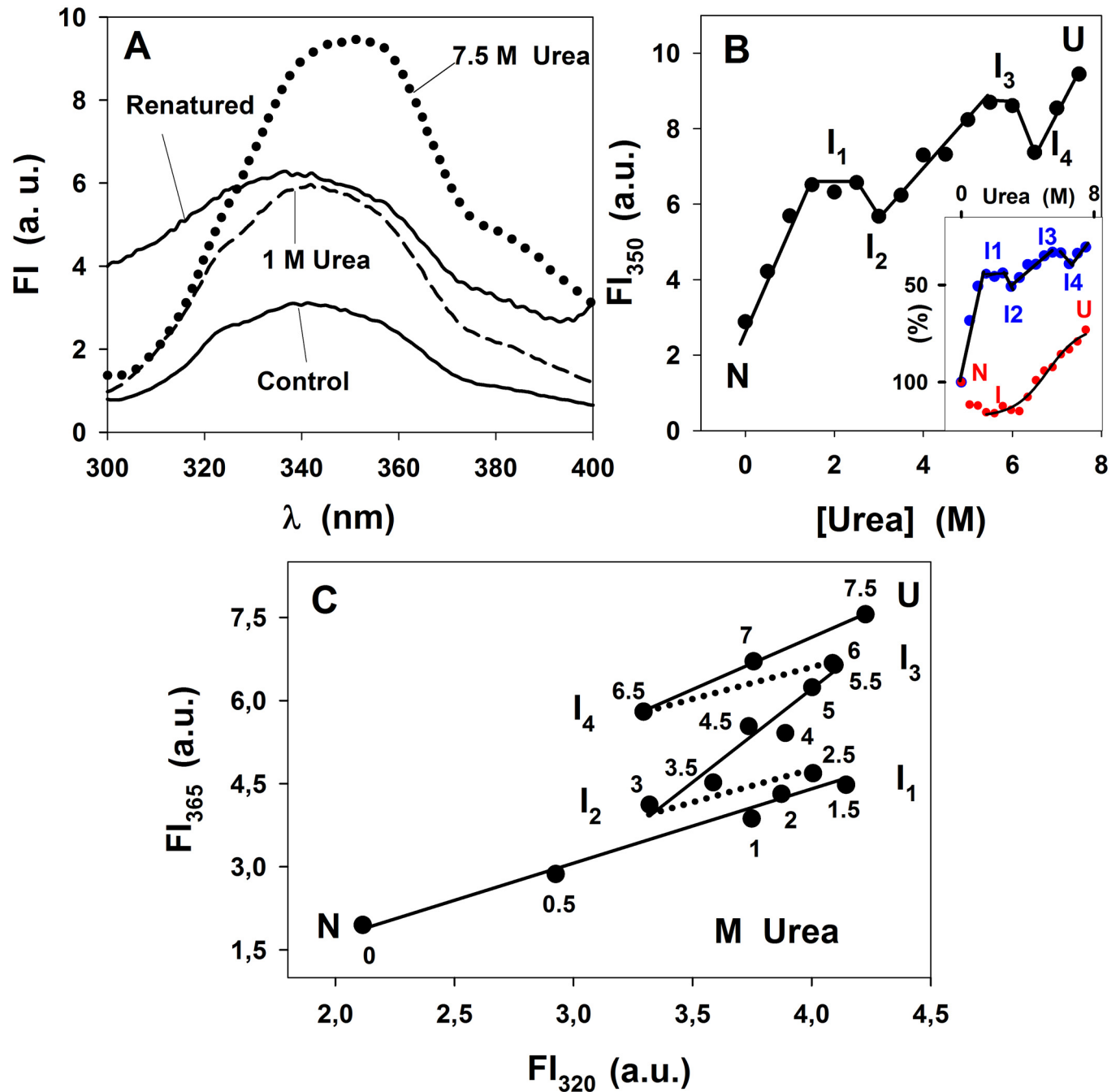
doi:10.1371/journal.pone.0158430.g003

its emission indicating that the reversion was only partial and approached the emission of the protein in the presence of 1 M denaturant (not shown). The effect of GdmCl on the fluorescence intensity emission at 350 nm ( $FI_{350}$ ) was studied since the FI is an extensive parameter, its value being proportional to the amount of the analyzed matter in the system [25] and is depicted in Fig 3B. The fit of the data obtained with 1.5–6.6 M GdmCl (Fig 3B) to Eq (2) gave a  $D_{1/2} = 3.8 \pm 0.2$  M GdmCl which is lower than the obtained for the secondary structure with the same chaotrope (see Table 1) indicating that the loss of tertiary structure precedes the loss of the secondary one during GdmCl-induced propeptide unfolding. The comparison of the effect of GdmCl in the secondary and tertiary structure of the propeptide (Fig 3, inset) and the fact that both curves do not superimpose corroborates the existence of intermediates. The intermediates  $I_1$  (detected in Fig 3B) and  $I_2$  (detected in Fig 3B and 3C) with decreased tertiary structure, still have a significant proportion of native secondary structure. The thermodynamic parameters were obtained above  $\sim 1$  M GdmCl since the effect of GdmCl in the propeptide was only partially reverted. The fitting of the same data in Fig 3B to Eq (1) gave  $\Delta G^0_{H_2O} = 12.8 \pm 3.3$   $\text{kJ} \cdot \text{mol}^{-1}$ , coincident with the value obtained with GdmCl in CD experiments whereas  $m = 3.4 \pm 0.8$   $\text{kJ} \cdot \text{mol}^{-1} \cdot \text{M}^{-1}$  is greater than estimated from the CD study, indicating that changes affecting the exposure of trp residues to the solvent show more cooperativity than those affecting the backbone of the protein. Those thermodynamic values would correspond to the transition  $I_1$ -U, being  $I_1$  the intermediate form of the protein seen around 2 M GdmCl.

As the phase diagram method is a powerful tool to uncover hidden intermediates [25], the variation of  $FI_{365}$  with  $FI_{320}$  for each GdmCl concentration was determined and depicted in Fig 3C, to detect intermediate states of the protein. Two sets of data were fitted by linear regression, corresponding to GdmCl concentration ranges of 0–2.8 M and 3.5–6.6 M indicating that, in addition to the N and U forms, there are two intermediate forms ( $I_1$  and  $I_2$ ) of the propeptide regarding its tertiary structure. Therefore, the phase diagram detects an intermediate state ( $I_1$ ) at  $\sim 2$  M GdmCl (also seen in Fig 3B), and another intermediate state ( $I_2$ ) at  $\sim 3.5$  M GdmCl not seen previously. Both intermediate states,  $I_1$  and  $I_2$ , hold altered tertiary structure while maintaining the secondary structure of the native state (see Fig 1B), a characteristic of the molten globule state [2]. Therefore, considering the results relative to the tertiary structure, the propeptide in the presence of GdmCl, would unfold according to a four-state transition as follows:  $N \rightarrow I_1 \leftrightarrow I_2 \leftrightarrow U$ , being irreversible the step  $N \rightarrow I_1$  under the experimental conditions we used to eliminate the GdmCl in the medium.

The effect of urea on the tertiary structure of the propeptide is shown in Fig 4.

As described above with GdmCl, the higher the urea concentration, the greater the fluorescence emission (Fig 4A). The effect of urea on the protein could not be reverted totally (solid line, renatured). The unfolding process of the propeptide with urea is shown in Fig 4B, where no cooperative transition was observed but discontinuities such as small plateaux and valleys, pointing to the existence of several intermediate states ( $I_1$ ,  $I_2$ ,  $I_3$  and  $I_4$ ). The comparison of the effect of urea in the secondary and tertiary structure of the propeptide (Fig 4B, Inset) indicates that whereas  $I_1$  and  $I_2$  (observed in fluorescence spectroscopy) shows the stabilized secondary structure corresponding to I (observed in CD spectroscopy), the intermediate  $I_3$  conserves 92.3% of the native CD signal whereas  $I_4$  exhibits a more pronounced loss of secondary structure. The phase diagram in Fig 4C confirms the existence of four intermediate states, two of



**Fig 4. Effect of urea on the fluorescence emission of SP-B<sub>N</sub>.** (A) Emission spectra of the protein in 0–7.5 M urea were recorded as in Fig 3. Spectra depicted are: without urea (control, solid line), with 1 M urea (dashed line), with 7.5 M urea (dotted line) and with 7.5 M urea after dialysis (renatured, solid line). (B) Variation of the fluorescence emission at 350 nm with 0–7.5 M urea. Connecting lines are not fitted. Inset: Comparison of secondary structure (%  $[\theta]^{220}$ , red circle, data and fitting from Fig 2B) and tertiary structure ( $FI_{350}N \cdot (FI_{350})^{-1} \cdot 10^2$ , blue symbol, data from Fig 4B being N the native form). N and U are omitted for clarity. (C) Phase diagram analysis of urea-induced equilibrium unfolding of SP-B<sub>N</sub>. Solid lines are fits of the data by linear regression: 0–2.5 M urea ( $r = 0.987$ ), 3–6 M ( $r = 0.97$ ) and 6.5–7.5 M ( $r = 0.999$ ). Dotted lines are not fitted.  $I_1, I_2, I_3$  and  $I_4$  are intermediate states in the propeptide urea-induced unfolding. Lines are not fitted. Numbers beside the points are denaturant concentration.

doi:10.1371/journal.pone.0158430.g004

them at low and moderate urea as those seen at low GdmCl concentration ( $I_1$  and  $I_2$ ).  $I_3$  was observed at 5.5–6 M urea and  $I_4$  at ~ 6.5 M urea. The analysis of the  $\lambda_{\max}$  dependence on urea indicates that there are four small plateaux between the N and U state (not shown), supporting the idea that the progressive exposure of trp residues to the polar medium, is reached in four steps. The high number of conformers seen in Fig 4C suggests that changes affecting the trp environment of the propeptide with urea are rather complex and compatible with a multi-state structural process. Thus regarding the tertiary structure, the propeptide would unfold with urea according to:  $N \rightarrow I_1 \leftrightarrow I_2 \leftrightarrow I_3 \leftrightarrow I_4 \leftrightarrow U$ , being irreversible the  $N \rightarrow I_1$  process. Two of these intermediate states ( $I_3$  and  $I_4$ ) display near native and a remarkable loss of secondary structure respectively (Fig 2B). In contrast,  $I_1$  and  $I_2$ , show also altered secondary structure, although this structure is stabilized since the  $[\theta]^{220}$  is more negative (Fig 2B).

## Bis-ANS Binding Studies

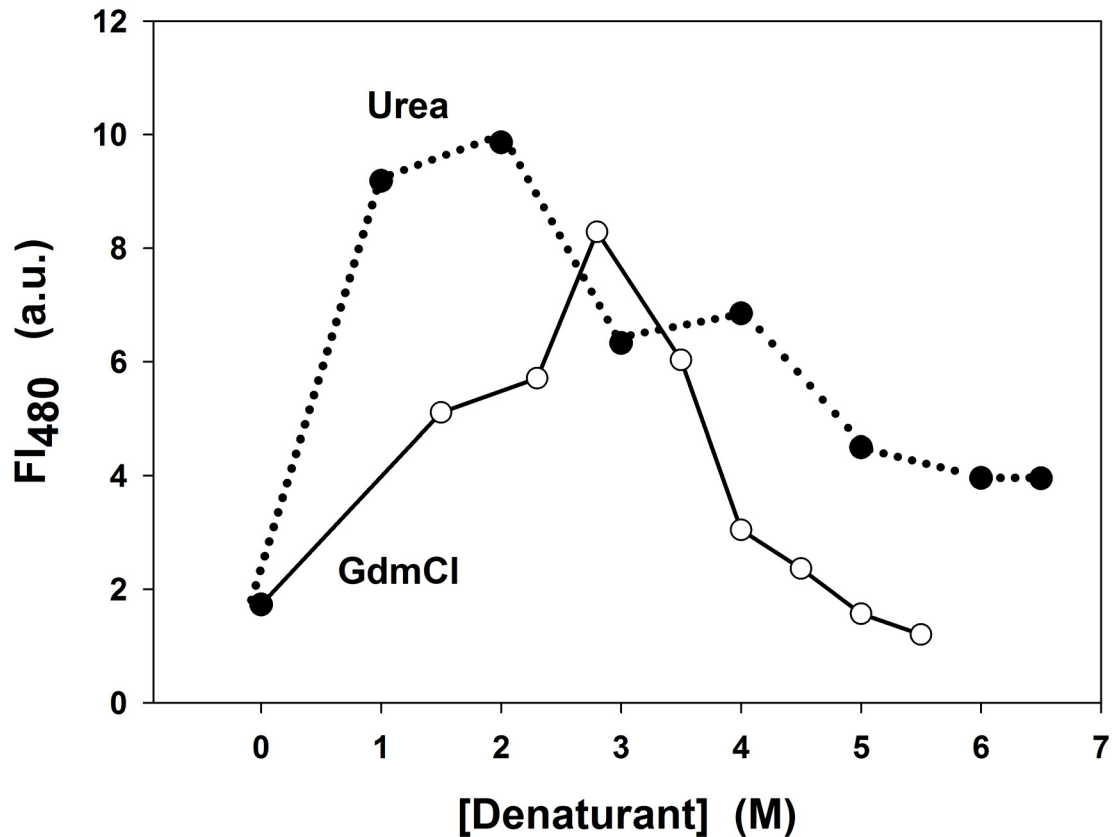
To check the molten globule nature of the intermediate states detected in the chemical-induced unfolding of the propeptide, we employed bis-ANS, a fluorescent probe that binds neither the native nor the unfolded state of proteins but is bound to intermediate structures with hydrophobic domains exposed to the medium [27]. Following the non-covalent binding of the probe to those domains, there is an increase in the fluorescence quantum yield due to the parallel orientation adopted by the two naphthyl rings of the probe. Therefore, bis-ANS can detect molten globule states of proteins holding secondary structure but with incomplete packing of the side chains in the hydrophobic core, meaning a fluctuating globular structure [28]. The effect that the environment of the propeptide has on the fluorescence emission of bis-ANS at 480 nm is shown in Fig 5.

Changes in the emission fluorescence intensity of bis-ANS were observed as urea or GdmCl concentration was increased. In the absence of denaturant, the fluorescence emission is poor but it increases with the denaturant concentration to reach maximum values at ~ 1.5–3 M ( $I_1$  and  $I_2$  intermediate states) to decrease thereafter. The decrease follows a sharp way with GdmCl as no additional intermediate state exists or in a soft way with urea according to the existence of  $I_3$  and  $I_4$  states approaching the unfolded form of the propeptide. Thus the probe bis-ANS shows emission at denaturant concentrations where intermediate states were detected, confirming them as molten globule or molten globule-like structures. Since the bis-ANS emission is higher with urea than with GdmCl, it follows that there is a higher number of exposed hydrophobic residues (where the probe can bind) in the propeptide upon unfolding with urea than upon unfolding with GdmCl.

## Effect of Temperature on SP-B<sub>N</sub> Secondary Structure

The effect of temperature on the structure of the propeptide was followed by far-UV CD in the presence of 150 mM NaCl (Fig 6A) or 500 mM NaCl (Fig 6B), to check also the influence of the ionic strength on protein thermal unfolding.

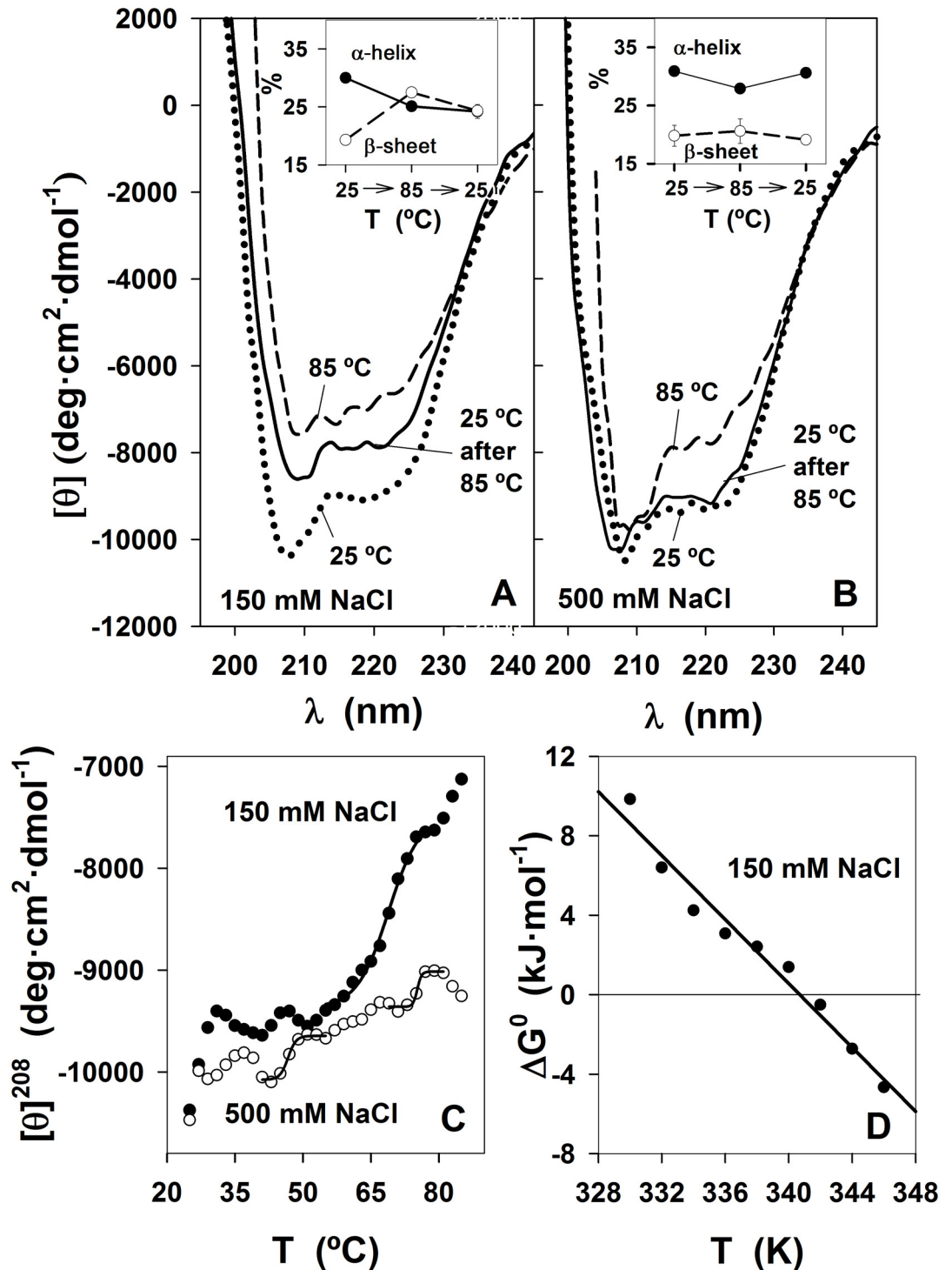
The spectrum of the sample was recorded at 25°C (dotted line), then at 85°C, (dashed line) and, after the T ramp the spectrum was recorded again at 25°C (solid line). The propeptide maintains ~ 71% of  $[\theta]^{208}$  with 150 mM NaCl with a decrease of 5% in the  $\alpha$ -helix content and a parallel increase in  $\beta$ -sheet content (Fig 6A, inset). At 500 mM NaCl, the only change observed at 85°C is a small decrease of the molar ellipticity at 220 nm (Fig 6B) without noteworthy changes in the  $\alpha$ -helix or  $\beta$ -sheet contents (Fig 6B, inset). The unfolding, whatever its extension is totally reversible after re-cooling at high salt (Fig 6B) and almost totally reversible at low salt (Fig 6A). The behaviour of SP-B<sub>N</sub> with temperature was monitored by recording  $[\theta]^{208}$  as shown in Fig 6C. We see that  $[\theta]^{208}$  at 25°C becomes less negative upon raising the



**Fig 5. Fluorescence emission of the probe bis-ANS in the presence of SP-B<sub>N</sub>.** Fluorescence intensity at 480 nm of the probe (11 μM) in the presence of SP-B<sub>N</sub> in 5 mM AMT buffer, 150 mM NaCl pH 7 with urea (filled circles, dotted line) and GdmCl (void circles, solid line) vs denaturant concentration.

doi:10.1371/journal.pone.0158430.g005

temperature to be more or less stabilized at ~ 38°C at low salt and at ~ 30°C at moderate salt, to wave again several times without reaching a clear plateau, indicative that the unfolded state was not reached at 85°C. Moreover, above ~70°C,  $[\theta]^{208}$  does not increase at 500 mM NaCl as much as it does with low salt, indicating that ionic strength stabilizes the secondary structure of the propeptide. We have considered multi-state transitions as if they were consecutive independent two-state transitions as we did when the denaturant was a chaotrope, and is usually done [26]. The fitting of the experimental data to Eq (2) at 150 mM NaCl in the 49–79°C range gave a midpoint transition of  $68.7 \pm 0.6^\circ\text{C}$ . When 500 mM NaCl was employed, three small transitions at 41–55°C, 51–69°C and 69–81°C could be detected and fitted to Eq (2), with midpoint transitions of  $46.7 \pm 0.2^\circ\text{C}$ ,  $62.7 \pm 1.6^\circ\text{C}$  and  $75.2 \pm 0.4^\circ\text{C}$  respectively. The observed changes in the propeptide secondary structure may be due to partial unfolding of protein domains or subdomains (small increases in  $[\theta]^{208}$ ) or to the dissociation of propeptide oligomers [15]. Alternatively, small decreases in  $[\theta]^{208}$  may be due to partial folding of unstructured domains existing in the connector arm of the propeptide (predictions in S2 Fig) as observed by Uversky in other proteins [29]. As the unfolding of the propeptide was not reached up to 87°C, its melting temperature could not be determined. The effect of temperature on the propeptide up to 85°C is practically reversible as seen in Fig 6A, thus the thermodynamic parameters for the transition at low salt could be calculated.  $\Delta G^0$  values were obtained by fitting data of Fig 6C in the 49–79°C interval to Eq (3) at each temperature with the exception of 51°C and temperatures above 73°C as the term  $(Y_N - Y_{\text{obs}})/(Y_{\text{obs}} - Y_U)$  in Eq (3) was negative and the natural logarithm could



**Fig 6. Dependence of the secondary structure of SP-B<sub>N</sub> on temperature.** (A) far UV-CD spectra of 0.115 mg·mL<sup>-1</sup> in AMT buffer, 150 mM NaCl pH 7 at 25°C (dotted line), 85°C (dashed line) and 85°C cooled back to 25°C (solid line). Inset: Plot of  $\alpha$ -helix (filled circles, solid line) and  $\beta$ -sheet (void circles, dashed line) content vs the temperature progress of the sample from 25°C to 85°C and back to 25°C. (B) Same as in A except that 500 mM NaCl was used. (C) Plot of  $[\theta]^{208}$  along the T ramp (25–85°C; data plotted every 2°C at 150 mM NaCl (filled circles) and 500 mM NaCl (void circles)). Data from 49–79°C at 150 mM NaCl were fitted to Eq (2) (solid line,  $r = 0.996$ ). Data from 41–55°C ( $r = 0.995$ ), 51–69°C ( $r = 0.981$ ) and 69–81°C ( $r = 0.989$ ) at 500 mM NaCl were also fitted to Eq (2) (solid lines through the data). (D) Dependence of  $\Delta G^0$  on temperature according to Eq (4).  $\Delta G^0$  values were obtained

according to Eq (3) from data (49–73°C) at 150 mM NaCl except 51°C since the term  $(Y_N - Y_{\text{obs}})/(Y_{\text{obs}} - Y_N)$  in Eq (3) was negative and the natural logarithm could not be determined. This term was also negative above 73°C. Data in plot were fitted by linear regression (solid line;  $r = 0.988$ ).

doi:10.1371/journal.pone.0158430.g006

not be calculated.  $\Delta G^0$  values were plotted vs the temperature in Fig 6D and data were fitted to Eq (4) (solid line). The intercept was the change in enthalpy for the transition,  $\Delta H^0 = 234.5 \text{ kJ}\cdot\text{mol}^{-1}$ , whereas the absolute value of the slope was the change in entropy  $\Delta S^0 = 0.69 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ . Those values indicate that the predominant forces involved in the interactions of the protein domains that suffer the thermal transition or in oligomer formation, are of enthalpic character, i.e. electrostatic interactions and/or hydrogen bonding. The calculated value of  $\Delta G^0$  at 25°C was  $28.9 \text{ kJ}\cdot\text{mol}^{-1}$ . The thermodynamic parameters of the protein in 500 mM NaCl could not be determined because of shortage of data to employ Eq (3). This was a consequence of the existence of small thermal transitions at moderate salt instead of the sole transition seen at low salt (Fig 6C).

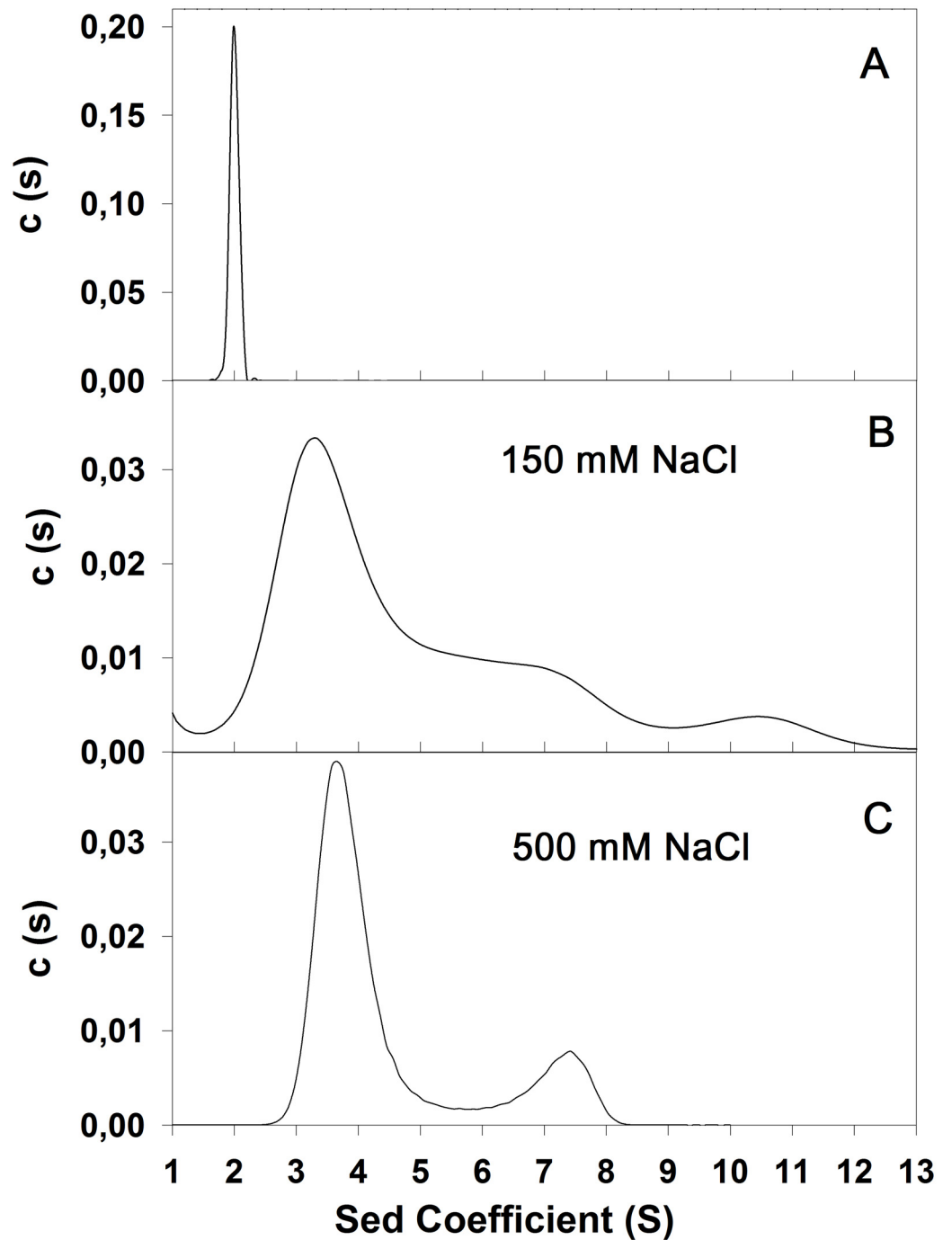
### Hydrodynamic Properties of SP-B<sub>N</sub>

In order to check the effect of the salt in the quaternary structure of the recombinant protein we performed boundary sedimentation experiments with  $0.15 \text{ mg}\cdot\text{mL}^{-1}$  protein in 5 mM Tris-HCl pH 7 under different salt conditions and results are depicted in Fig 7.

The raw sedimentation coefficients ( $s$ ) and the sedimentation coefficients in standard conditions, 20°C and water ( $s_{20,w}$ ) are summarized in Table 2.

The sedimentation coefficient distribution,  $c(s)$  in the absence of NaCl is depicted in Fig 7A. It corresponds to a symmetrical peak with  $s = 1.96 \pm 0.26 \text{ S}$  and  $s_{20,w} = 2.01 \pm 0.27 \text{ S}$ , corresponding to an apparent molecular mass of  $\sim 19.7 \text{ kDa}$  which is well in line with a SP-B<sub>N</sub> monomer since the molecular weight of SP-B<sub>N</sub> determined by mass spectrometry is 19,883 Da [13] and the estimated from the amino acid composition is 19,902 Da. The frictional ratio  $f/f_0 = 1.6$  would indicate that the species has an elongated shape, equivalent to a prolate ( $a/b = 7.86$ ) or oblate ( $a/b = 8.74$ ) ellipsoid although we cannot discard that the extremely low ionic strength the buffer provides is affecting the species sedimentation. Moreover, only 6% of the loaded protein was detected which indicates a massive protein precipitation. In the presence of 500 mM NaCl (Fig 7C) the protein (35% of the loaded protein) sediments mainly with an apparent sedimentation coefficient of  $3.74 \pm 0.76 \text{ S}$  and  $6.95 \pm 0.61 \text{ S}$ . These peaks are compatible with dimeric ( $\sim 40 \text{ kDa}$ ) and hexameric forms ( $\sim 123 \text{ kDa}$ ) of the propeptide, respectively, although oligomers with another stoichiometry can coexist at lower concentration, since the peaks are not really slender. The lower averaged frictional ratio,  $f/f_0 = 1.1$  at this salt concentration, point to monomers associating in a side-by-side rather than end-to-end manner to form oligomers more compact than the monomer. Analysis of the sedimentation profile of SP-B<sub>N</sub> was also carried out with 150 mM NaCl and results in Fig 7B indicate a great heterogeneity in the sample. The average sedimentation coefficient was only determined for the species constituting the major part of the preparation ( $s = 3.49 \pm 0.66 \text{ S}$ ; averaged  $f/f_0 = 1.34$ ; appM  $\sim 39.6 \text{ kDa}$ ) indicating that the dimer co-exists with several other species whose sedimentation coefficient could not be evaluated due to lack of resolution of the peaks. Moreover, the maxima in the sedimentation coefficient distribution do not necessarily corresponds to the precise values of the actual sedimentation coefficients of each species because the positions of these maxima are sensitive to the rate of monomer-oligomer association/dissociation as has been observed for other proteins [30]. Considering the sedimentation coefficients obtained in the experiment, it appears that monomers, anions at pH 7, would counteract the electrostatic repulsion between them,





**Fig 7. Dependence of the sedimentation velocity of SP-B<sub>N</sub> on ionic strength.** (A) Sedimentation coefficient distribution at 20°C and 48,000 rpm of the propeptide (0.15 mg·mL<sup>-1</sup>) in 5 mM Tris buffer pH 7. (B) Same conditions as described in A with 150 mM NaCl. (C) Same conditions as described in A with 500 mM NaCl.

doi:10.1371/journal.pone.0158430.g007

hiding their hydrophobic regions through mutual interaction to form dimers and higher oligomers of a globular shape, in the presence of low or moderate salt in the medium. SP-B<sub>N</sub> monomers are only detected in absence of added salt although the vast majority of the protein suffers precipitation propitiated by the low ionic strength.

**Table 2. Sedimentation velocity parameters of SP-B<sub>N</sub>.**

NaCl (mM)	Parameters			
	s (S)	s <sub>20,w</sub> (S)	f/f <sub>0</sub>	appM (kDa)
	1.96±0.26	2.01±0.27	1.6	19.7
150	3.49±0.66	3.60±0.67	1.34	39.6
500	3.74±0.76	4.15±0.84	1.1	44
	6.95±0.61	7.71±0.67	1.1	123

The experimental sedimentation coefficients (s) were converted to standard conditions (s<sub>20,w</sub>; 20°C, water) and expressed in Svedberg. f/f<sub>0</sub> is the frictional ratio of the protein to a sphere and is an average of the different species in the sample, and appM is the apparent molecular mass of the species.

doi:10.1371/journal.pone.0158430.t002

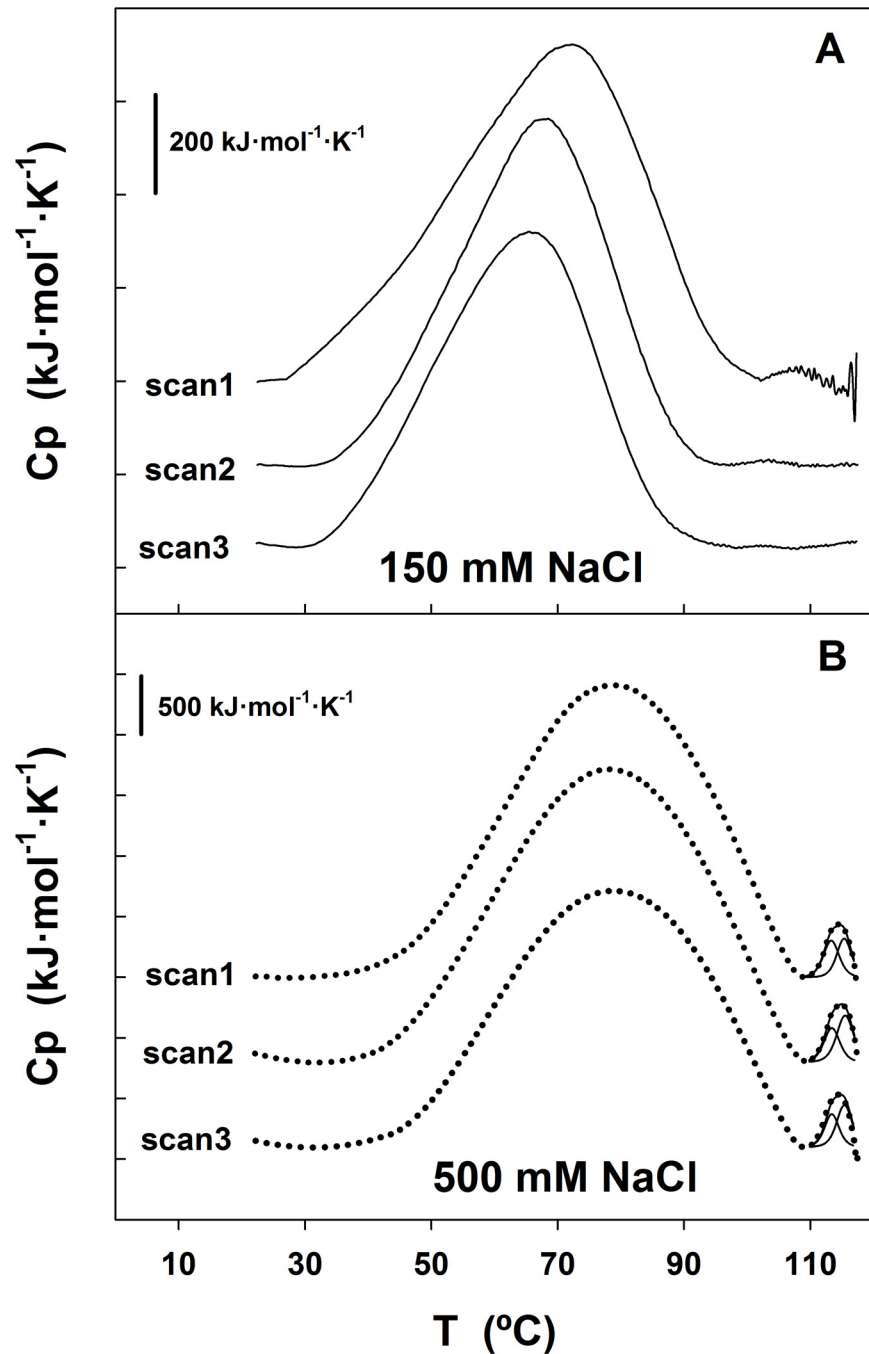
### Calorimetric Behaviour of SP-B<sub>N</sub>

Differential scanning calorimetry (DSC) was employed for the thermodynamic characterization of the structural changes accompanying the unfolding process of 0.3 mg·mL<sup>-1</sup> SP-B<sub>N</sub> in the 20–120°C interval at 60°C·h<sup>-1</sup> scan rate (Fig 8).

The variation of the excess heat capacity (C<sub>p</sub>) of the protein with the temperature at 150 mM NaCl in Fig 8A (scan1) is characterized by two endothermic peaks whose integration gave transition temperatures that we considered as apparent T<sub>m</sub> along this work: T<sub>m1</sub> = 72.2°C and T<sub>m2</sub> = 108.5°C respectively. After cooling and rescanning the sample twice, the first transition was observed in the successive scans with the T<sub>m1</sub> moving to lower values and the calorimetric enthalpy change obtained by numerical integration of the peak (ΔH<sub>cal</sub>) decreasing as if less protein were available in the rescans to suffer this thermal transition (ΔH<sub>cal</sub> was ~18% smaller in scan 2 and ~24% in scan 3; Table 3).

When two transitions are observed in the thermogram of an oligomeric polypeptide, the first one is usually attributed to oligomer dissociation and the second to monomeric protein unfolding [31]. The data corresponding to the first transition could not be fit to the oligomer dissociation equations provided by Origin. The cause may be that those Eqs apply to the dissociation of oligomers with definite stoichiometry such as dimer, hexamer, u.s.w. whereas SP-B<sub>N</sub> is a complex mixture of oligomers of different sizes as seen in analytical ultracentrifugation (Fig 7C) and the endothermic peak must be due to the dissociation of more than one oligomer type. Regarding the second transition, the small peak observed in scan 1 was unobserved after rescanning, indicating that this calorimetric process is irreversible. There is a small exothermic contribution following this endothermic peak, which suggests that the protein is suffering unfolding. Moreover, the noisy traces seen at higher temperatures are likely due to convection of clump aggregates into the DSC cell. Hence, it is probable that aggregation of unfolded protein causes the irreversibility. As there is no clear exothermic peak, we suggest that the cooling following unfolding might allow a partial refolding which is seen, after rescanning, only in the first transition as oligomers of distinct (and greater) n-mers dissociation. We tried to find if the aggregation process was under kinetic control by varying the scan rate but the thermograms obtained at 15°C·h<sup>-1</sup> and 30°C·h<sup>-1</sup> showed also protein aggregation at higher temperatures (not shown). Therefore, the dependence of T<sub>m2</sub> on the scan rate could not be established. The activation energy for the irreversible transition (one-step process from native to irreversibly inactivated state of the protein) was obtained with Eq (6), E<sub>A</sub> = 703.8 kJ·mol<sup>-1</sup> which is a high value and points to the protein showing a high barrier for the inactivation.

To study the influence of the increased ionic strength in the calorimetric behaviour of SP-B<sub>N</sub>, the experiment was repeated under the same conditions except that 500 mM NaCl were employed. Results are depicted in Fig 8B. The endothermic peaks observed in the thermogram



**Fig 8. Thermal unfolding of SP-B<sub>N</sub> followed by DSC.** (A) Variation of the excess heat capacity ( $C_p$ ) of SP-B<sub>N</sub> ( $0.3 \text{ mg}\cdot\text{mL}^{-1}$ ) with temperature at  $60^{\circ}\text{C}\cdot\text{h}^{-1}$  in 5 mM Tris-ClH 150 mM NaCl pH 7 (scan 1, solid line). The heated sample was cooled back,  $C_p$  recorded (scan 2), cooled again and  $C_p$  recorded (scan 3). (B) Same as in A except that the sample contained 500 mM NaCl. The dotted lines correspond to the experimental data, the solid lines through these points are the cumulative Gauss curves according to the model of Non Two-States and the other two solid lines were obtained by deconvolution analysis with Eq (5).

doi:10.1371/journal.pone.0158430.g008

**Table 3. Parameters of the SP-B<sub>N</sub> calorimetric transitions measured by DSC.**

NaCl (mM)	Transition 1		Transition 2			
	T <sub>m1</sub> <sup>a</sup> (°C)	ΔH <sub>cal 1</sub> <sup>a</sup> (kJ·mol <sup>-1</sup> )	T <sub>m2</sub> <sup>a</sup> (°C)	ΔH <sub>cal 2</sub> <sup>a</sup> (kJ·mol <sup>-1</sup> )	ΔH <sub>VH</sub> <sup>b</sup> (kJ·mol <sup>-1</sup> )	E <sub>A</sub> <sup>c</sup> (kJ·mol <sup>-1</sup> )
<b>150</b>						
scan1	72.2	26,269	108.5	150.4	ND	703.8
scan2	68.4	21,411	ND	ND	ND	ND
scan3	65.6	19,929	ND	ND	ND	ND
<b>500</b>						
scan1	78.8	90,899	114.4 (113.3; 115.3) <sup>b</sup>	1,841 (994; 981) <sup>b</sup>	(1,505; 1,627) <sup>b</sup>	ND
scan2	77.6	93,291	114.8 (113.4; 115.5) <sup>b</sup>	2,089 (1,352; 930) <sup>b</sup>	(1,422; 1,495) <sup>b</sup>	ND
scan3	78.7	81,413	114.5 (113.4; 115.4) <sup>b</sup>	1,782 (1,115; 813) <sup>b</sup>	(1,547; 1,649) <sup>b</sup>	ND

<sup>a</sup>: The apparent transition temperatures and the change in calorimetric enthalpies of the reversible and irreversible processes were determined by integrating the corresponding peak in Fig 8.

<sup>b</sup>: The van 't Hoff enthalpy, the apparent T<sub>m</sub> and the calorimetric enthalpy of the subpeaks (in parenthesis) were determined by deconvolution analysis, fitting the data in Fig 8B to Eq (5).

<sup>c</sup>: The activation energy of the irreversible transition in Fig 8A was determined according to Eq (6). ND is not determined.

doi:10.1371/journal.pone.0158430.t003

in scan 1 were also observed through rescanning the sample, their T<sub>m</sub> being practically the same (Table 3) which indicates that both thermal transitions are reversible. Comparing the T<sub>m</sub> values obtained in scan 1 at 500 mM NaCl (78.8°C and 114.4°C) with the obtained at 150 mM NaCl (72.2°C and 108.5°C), we see that the increase in ionic strength shifts both T<sub>m</sub> to higher temperatures. It appears that the protein has become stabilized by the salt and this effect is also reflected in ΔH<sub>cal</sub> values which are higher at 500 mM NaCl than at 150 mM NaCl (Table 3).

The reversibility of the second calorimetric transition allowed the thermodynamic treatment of the data. As the Cp curve reflects the energetic variations associated with differently populated states in thermodynamic equilibrium with each other, the number of intermediate states and the thermodynamic variables associated with the second peak were calculated by using an appropriate mechanical-statistical deconvolution algorithm [32]. The reversible transitions at each successive scan were subjected to thermodynamic deconvolution analysis and the experimental data were fitted to the Non Two State model with two peaks which accounted for the best fitting (the lowest χ<sup>2</sup> / DoF) according to Eq (5). This indicates that SP-B<sub>N</sub> does not unfold following an all-or-none transition between the native and the denatured form but that some domains in the protein are thermodynamically more stable than others and melt at specific temperatures.

The ΔH<sub>cal</sub>, the van 't Hoff enthalpy (standard enthalpy change of a one-step reaction; ΔH<sub>VH</sub>) and the apparent T<sub>m</sub> for the two subpeaks obtained after the deconvolution process are summarized in Table 3. We must remark that the deconvolution algorithm employed do not take into account the eventual change in molecularity during the process. The enthalpies ratio ΔH<sub>VH</sub>/ΔH<sub>cal</sub> is 1.51 and 1.65 for both subpeaks of scan 1 (Fig 8B and Table 3). These values, above 1, indicate that the cooperativity unit for SP-B<sub>N</sub> unfolding is not the monomer but an oligomer, that is, the real stoichiometry of the protein is higher than we have assumed (the monomer molarity was used in the normalization during data processing). Therefore, intermolecular processes seem to be involved not only in the first but also in the second thermal transition. The average of the ratio values mentioned above: ΔH<sub>VH</sub>/ΔH<sub>cal</sub> ~ 1.58, is fairly close to 1.6 which is the value calculated for the dissociation of a tetramer (n = 4) according to the equation ΔH<sub>VH</sub>/ΔH<sub>cal</sub> = 2n (n+1)<sup>-1</sup>. The second transition would involve tetramer dissociation, coupled or not, to protein unfolding. However, some reservations must be made to the ratio ΔH<sub>VH</sub>/ΔH<sub>cal</sub> as

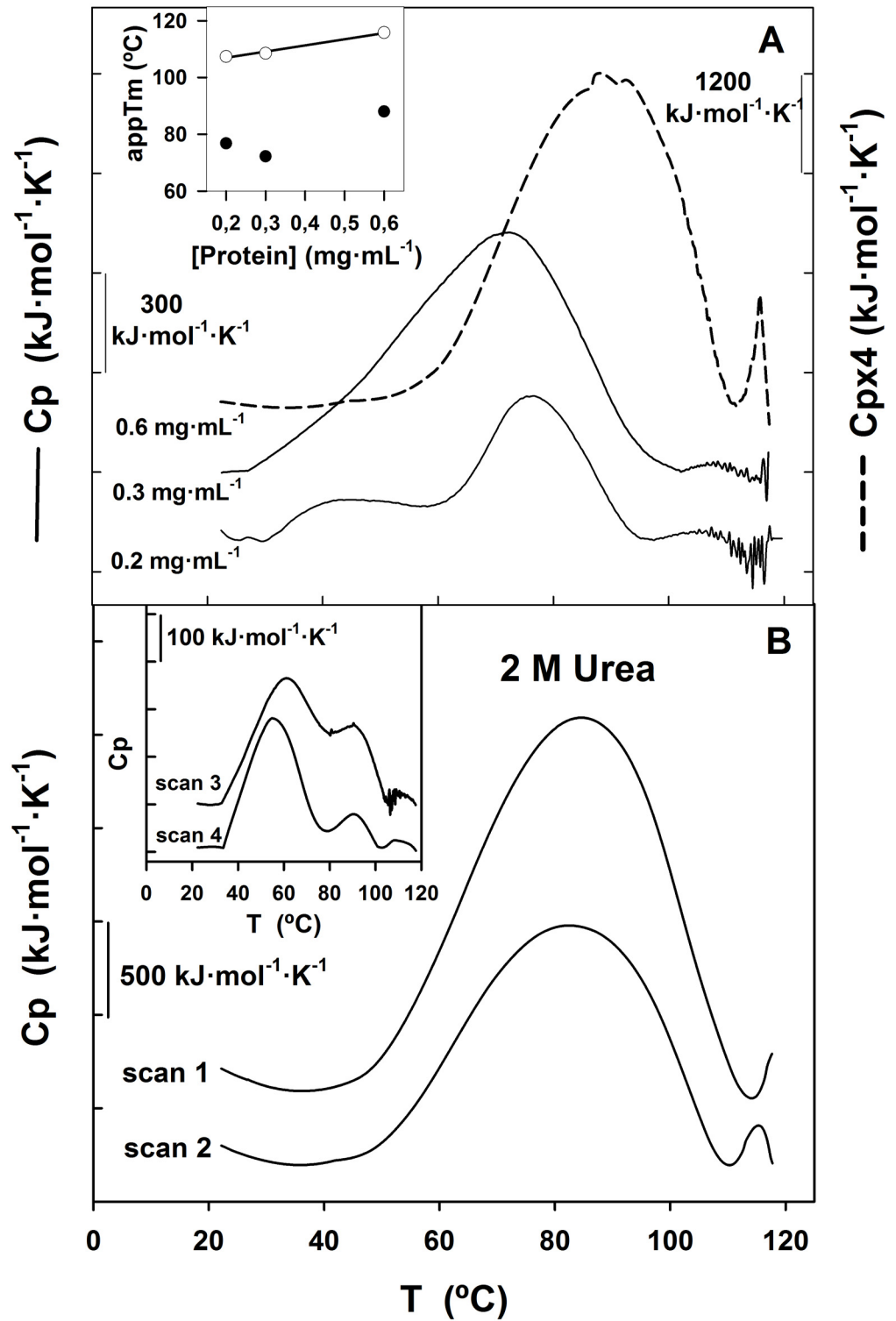
indicative of tetramer dissociation: first, we cannot discard that oligomers holding different molecularity coexist and suffer this thermal transition, and second, the unfolding reaction may also contain small irreversible and not detected steps [33].

Once established the reversible condition of the second transition at 500 mM NaCl, we searched the character of the irreversible condition of the same peak at 150 mM NaCl (see Fig 8A). To elucidate if the irreversibility was due to intermolecular interactions taking place between protein molecules, we studied the dependence of the  $T_m$  on protein concentration as well as the effect of 2 M urea on the  $T_m$  (Fig 9).

The propeptide (0.2 and 0.6 mg·mL<sup>-1</sup>) was subjected to DSC with any other experimental condition as described for 0.3 mg·mL<sup>-1</sup> (Fig 8A) and results obtained were depicted in Fig 9A. The thermogram of 0.2 mg·mL<sup>-1</sup> protein shows an endotherm ( $T_{m1} = 76.8^\circ\text{C}$ ) with a broad shoulder  $\sim 46^\circ\text{C}$  which suggests that oligomers of several sizes are dissociating at temperatures quite apart at low protein concentration. The second peak ( $T_{m2} = 107.4^\circ\text{C}$ ) shows a noisy trace and is followed by an exothermic contribution pointing to protein aggregation. The 0.6 mg·mL<sup>-1</sup> thermogram exhibits two endotherm peaks with  $T_{m1} = 81.7^\circ\text{C}$  and  $T_{m2} = 115.8^\circ\text{C}$  being the sharp profile of this peak quite different to the observed at lower protein concentrations. If the first thermal transition is due to oligomers dissociation and the second only to protein unfolding, the increase in protein concentration would increase only  $T_{m1}$  while it would not affect  $T_{m2}$  [34]. Regarding the first transition in Fig 9A, the  $T_{m1}$  is higher with 0.6 than with 0.3 mg·mL<sup>-1</sup> protein but the different profile the endotherm shows with 0.2 mg·mL<sup>-1</sup> (because of the shoulder) avoids the whole comparison. The  $T_{m2}$  increases linearly with protein concentration (Fig 9A, inset, void circles) indicating that the process includes a bimolecular step and that oligomer dissociation contributes to this thermal transition which may include or not, protein unfolding.

Seeking further confirmation, DSC experiments were carried out with 0.6 mg·mL<sup>-1</sup> protein in the presence of 2 M urea and 150 mM NaCl at 60°C·h<sup>-1</sup>. Mild urea concentrations are usually employed to break intermolecular protein associations without affecting protein unfolding. The position of the  $T_m$  is strongly influenced by the molecularity (n), the  $T_m$  being moved towards higher temperatures as the molecularity decreases [35]. As 2 M urea has a dissociating effect on oligomers, their molecularity would be reduced and  $T_{m1}$  would increase whereas  $T_{m2}$  would be affected only if oligomer dissociation is involved in the second thermal transition. We see in Fig 9B (scan 1) that  $T_{m1} = 84.6^\circ\text{C}$  and thus, it is increased regarding the control value at the same protein concentration ( $T_{m1} = 81.7^\circ\text{C}$ , Fig 9A, dashed line), as was expected. This value is displaced to lower temperatures in the rescans: 82.5°C (scan 2), 61.2°C (scan 3, inset), 54.7°C (scan 4, inset), which indicates that after cooling and reheating, oligomers of higher molecularity are being formed through protein association as it has been proposed for other proteins [36]. Moreover, scan 3 and scan 4 also shows shoulders at 93°C and 91.5°C respectively indicating that oligomers of lower molecularity are being segregating from the bulk in the dissociation process. The  $\Delta H_{\text{cal}}$  of the first transition also decreases regarding control (Fig 9A) with urea, in scan 1, and in the successive rescans (see the scales in Fig 9B).

Regarding  $T_{m2}$ , it could not be calculated in scan 1 since urea displaces the thermal transition towards higher temperatures and the peak is no longer seen. After rescanning the sample, the peak is again in the temperature range of the calorimeter, being  $T_{m2}$ : 115.3°C, 110.5°C and 108.7°C in scan 2, scan 3 and scan 4 respectively. It appears that mild urea has disrupted whatever the oligomer/s contributing to this thermal transition in its absence. Upon successive rescans, the protein would associate forming the oligomers responsible of the second thermal transition, the  $T_{m2}$  being displaced towards the low temperature side of the thermogram. From the experiments with 2 M urea, it follows that the real  $T_m$  of the monomeric propeptide may be higher than 117°C and is not detected in the temperature limits of the calorimeter. Attempts of



**Fig 9. Effect of protein concentration and urea on the thermal behaviour of SP-B<sub>N</sub> followed by DSC.** (A) Thermograms (scan 1) of 0.2 and 0.3  $mg \cdot mL^{-1}$  protein (solid lines) and 0.6  $mg \cdot mL^{-1}$  protein (dashed line) in 5 mM Tris-ClH 150 mM NaCl pH 7 at  $60^{\circ}C \cdot h^{-1}$ . Inset: Dependence of the apparent  $T_{m1}$  (filled circles) and apparent  $T_{m2}$  (void circles) on protein concentration. The line is the fit by linear regression ( $r = 0.993$ ). (B) Thermograms (scan 1 and scan 2) of 0.6  $mg \cdot mL^{-1}$  protein in 5 mM Tris-ClH 150 mM NaCl pH 7 with 2 M urea at  $60^{\circ}C \cdot h^{-1}$ . Inset: scan 3 and scan 4 of the same experiment.

doi:10.1371/journal.pone.0158430.g009

determining the calorimetric behaviour of the monomeric propeptide ( $0.3 \text{ mg}\cdot\text{mL}^{-1}$ ,  $60^\circ\text{C}\cdot\text{h}^{-1}$ ) in the absence of salt, were useless since the protein is prone to aggregate at those conditions.

## Free Thiol Titration

The determination of free thiols in SP-B<sub>N</sub> with DTNB, with or without 5.7 M GdmCl, gave less than and  $0.15 \pm 0$  mol cys per mol of protein indicating that the five possible disulfide bridges have been formed in the propeptide. The titration of the fusion protein, MBP-SP-B<sub>N</sub> gave  $10.22 \pm 0.4$  and  $1.96 \pm 0.04$  in the presence and absence of 5.7 M GdmCl respectively, indicating that no disulfide bond has been formed in the fusion protein although there are two cys accessible to the reagent and eight hidden to it, in the absence of the denaturant. The protein MBP released from the fusion by proteolysis with FXa contains no cys and its titration yielded in fact  $0.53 \pm 0.08$  mol cys per mol of protein. Therefore, the ten cys determined in the fusion belong to the propeptide indicating that whereas SP-B<sub>N</sub> remains attached to MBP, no disulfide bond was formed. MBP possesses a natural protein-binding site, which is employed naturally to interact with proteins involved in maltose signalling and whose relevance to bind the fused polypeptides has been already proposed [37]. It is possible that SP-B<sub>N</sub>, as a partner of MBP under the reduced conditions during purification, become sequestered in a way that prevents the formation of disulfide bridges once the reducing agent was eliminated by dialysis. After cleavage of the fusion by Factor Xa, O/N at room temperature, the propeptide would fold forming 5 disulfide bonds. MBP has been proposed to act as a molecular chaperone in the context of fusion proteins by binding reversibly to aggregation-prone folding intermediates of passenger proteins to prevent their self-association [37]. MBP could then efficiently protect SP-B<sub>N</sub> of undesirable folding and aggregation until, once released from the fusion by proteolysis, the propeptide would be able to fold both independent and properly.

## Discussion

The folding of a protein comes from a balance of forces between the interaction of the protein with itself and the interaction with the medium. The disruption of this balance by denaturing environmental conditions (heat, acid, alkali or organic denaturants among others) is followed by protein unfolding [38]. The folding of small proteins occurs mainly in a highly cooperative fashion (two-state process) without the presence of detectable populated intermediates [39]. In contrast, larger proteins with more than 100 residues in size usually fold in three-state transitions involving populated intermediates containing structured regions corresponding to domains or subdomains in the native state proteins [40]. At equilibrium conditions, a protein molecule can be transformed in one of several compact forms, the highly ordered molten globule (ho-MG, also named N\*), the molten globule (MG), the swollen MG (MG<sub>sw</sub>) and the pre-molten globule (pre-MG, one structural form of the coil state) [3, 25]. They are different structural conformers of the protein but the MG is being considered as the only phase state existing in addition to the native (N) and the unfolded (U) states [4]. The size of the recombinant NH<sub>2</sub>-terminal propeptide of the precursor of SP-B, holding 177 amino acid residues, can sustain folding through a non two-state process. Statistical analysis of 154 proteins carried out by Uversky conclude that the set of intermediate(s) forming proteins shows a mean hydrophobicity of  $0.446 \pm 0.023$  whereas  $0.422 \pm 0.017$  was obtained for the set of proteins unable to adopt a partially folded conformation [41]. The mean hydrophobicity of the propeptide, is 0.456 which is compatible with the propeptide having a theoretical tendency to form intermediate(s) states. Moreover, when both the mean hydrophobicity and the mean net charge (0.0708) of the propeptide are combined, the protein would be located in the area occupied by the proteins folding with intermediate states within the Uversky charge-hydrophobicity phase space [41].

To explore this possibility and to detect the possible existence of intermediate states along the process, urea and GdmCl have been used to follow the unfolding transitions of SP-B<sub>N</sub> at 25°C through changes in secondary and tertiary structures. Employing the phase diagram method, intermediate states at ~ 2 M and 3.5 M GdmCl have been detected showing native-like secondary and altered tertiary structures. They could correspond to the ho-MG (N\*) and MG states respectively as they bind to bis-ANS. Four intermediate states were detected at ~ 2 M, ~ 3 M, ~ 5.5 M and 6.5 M urea, all of them holding more or less altered tertiary structure and binding bis-ANS, which confirms their character of MG-like states. At low and moderate urea, the scenario is more complex since an increase in negative ellipticity points to protein stabilization (intermediates I<sub>1</sub> and I<sub>2</sub>) as other proteins have shown [42]. At higher urea, I<sub>3</sub> shows native-like secondary structure and is compatible with the MG state whereas I<sub>4</sub> shows ~ 80% of the CD signal approaching the pre-MG state, which would be half-way between the MG and the U states and is characterized by holding destabilized secondary structure (~ 50% of the N structure) and weaker interactions with bis-ANS [4].

Unfolding is a thermodynamic process involving a change in free energy and changes in conformation between the native and the unfolded states [43]. To quantify protein stability, the conformational stability is commonly used (the Gibbs energy change;  $\Delta G^0_{\text{H}_2\text{O}}$ ), which is used to compare stabilities of closely related proteins [44]. Relative to the actual validity of  $\Delta G^0_{\text{H}_2\text{O}}$  values obtained from unfolding transitions with chaotropes, it happens that, for an increasing number of proteins, the free energy change derived from hydrogen exchange experiments has been shown to be equivalent to the values obtained from calorimetric data or using spectroscopic probes [45]. The  $\Delta G^0_{\text{H}_2\text{O}}$  of the propeptide is ~ 12.7 kJ·mol<sup>-1</sup> in CD assays (with any chaotrope) or in fluorescence emission assays (with GdmCl), suggesting that chemical induced unfolding deals with a same basic phenomenon and that the distribution of species, which comprises the unfolded state, must be the same in both modes of unfolding. This  $\Delta G^0_{\text{H}_2\text{O}}$  is relatively low compared with the values reported for common globular proteins whose native state appears to be stabilized by 17.7–83.6 kJ·mol<sup>-1</sup> [46] or 21–41.8 kJ·mol<sup>-1</sup> [47]. Thereby, the propeptide is only marginally stable, which may be necessary to achieve its biological function, either chaperoning mature SP-B or potentially acting in the vicinity of membranes. The *m* values (a measure of the cooperativity of the denaturation) reflect the change in solvent exposure of the protein during the transition [47]. As the *m* values are close in CD assays with GdmCl and urea in the N-U and I-U transitions respectively, it follows that changes affecting the backbone of the protein show the same cooperativity disregarding the denaturant whereas those affecting the exposure of trp residues to the solvent with GdmCl are more cooperative since *m* is higher.

The effect of temperature on the propeptide secondary structure detected several thermal transitions, none of them leading to T<sub>m</sub> determination since thermal unfolding was not achieved. Instead, these transitions exhibited midpoints around 68°C at 150 mM NaCl and 46°C, 62°C and 75°C at 500 mM NaCl. It is known that multi-domain proteins usually unfold step-wise, with the domains unfolding individually, either independently or with varying degrees of interactions between them whereas in oligomeric proteins, the subunits dissociation usually precedes the monomer unfolding [48]. Therefore, the thermal transitions observed in SP-B<sub>N</sub> may correspond to the unfolding of certain protein domains or subdomains and/or to oligomers dissociation whereas the overall protein shows resistance to temperature-induced unfolding up to ~ 86°C. DSC studies at 150 mM NaCl evidenced two transitions, one reversible with apparent T<sub>m1</sub> = 72.2°C due probably to oligomers dissociation and perhaps to subdomains unfolding and a second irreversible transition with apparent T<sub>m2</sub> = 108.5°C. Results obtained with different protein concentrations or with urea, point to oligomers dissociation taking place in both transitions. The irreversible character of the process causing the second



peak at 150 mM NaCl disappeared at 500 mM NaCl, suggesting that Na ions could weak the repulsive forces the propeptide holds at neutral pH ( $pI = 4.45$ ) favouring compaction and oligomerization and hindering aggregation at high temperature. The fitting of data in the second transition of the reversible process to Non Two-States with two subpeaks equation confirms the existence of intermediate states in the process and the ratio  $\Delta H_{VH}/\Delta H_{cal} \sim 1.6$  of the subpeaks suggests the dissociation of tetramers, maybe in addition to protein unfolding. Sedimentation data obtained at this salt concentration detected mainly species compatible with dimers and hexamers but the lack of slenderness of the peaks points to other coexisting species whereas crosslinking experiments (S3 Fig) evidenced dimers and trimers.

The high activation energy for the irreversible transition ( $E_A = 703.8 \text{ kJ}\cdot\text{mol}^{-1}$ , Fig 9A and Table 3) indicates that a big energy barrier is present in the propeptide making its unfolding slow and irreversible since the  $E_A$  for many proteins is in the 100–500  $\text{kJ}\cdot\text{mol}^{-1}$  range [49]. The thermal resistance of the propeptide to unfold suggests stabilization by intrahelical disulfide bonds and ulterior confirmation comes from the impossibility of titrating thiol groups in cys residues with or without 5.7 M GdmCl. Thus, the propeptide (S1 Fig) must hold three disulfide bonds in the SAPB domain, by homology with those existing in saposins, which provides them with extreme stability against thermal denaturation and degradation by proteases [16, 17]. The SAPB module of the propeptide has been isolated from rat lung and may display, by homology, a disulfide bond pattern according to the exhibited by saposins [12]. The exact localization of the remaining two bonds in the SAPA domain of the propeptide is still to be determined. The structure of the different intermediates in the folding pathway could reflect different dispositions and compactions of the SAPB module (taking  $\sim 1/2$  of the total sequence of SP-B<sub>N</sub>) with respect to their flanking segments (SAPA module and the connector arm). Denaturation of SP-B<sub>N</sub> induced by urea or GdmCl ends in structures still preserving a substantial amount of secondary structural elements (CD experiments) which could be concentrated in the disulfide-cross-linked SAPB module. This seems to indicate that the disulfide bonds formed according to the expected saposin pattern could provide a surplus of stability to the core (probably the most important part, from a functional point of view, of the protein). In addition, the acidification of the medium causes the formation of a coiled-coil structure with a stretch of the SAPB domain involved in it [15]. Moreover, the presence of salt in the buffer may be important for neutralizing the acidic residues (28%) concentrated in the  $^{148}\text{P-D}^{174}$  stretch of the connector arm, where the protein must hold disorder, for promoting stabilization upon oligomerization.

This complex scenario of protein unfolding is likely related with the role of the N-terminal propeptide to modulate processing and assembly of mature SP-B into pulmonary surfactant complexes, along the pathway of surfactant biogenesis in type II pneumocytes. We propose that the intermediates resolved in the structure of SP-B<sub>N</sub> could reflect true intermediate structures with a functional meaning once processing of proSP-B is triggered in acidic subcellular compartments, in the presence of phospholipid surfactant membranes. Partly folded intermediates may be required to expose disulfide-stabilized saposin domains for assembly into membranes, involving or not changes in the oligomerization state, as a consequence of conformational changes affecting the connector segments. Proteolysis targets are located within those intervening sequences, suggesting that partial unfolding, which could be originated by acidified pH under the physiological context of pneumocytes, could be strictly necessary to trigger maturation and assembly. This is relevant not only to understand the complex process of surfactant assembly *in vivo*, but to optimize the production of surfactant protein SP-B analogues by recombinant strategies, something that still remains elusive. Our results suggest that processing of recombinant propeptides *in vitro*, an apparently limiting step, may need coupling proteolytic maturation with conditions facilitating the acquisition of the partly folded molten

globule-like intermediates that expose both the cleavage sites and the hydrophobic membrane-active modules.

## Supporting Information

**S1 Fig. Scheme of human preproSP-B structure.** Numbers indicate the amino acid position at the beginning and end of the Saposin A-type like module (SAPA) or the B-type (SAPB). Signal peptide encompasses amino acids 1–23, NH<sub>2</sub>-terminal propeptide: 24–200 and mature SP-B: 201–279. Solid lines between cysteine residues indicate disulfide bonds and dashed lines indicate putative disulfide bonds by homology with bonds in saposin modules of preproSaposin. Red cys in mature SP-B serves to dimerize mature SP-B through interchain disulphide bond. Potential glycosylation site in the NH<sub>2</sub>-terminal propeptide (N<sup>129</sup>) is omitted as the protein is produced in bacteria. The second site (N<sup>311</sup>) in the COOH-terminal propeptide is also omitted. (DOC)

**S2 Fig. Prediction of disordered regions in SP-B<sub>N</sub>.** Disordered regions in SP-B<sub>N</sub> have been predicted by Spritz v0.1 [1]. The total % disorder is 25.42, comprising three patches of 4, 22 and 19 amino acids respectively as is shown in the propeptide sequence (D means disorder and O means globular). The bigger patches correspond to connector arm. [1] Vullo A, Bortolami O, Pollastri G, Tosato SC, Spritz A: a server for the prediction of intrinsically disordered regions in protein sequences using kernel machines. *Nucleic Acids Res.* 2006;34: 164–168. (DOC)

**S3 Fig. Time progress of SP-B<sub>N</sub> cross-linked oligomers formation.** The propeptide was incubated with 0.08% (v/v) glutaraldehyde (+GA) for 20 s and 120 s or without it (-GA) at 20°C (control sample, ~ 20 kDa). The positive control was 2 µg of the propeptide fused to the Maltose Binding Protein (F, ~ 62 kDa). To confirm and clarify the SP-B<sub>N</sub> oligomers, we carried out experiments with the bifunctional cross-linking agent glutaraldehyde. SP-B<sub>N</sub> samples (0.26 mg·mL<sup>-1</sup>) in 20 mM Tris-HCl, 500 mM NaCl pH 7.0 were incubated at 20°C with 0.08% (v/v) glutaraldehyde (GA, Amersham) during different times. This is a protein concentration above the concentration employed when only intramolecular crosslinking is desired (~ 0.1 mg·mL<sup>-1</sup>) [1]. The reaction was finished by addition of 5 µL of loading buffer (5x; SDS, reducing conditions) to 20 µL of the sample and 5.2 µg of protein was applied to 12% SDS-PAGE. SP-B<sub>N</sub> was detected by immunoblotting with a primary monoclonal anti-proSP-B antibody (a gift of Dr Weaver, University of Cincinnati, USA) and bound antibody was probed with a secondary anti-Mouse IgG peroxidase-conjugate antibody (Sigma) as described [2]. The recombinant SP-B<sub>N</sub> in absence of GA shows a band of ~ 20 kDa corresponding to the monomer and a trace band migrating as the positive control (fusion protein) which is seen due to deliberate overexposure. Incubation of SP-B<sub>N</sub> with 0.08% (v/v) GA for 20 s and 120 s decreased progressively the ~ 20 kDa band signal whereas new band signals were detected. Those bands were assigned to putative dimer and trimer cross-linked forms of SP-B<sub>N</sub> respectively. In samples exposed to higher GA concentrations or incubated longer times at the same concentration, SP-B<sub>N</sub> was cross-linked to produce higher molecular mass species which either did not enter or remained at the non-linear part of the gel (not shown). Attempts of detecting any SP-B<sub>N</sub> cross-linked species staining the gel with Coomassie, failed due to the small quantities of oligomers entering the gel. By the other hand, lower GA concentrations than 0.08% (v/v) were not effective to crosslink the protein. [1] Ohno H, Kurusu F. Cytochrome C cross-linked with glutaraldehyde. Electrochemical response in poly(ethylene oxide) oligomers. *Chem Lett.* 1996;8: 693–694. [2] Palacios A, González B, Alonso S, Pérez-Gil J, Estrada P. Production of a recombinant form of the propeptide NH<sub>2</sub>-Terminal of the precursor of pulmonary surfactant protein B. *Enzyme*

Microb Technol. 2006;40: 85–92.  
(DOC)

## Acknowledgments

This research has been supported by grants from the Spanish Ministry of Economy and Competition (BIO2012-30733, BIO2015-67930-R), and the Regional Government of Madrid (P2013/MIT-2807). The authors are indebted to Drs. German Rivas from Centro de Investigaciones Biológicas of the Spanish Research Council (CSIC) for their assistance to carry out the analytical centrifugation experiments.

## Author Contributions

Conceived and designed the experiments: ABH JPG PE. Performed the experiments: ABH PE. Analyzed the data: ABH JPG PE. Contributed reagents/materials/analysis tools: ABH JPG PE. Wrote the paper: PE JPG.

## References

1. Dill KA. Perspectives in Biochemistry. *Biochemistry* 1990; 29: 7133–7139.
2. Ptitsyn OB. Molten globule and protein folding. *Adv Prot Chem.* 1996; 47: 83–229.
3. Redfield C, Smith RA, Dobson CM. Structural characterization of a highly-ordered molten globule at low pH. *Nature Struct Biol.* 1994; 1: 23–29. PMID: [7656002](#)
4. Uversky VN. Diversity of equilibrium compact forms of denatured globular proteins. *Prot Pept Lett.* 1997; 4: 355–367.
5. Guo SY, Guo Z, Chen BY, Guo Q, Ni SW, Wang XC. Urea induced inactivation and unfolding of arginine kinase from the sea cucumber *Stichopus japonica*. *Biochemistry (Moscow)* 2003; 68: 1267–1271.
6. Ishimaru D, Lima LMTR, Maia LF, Lopez PM, Bom APA, Valente AP, et al. Reversible aggregation plays a crucial role on the folding landscape of p53 core domain. *Biophys J.* 2004; 87: 2691–2700. PMID: [15298872](#)
7. Weaver TE, Conkright JJ. Function of surfactant proteins B and C. *Ann Rev Physiol.* 2001; 63: 555–578.
8. Lin S, Akinbi HT, Breslin JS, Weaver TE. Structural requirements for targeting of surfactant protein B (SP-B) to secretory granules in vitro and in vivo. *J Biol Chem.* 1996; 271: 19689–19695. PMID: [8702672](#)
9. Brasch F, Johnen G, Winn-Brasch A, Guttentag SH, Schmiedl A, Kapp N, et al. Surfactant protein B in type II pneumocytes and intra-alveolar surfactant forms of human lungs. *Am J Respir Cell Mol Biol.* 2004; 30: 449–458. PMID: [12972403](#)
10. Patthy L. Homology of the precursor of pulmonary surfactant-associated protein SP-B with prosaposin and sulfated glycoprotein. *J Biol Chem.* 1991; 266: 6035–6037. PMID: [1688355](#)
11. Zhai Y, Saier MH Jr. The amoebapore superfamily. *Biochim Biophys Acta* 2000; 1469: 87–99. PMID: [10998571](#)
12. Yang L, Johansson J, Ridsdale R, Willander H, Fitzen M, Akinbi HT, et al. Surfactant protein B propeptide contains a saposin-like protein domain with antimicrobial activity at low Ph. *J Immunol.* 2010; 12: 975–983.
13. Palacios A, González B, Alonso S, Pérez-Gil J, Estrada P. Production of a recombinant form of the propeptide NH<sub>2</sub>-Terminal of the precursor of pulmonary surfactant protein B. *Enzyme Microb Technol.* 2006; 40: 85–92.
14. Bañares-Hidalgo A, Bolaños-Gutierrez A, Gil F, Cabré EJ, Pérez-Gil J, Estrada P. Self-aggregation of a recombinant form of the propeptide NH<sub>2</sub>-terminal of the precursor of pulmonary surfactant protein SP-B: a conformational study. *J Ind Microbiol Biotechnol.* 2008; 35: 1367–1376. doi: [10.1007/s10295-008-0437-3](#) PMID: [18797948](#)
15. Bañares-Hidalgo A, Pérez-Gil J, Estrada P. Acidic pH triggers conformational changes at the NH<sub>2</sub>-terminal propeptide of the precursor of pulmonary surfactant protein B to form a coiled coil structure. *Biochim Biophys Acta Biomembranes* 2014; 1838: 1738–1751.

16. Qi X, Leonova T, Grabowski GA., Functional saposins expressed in *Escherichia coli*. Evidence for binding and activation properties of saposin C with acid  $\beta$ -glucosidase. *J Biol Chem*. 1994; 269: 16746–16753. PMID: [8206997](#)
17. Faull KF, Higginson J, Waring AJ, Johnson J, To T, Whitelegge JP, et al. Disulfide Connectivity in cerebroside sulfate activator is not necessary for biological activity of  $\alpha$ -helical content but is necessary for trypsin resistance and strong ligand binding. *Arch Biochem Biophys*. 2000; 376: 266–274. PMID: [10775412](#)
18. Available: [www.expasy.org/protparam](http://www.expasy.org/protparam).
19. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982; 157: 105–132. PMID: [7108955](#)
20. Sreerama N, Woody RW. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON and CDSSTR methods with an expanded reference set. *Anal Biochem*. 2000; 287: 252–260. PMID: [11112271](#)
21. Laue TM, Shah BD, Ridgeway TM, Pelletier SL. Analytical centrifugation. In: Harding S, Rowe A, Horton J, editors. *Biochemistry and Polymer Science: Royal Society of Chemistry, Cambridge, UK*, 1992, pp. 90–125.
22. Sánchez-Ruiz JM. Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophys J*. 1992; 61: 921–935. PMID: [19431826](#)
23. Ellman GL. A colorimetric method for determining low concentrations of mercaptans. *Arch Biochem Biophys*. 1958; 74: 443–450. PMID: [13534673](#)
24. Greene RF Jr, Pace CN. Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme,  $\alpha$ -chymotrypsin and  $\beta$ -lactoglobulin. *J Biol Chem*. 1974; 17: 5388–5393.
25. Kuznetsova IM, Turoverov K, Uversky VN. Use of the phase diagram method to analyze the protein unfolding-refolding reactions: fishing out the “invisible” intermediates. *J Proteome Res*. 2004; 3: 485–494. PMID: [15253430](#)
26. Gozdek A, Stankiewicz-Drogon A, Poznanski J, Boguszewska-Chachulska AM. Circular dichroism analysis for multidomain proteins: studies of the irreversible unfolding of *Hepatitis C* virus helicase. *Acta Biochem Polonica*. 2008; 55: 57–66.
27. Lima MR, Zingali RB, Foguel D, Monteiro RQ. New insights into conformational and functional stability of human  $\alpha$ -thrombin probed by high hydrostatic pressure. *Eur J Biochem*. 2004; 271: 3580–3587. PMID: [15317594](#)
28. Chapeaurouge A, Johansson JS, Ferreira ST. Folding of a de novo designed native-like four-helix bundle protein. *J Biol Chem*. 2002; 277: 16478–16483. PMID: [11832477](#)
29. Uversky VN, Dunker AK. Understanding protein non-folding. *Biochim Biophys Acta* 2010; 1804: 1231–1264. doi: [10.1016/j.bbapap.2010.01.017](https://doi.org/10.1016/j.bbapap.2010.01.017) PMID: [20117254](#)
30. González JM, Vélez M, Jimenez M, Alfonso C, Schuck P, Mingorance J, et al. Cooperative behaviour of *Escherichia coli* cell-division protein FtsZ assembly involves the preferential cyclization of long single-stranded fibrils. *Proc Nat Acad Sci*. 2005; 102: 1895–1900. PMID: [15684053](#)
31. Menéndez M, Gasset M, Laynez J, López-Zumel C, Usobiaga P, Töpfer-Peterse E, et al. Analysis of the structural organization and thermal stability of two spermadhesins. Calorimetric, circular dichroic and Fourier-transform infrared spectroscopic studies. *Eur J Biochem*. 1995; 234: 887–896. PMID: [8575449](#)
32. Freire E, Biltonen R. Thermodynamics of transfer ribonucleic acids: the effect of sodium on the thermal unfolding of yeast tRNA<sup>Phe</sup>. *Biopolymers* 1978; 17: 463–479.
33. Rösgen J, Hinz HJ. Thermodynamic analysis of heat capacity curves. In: Kemp RB, editor. *Handbook of thermal analysis and calorimetry*, Elsevier, New York, 1999, pp. 65–107.
34. Lyubarev AE, Kurganov BI. Analysis of DSC data relating to proteins undergoing irreversible thermal denaturation. *J Therm Anal Calorim*. 2000; 62: 51–62.
35. Milardi D, Carmelo LR, Grasso D. Theoretical basis for differential scanning calorimetric analysis of multimeric proteins. *Biophys Chem*. 1996; 62: 95–108. PMID: [17029806](#)
36. Burgos I, Dassie SA, Fidelio GD. Thermodynamic model for the analysis of calorimetric data of oligomeric proteins. *J Phys Chem. B* 2008; 112: 14325–14333. doi: [10.1021/jp804465c](https://doi.org/10.1021/jp804465c) PMID: [18939789](#)
37. Fox JD, Kapust RB, Waugh DS. Single amino acid substitutions on the surface of *Escherichia coli* maltose-binding protein can have a profound impact on the solubility of fusion proteins. *Prot Sci*. 2001; 10: 622–630.
38. Schiffer CA, Dötsch V. The role of protein solvent interaction in protein unfolding. *Curr Opin Biotechnol*. 1996; 7: 428–432. PMID: [8768902](#)

39. Chapeaurouge A, Johansson JS, Ferreira ST. Folding intermediates of a model three-helix bundle protein. Pressure and cold denaturation studies. *J Biol Chem*. 2001; 276: 14861–14866. PMID: [11278529](#)
40. Radford SE. Protein folding: progress made and promises ahead. *Trends Biochem Sci*. 2000; 25: 611–618. PMID: [11116188](#)
41. Uversky VN. Cracking the folding code. Why do some proteins adopt partially folded conformations, whereas other don't? *FEBS Lett*. 2002; 514: 181–183. PMID: [11943147](#)
42. López G, Bañares-Hidalgo A, Estrada P. Xylanase II from *Trichoderma reesei* QM 9414: conformational and catalytic stability to chaotropes, trifluoroethanol and pH changes. *J Ind Microbiol Biotechnol*. 2011; 38: 113–125. doi: [10.1007/s10295-010-0836-0](#) PMID: [20838850](#)
43. Nath D, Rao M. Artificial chaperone mediated refolding xylanase from an alkalophilic thermophilic *Bacillus* sp. *Eur J Biochem*. 2001; 268: 5471–5478.
44. Pfeil W. Protein stability and folding: a collection of thermodynamic data. Springer-Verlag, Berlin, 1998.
45. Clarke J, Itzhaki LS. Hydrogen exchange and protein folding. *Curr Opin Struc. Biol*. 1998; 8: 112–118.
46. Tanford C. Isothermal unfolding of globular proteins in aqueous urea solutions. *J Am Chem Soc*. 1994; 86: 2050–2059.
47. Pappa HS, Cass AEG. A step towards understanding the folding mechanism of horseradish peroxidase. Tryptophan fluorescence and circular dichroism equilibrium study. *Eur J Biochem*. 1993; 212: 227–235.
48. Creighton TE. Protein folding. *Biochem J*. 1990; 270: 1–16. PMID: [2204340](#)
49. Fan DJ, Ding YW, Zhou JM. Structural rearrangements and the unfolding mechanism of a trigger factor mutant studied by multiple structural probes. *Biochem Biophys Acta* 2009; 1794: 994–952.