Temperature-Induced Self-Assembly of the Group B Streptococcus (GBS) Fusion Antigen GBS-NN

Fabrice Rose,[†] Silke Roovers,[†,‡] Mathias Fano,[§] Stine Harloff-Helleberg,[†] Jacob J.K. Kirkensgaard,[∥] Kim Hejnaes,[⊥] Per Fischer,[⊥] and Camilla Foged*,[†]

[†]Department of Pharmacy and [§]Bioneer:FARMA, Faculty of Health and Medical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen Ø, Denmark

[‡]Laboratory of General Biochemistry and Physical Pharmacy, Ghent University, Ottergemsesteenweg 460, 9000 Ghent, Belgium

[∥]Niels Bohr Institute, Faculty of Science, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen Ø, Denmark

[⊥]MinervaX ApS, Ole Maaløes Vej 3, DK-2200 Copenhagen N, Denmark

S *Supporting Information*

ABSTRACT: Group B *Streptococcus* (GBS) is a leading cause of serious bacterial neonatal infections worldwide, which provides an unmet medical need for a globally effective vaccine. The recombinant GBS fusion antigen GBS-NN contains the N-terminal regions of the GBS Rib and Alpha C proteins. It shows promising immunogenicity eliciting protective immunity in mice and encouraging results in early human clinical trials. Understanding the physical stability of GBS-NN containing conformational B-cell epitopes is crucial to ensure optimal vaccine stability, efficacy, and safety. We initially discovered that GBS-NN is prone to form higher-order structures at elevated temperatures. We therefore investigated the self-assembly behavior of GBS-NN and characterized the higher-order conformational structures as a function of temperature. In the native state, GBS-NN exists as a monomer and has a secondary structure containing α-helix and β-sheet. Langmuir studies demonstrated that the native protein is highly surface-active and forms a monolayer film at the air−water interface because of its amphipathic properties. The conformational stability of GBS-NN was measured as a function of temperature. GBS-NN has an unusual thermal behavior with a phase transition of approximately 61 °C, which is not accompanied by any major changes in the secondary structure. However, the antigen showed irreversible self-assembly as a function of temperature into higher-order structures with a hydrodynamic diameter of approximately 100 nm. Cryo-transmission electron microscopy analyses demonstrated that these self-assemblies consist of vesicular, ring-like structures with a hollow aqueous interior. Therefore, GBS-NN is a physically stable monomeric protein but is prone to temperature-induced self-assembly above 61 °C.

KEYWORDS: *antigen, biophysics, drug design, protein conformation, protein engineering, protein folding, protein self-assembly, protein stability, Streptococcus, vaccine development*
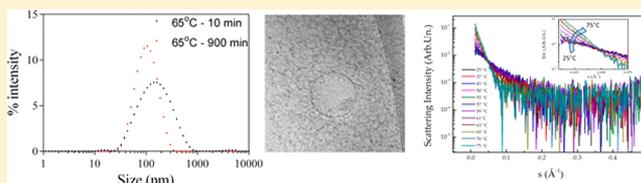
INTRODUCTION

The Gram-positive bacterium group B *Streptococcus* (GBS) is a leading cause of serious neonatal bacterial infections globally. GBS, also referred to as *Streptococcus agalactiae*, is an opportunistic pathogen that colonizes the host via the gastrointestinal and genitourinary tracts, respectively, in up to 25% of pregnant women,[1] and it can cause serious in-uterus infections of the fetus and stillbirths as well as transmission to the neonates during birth.[2] Despite significant improvements in antibiotic prophylaxis and neonatal care, GBS is still the leading cause of neonatal sepsis and meningitis.

Prevention of GBS infection by vaccination is a promising alternative to treatment with antibiotics, and there is ongoing research to develop GBS vaccines.[3] However, vaccine development is challenged by the existence of at least nine antigenically and structurally different GBS serotypes, thus providing a need for the development of a serotype-independent and globally effective vaccine. The serotype of GBS is determined by the capsular polysaccharide (CPS),[4] and CPS-tetanus toxoid conjugate vaccines have been shown to be safe and immunogenic but do not offer cross-protection against other GBS serotypes, which limits their utility.

For these reasons, cell surface proteins of GBS have been investigated as alternative vaccine candidates due to their ability to induce cross-protective immunity.[5,6] Most clinically important GBS strains express the so-called alpha-like proteins (Alp), to which the streptococcal surface Rib protein and the alpha C protein (ACP) belong, providing a unique opportunity for serotype-independent vaccines.[7−9] Immunization of mice with both proteins in combination with Freunds adjuvant and Alum, respectively, conferred promising cross-protective

immunity against a subsequent challenge with GBS.[10,11] The ACP is a virulence factor and an invasin that has been shown to interact with cervical epithelial cells via a glycosaminoglycan (GAG) binding region and an $\alpha_1\beta_1$-integrin binding region, facilitating GBS entry into host cells and subsequent cellular translocation.[12] Full-length ACP consists of an N-terminal domain (NtACP), a repeat region, and a C-terminal domain anchoring the protein to the cell wall. Soluble NtACP has been shown to block cellular entry and contains an integrin binding site.[13] The crystal structure of NtACP has been determined:[14] the NtACP molecule is elongated with overall dimensions of $8.2 \times 3.4 \times 2.7$ nm, and it consists of two major domains (D); a membrane-distal N-terminal β-sandwich domain 1 (D1) and a C-terminal three-helix bundle domain 2 (D2). D1 is composed of eight β-strands arranged into three β-sheets enclosing a hydrophobic core. D2 has three antiparallel α-helices arranged in a 3.5 nm long left-handed three-helix bundle with a highly conserved hydrophobic core.

The recombinant fusion protein antigen GBS-NN, consisting of the nonimmunodominant N-terminal regions of Rib and ACP, was recently designed.[15] The regions are highly conserved among GBS strains, and the fusion protein has been shown to be strongly immunogenic and capable of eliciting antibodies and cross-protective immunity.[15] A vaccine based on monomeric GBS-NN adjanted to Alhydrogel has recently been tested in phase I first in human trials including 240 healthy adult women (NCT ID: NCT02459262). In addition to demonstrating a very favorable safety profile, the vaccine has proven to be highly immunogenic by inducing high levels of functionally active IgG1 and IgA antibodies already 2 weeks after the first dose. Therefore, this vaccine is highly promising, because it is expected that this formulation will confer protection against almost 100% of all GBS isolates.

Recombinant fusion antigens represent an interesting and challenging class of proteins, because their design is based on a specific desired immunogenicity profile, often without taking the structural properties of the constructs into account. This makes a subsequent understanding of their physical stability crucial to ensure optimal (i) vaccine stability during manufacture and storage, (ii) efficacy (immunogenicity), and (iii) safety. This is particularly relevant for antigens with conformational B-cell epitopes, for which it is well-known that immunogenicity is dependent on a specific three-dimensional structure of the antigen. We initially discovered that GBS-NN is prone to form higher-order conformational structures at elevated temperatures. The purpose of this study was therefore (i) to understand the solution properties of the native state of the fusion protein, (ii) to investigate further the temperature-induced aggregation of GBS-NN, and (iii) to characterize the biophysical properties of these higher-order structures.

■ EXPERIMENTAL SECTION

Expression and Purification of GBS-NN. Purified, recombinant GBS-NN (Supplementary Data, Figure S1) was provided by Biovian (Turku, Finland). In brief, the fusion protein was expressed in *Escherichia coli* and kindly provided by MinervaX. Cell extracts were prepared by high-pressure homogenization. The protein was purified by anionic exchange chromatography, followed by hydrophobic interaction chromatography. Finally, the buffer was exchanged to PBS [10 mM sodium phosphate (Sigma-Aldrich) and 0.15 M NaCl (Merck Millipore)], pH 7.2, by gel filtration using disposable PD-10 gel filtration columns (GE Health Care, Little Chalfont, U.K.)

according to the manufacturer's instructions. The purity of the antigen was >95%, as determined by SDS-PAGE and size exclusion chromatography (SEC), respectively. The concentration of GBS-NN was determined by UV spectroscopy at 280 nm by using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, U.S.A.) applying a molar extinction coefficient ε of 30 370 M^{-1} cm^{-1} ($A_{0.1\%}$ = 0.794 mg mL^{-1} cm^{-1}), which was calculated using the online ProtParam tool program (SIB ExPASy Bioinformatics Resources Portal, the Swiss Institute of Bioinformatics, http://web.expasy.org/protparam/). This program was also used to calculate the theoretical isoelectric point (pI) of the protein.

Small-Angle X-ray Scattering (SAXS). Solutions with different concentrations of GBS-NN in PBS were measured at the synchrotron beamline BM-29 (Grenoble, France) at an operating electron energy of 12.5 keV. The scattering patterns were recorded with a 2D image plate detector (Pilatus 1M, Dectris, Baden-Daettwil, Switzerland). The camera was kept under vacuum during data collection to minimize the background scattering. The samples were measured in custom-built sample holders and thermostated to 25 °C with a circulating water bath (Julabo, Seelbach, Denmark). The covered s-range was 0.025 to 5.0 nm^{-1}, and the sample exposure time was 10 times 1 s frame. The 2D scattering data were azimuthally averaged, normalized to the incident radiation intensity and the sample exposure time, and corrected for background and detector inhomogeneities using the software BioXTAS RAW.[19] The radially averaged intensity $I(s)$ is given as a function of the scattering vector s according to eq 1 with the length

$$s = \frac{4\pi}{\lambda} \times \sin(\theta) \tag{1}$$

where λ is the X-ray wavelength, and 2θ is the scattering angle.

The X-ray measurements as a function of temperature were performed by using a SAXSLab instrument (JJXray, Lyngby, Denmark) equipped with a 100XL + microfocus sealed X-ray tube (Rigaku, Tokyo, Japan) producing a photon beam with a wavelength of 1.54 Å. The scattering patterns were recorded with a two-dimensional (2D) 300 K Pilatus pixel area detector (Dectris, Baden, Switzerland). The samples were measured in a temperature-controllable sample stage (Linkam, Tadworth, U.K.). The heating rate was 10 °C/min, and the investigated temperature range was 5−80 °C. The samples were left for 10 min to equilibrate before measuring at each specific temperature step. After cooling to 5 °C, the samples were left for 1 h to equilibrate. The 2D data did not show any angular dependency and were thus azimuthally averaged, normalized by the incident radiation intensity, the sample exposure time, and the transmission, and corrected for background and detector inhomogeneities by using the software SAXSGUI. Data analysis was performed as described above.

Hydrodynamic Radius. The average intensity-weighted hydrodynamic radius (R_h) and the polydispersity index (PDI) of the undiluted protein in PBS were determined by dynamic light scattering (DLS) by using the photon correlation spectroscopy technique. The measurements were performed at 25 °C by using a Zetasizer Nano ZS (Malvern Instruments, Worcestershire, U.K.) equipped with a 633 nm laser and 173° detection optics. For viscosity and refractive index, the values of pure water were used. Malvern DTS v.6.20 software (Malvern Instruments) was used for data acquisition and analysis. Cumulant fitting analysis of the autocorrelation function of

the scattered intensity was used to calculate the R_h. A Nanosphere Size Standard (60 ± 6 nm, Duke Scientific Corporation, Palo Alto, CA, U.S.A.) was used to verify the performance of the instrument.

Circular Dichroism. Far-UV CD measurements were performed in a 0.1 mm quartz cell using an Olis DSM 10 spectrophotometer (Olis, Bogart, GA, U.S.A.). The protein concentration of the samples was 0.2 mg/mL in PBS. All spectra were recorded in the region from 198 to 260 nm using a step size of 1 nm, a fixed bandwidth of 0.5 nm, and a constant integration time of 3 s, eventually resulting in a scanning speed of 5 nm/min. The measurements were performed at room temperature (rt) and at specified temperatures in the range of 37 to 67 °C. All spectra were an average of three scans, and they were background-corrected and transformed into molar ellipticity (θ) on a per residue basis. The shown spectra have not been smoothed.

Fourier Transform Infrared (FTIR) Spectroscopy. Infrared spectra were recorded using a Bomen MB 104IR spectrophotometer (ABB, Bomen, Quebec, Canada). The samples (12 μL, 8.64 mg/mL) were placed in a cell with CaF_2 crystal windows with a path length of 6 μm. For each spectrum, a 256 scan interferogram was collected in the single beam mode with a resolution of 4 cm^{-1} at rt. Background spectra in the region from 1850 to 2200 cm^{-1} and the water vapor spectra in the region of 1500 to 1700 cm^{-1} were subtracted from the protein spectra to obtain a flat baseline, according to previously published criteria.[16] The second derivative spectra were obtained with a 13-point Savitsky-Golay derivative function, and the baseline was corrected with a minimum of points. In addition, all spectra were area-normalized as reported previously[17] in the amide I region from 1595 to 1705 cm^{-1} using the Bomen-Grams software (Galactic Industries, Salem, NH, U.S.A.).

Langmuir Isotherm. A monolayer of GBS-NN was formed at 22 °C by dropwise spreading of a total amount of 4.5 nmol of protein in 60 μL of PBS (pH 7.4) on an aqueous subphase in a KSV Minitrough 1 (KSV Instruments Ltd., Helsinki, Finland) with a total surface area of 243 cm^2 using a Hamilton micro syringe. The aqueous subphase consisted of PBS (pH 7.4). Compression of the monolayer was initiated 20 min after spreading the protein to ensure equilibrium conditions. The monolayer was compressed with a barrier speed of 10 mm/min, and the surface pressure π was measured by using a Wilhelmy platinum plate (KSV Instruments Ltd.). The sample was compressed once in three independent experiments ($n = 3$). KSV software (KSV Instruments Ltd.) was used for data analysis. The phase transition was estimated from the compression modulus (C_S^{-1}) versus π dependency, where C_S^{-1} is defined according to eq 2

$$C_s^{-1} = -A(\mathrm{d}\pi/\mathrm{d}A) \tag{2}$$

where A is the mean molecular area. A characteristic minimum for the C_S^{-1} versus π dependency for the monolayer reflects the phase transition of the monolayer.

Differential Scanning Calorimetry (DSC). The thermal stability of the protein was determined by DSC, which was performed using a NanoDSC (TA Instruments, Lindon, UT, U.S.A.) with a cell volume of 299 μL. Thermograms were recorded in the temperature range of 20 to 90 °C at a scan rate of 1 °C/min, a cell pressure of 3 atm, and an equilibrium time of 600 s between each scan. All samples and references were degassed for 15−20 min immediately before loading. Buffer−

buffer scans of PBS were completed prior to loading of the protein solution (1.2 and 4.8 mg/mL, respectively) in the sample cell. The reversibility of the thermotropic phase behavior was examined by performing a second heating scan after cooling to 20 °C. The NanoAnalyze Data Analysis program version 2.2.0 (TA Instruments) was used to subtract the curve for the second buffer scan from each individual sample scan. The Origin 7 SR2 scientific plotting software (OriginLab Corporation, Northhampton, MA, U.S.A.) was used for baseline correction and data analysis. The calorimetric enthalpy (ΔH_{cal}) is given by the following eq 3

$$\Delta H_{cal} = \int_{T_0}^{T} C_p \mathrm{d}T \tag{3}$$

where C_p is the excess heat capacity. Accordingly, the ΔH_{cal} was calculated as the area under the baseline-corrected C_p curve obtained for each sample. The transition midpoint, T_m, is defined as the temperature at which the transition is half completed, whereas the T_{max} is the temperature at which the C_p is at its maximum.

Aggregation Kinetics. The scattering intensity during temperature-induced GBS-NN aggregation was measured by DLS essentially as described previously[18] employing a Zetasizer Nano ZS with a 633 nm light beam and operating in the back scattering configuration (173°). The position and the attenuation of the laser were fixed at 4.65 mm and 8, respectively. The experiments were performed at different temperatures using a total sample volume of 800 μL in a PMMA UV-Grade cuvette (Kartell, Noviglio, Italy). A total of approximately 300 measurements were recorded for 120 s at a time interval of 180 s. Triplicate measurements were performed at 55, 60, and 65 °C, respectively, to validate the standard deviation of the measurements. The curves were normalized using the following equation (zero is defined as the smallest value of the data set, and one is defined as the largest value of the data set)

$$y = \frac{y - \min(y)}{\max(y) - \min(y)} \tag{4}$$

The initial slope of the different curves was estimated by linear regression of the first part of the curve from $t = 0$ to $t = 20$ min, except for the curves generated at 37 and 45 °C, for which the linear regression was performed using the full curve. The natural log of the slope was plotted as a function of the inverse of the temperature to construct an Arrhenius plot. The activation energy (E_a) was calculated using the Arrhenius equation

$$1/\tau c = A\exp(-E_a/kT) \tag{5}$$

where $1/\tau c$ is the rate constant, A is the pre-exponential factor, k is the Boltzmann constant, and T is the absolute temperature.

Intrinsic Fluorescence and Static Light Scattering (SLS). The intrinsic fluorescence and SLS of the protein were measured using an Avacta Optim 1000 (Avacta Analytical LTd, York, U.K.). Optim-compatible microcuvette arrays were loaded with a volume of 9 μL of protein solution (1.2 mg/mL). The samples were excited with 266 and 473 nm laser beams, respectively, for 800 ms with slits of 50 μm. Fluorescence and SLS signals were measured as a function of temperature in the range of 20 to 90 °C at a rate of 0.3 °C/min, and a hold time of 10 s/well. The spectra were recorded from 240 to 490 nm. Peaks at 266 and 473 nm are specific to SLS.
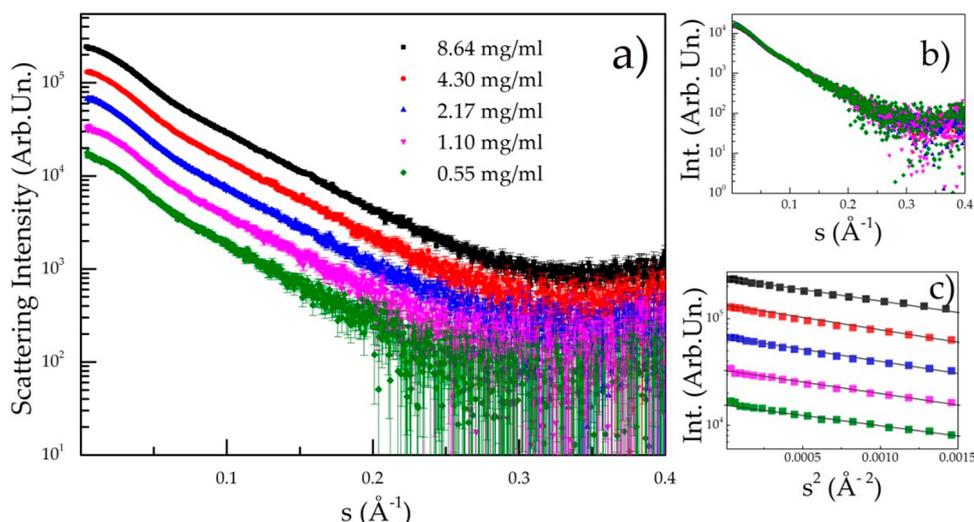
Figure 1. SAXS of the native state of GBS-NN at 25 °C at 0.55 (green), 1.10 (pink), 2.17 (blue), 4.30 (red), and 8.64 (black) mg/mL in PBS. Scattering intensity as a function of *s* at different protein concentrations (a). Data from panel (a) normalized for protein concentration (b). Data show an overlap of the profiles. Guinier region and fittings of the data at different protein concentrations as obtained using PRIMUS (c).

The changes in the fluorescence were determined from the barycentric mean (BCM), which was calculated according to eq 6

$$\mathrm{BCM}(\lambda_{\mathrm{bcm}}) = \frac{\sum_\lambda \lambda I(\lambda)}{\sum_\lambda I(\lambda)} \tag{6}$$

where λ is the wavelength, and I is the intensity. The T_m was calculated by using the second differential, the data was smoothed, and a linear regression curve was fitted locally to the root to reduce the effect of noise in the data. The aggregation onset temperature (T_agg) was defined and calculated as the temperature at which the signal reached a threshold of 10% of the maximum value.

8-Anilino-1-naphthalenesulfonate (ANS) Fluorescence Spectroscopy. A concentration of 50 μM of ANS in PBS was added to a protein solution at 1.2 mg/mL (final concentrations) and equilibrated for 30 min before the measurements. The 266 nm laser beam of the Optim (for 1000 ms with slits of 100 μm) was used for excitation of the ANS, and a spectrum from 290 to 543 nm was recorded in the temperature range of 20 to 90 °C at a rate of 0.3 °C/min and a hold time of 10 s/well. The intensity at 470 nm was plotted as a function of the temperature. A control with ANS alone was measured, and no significant influence on the spectrum was found. The measurements were performed on three wells and were repeated twice. The mean value and standard deviation were calculated for a total of six measurements. The T_m was calculated by using the first derivative of each well.

Thioflavine T (ThT) Fluorescence Spectroscopy. A concentration of 20 μM of ThT in Milli-Q water was added to a protein solution at 1.2 mg/mL (final concentrations) and equilibrated for 30 min before the measurements. The 473 nm laser beam of the Optim (for 1000 ms with slits of 100 μm) was used for excitation of the ANS, and a spectrum from 290 to 543 nm was recorded in the temperature range of 20 to 90 °C at a rate of 0.3 °C/min, and a hold time of 10 s/well. The intensity at 480 nm was plotted as a function of temperature. A control with ThT alone was measured, and no significant influence on the spectrum was found. The measurements were performed on three wells and repeated twice. The T_m was calculated by

using the first derivative of the measurements of each well. The average value and standard deviation were calculated using a total of six measurements.

Transmission Electron Microscopy (TEM). Morphological analysis was carried out using a TWIN transmission electron microscope CM100 with a tungsten emitter (Philips, Eindhoven, The Netherlands). The samples were treated as follows before imaging: a volume of 1000 μL of a GBS solution at 1.2 mg/mL was heated at 55 and at 65 °C, respectively, for at least 24 h. The samples were mounted onto TEM grids (carbon-coated mica surface) and transferred to phosphotungstic acid at 2% (w/v) for 2 min. Blank grids were put on top, and liquid was removed with a filter paper, followed by air drying for 5−10 min.

cryo-TEM. Morphological analysis was also carried out by cryo-TEM using a Tecnai G2 20 TWIN transmission electron microscope (FEI, Hillsboro, OR, U.S.A.). The samples were treated as follows before imaging: two volumes of 1000 μL of a GBS solution at 1.2 mg/mL were heated at 55 and 65 °C, respectively, for at least 2 h. Samples for cryo-TEM were prepared with a FEI Vitrobot Mark V under controlled temperature and humidity conditions within an environmental vitrification system. A small droplet (3 μL) of each sample was deposited onto glow-discharged 300 mesh holey carbon grids (Formvar/Carbon) and spread carefully. Excess liquid was removed resulting in the formation of a thin (10−500 nm) sample film. The samples were immediately plunged into liquid ethane and kept at −174 °C and subsequently transferred in liquid nitrogen to a Gatan Cryo holder connected to the electron microscope. The sample temperature was continuously kept below −180 °C. All observations were made in the bright field mode at an acceleration voltage of 120 kV. Digital images were recorded with a 4 × 4 K CCD Eagle camera from FEI.

Statistics. Statistical calculations were performed by using GraphPad Prism 6 (GraphPad Software, Inc., La Jolla, CA, U.S.A.) by a one-way analysis of variance at a 0.05 significance level followed by means comparison by applying the Tukey's test.
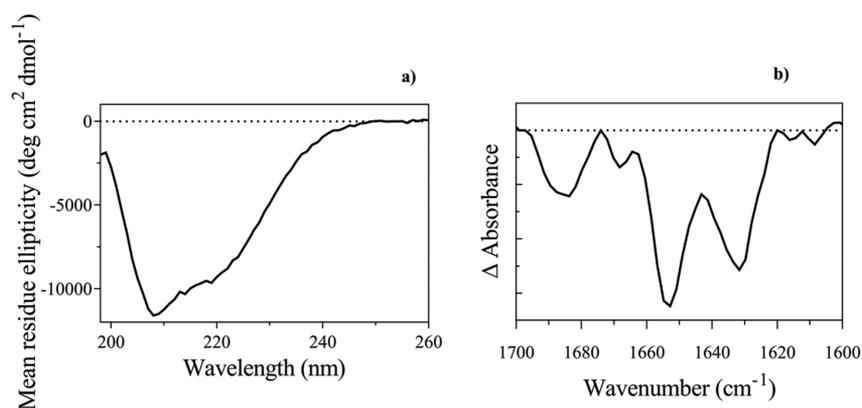
Figure 2. Secondary structure of the native state of GBS-NN. Representative far-UV CD spectrum of GBS-NN (a). Representative FTIR spectrum of GBS-NN (b).
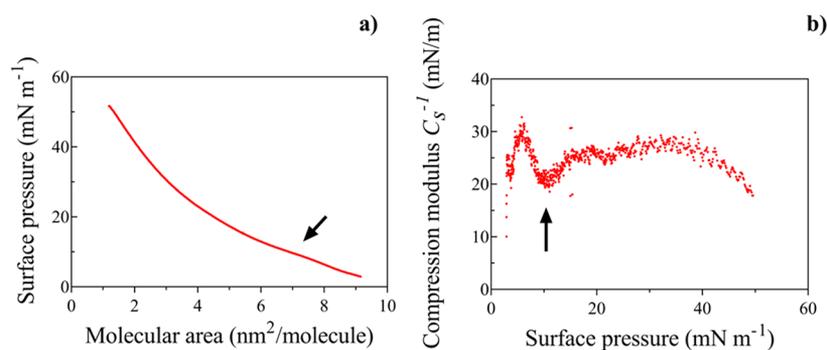
Figure 3. π–A isotherm of a Langmuir monolayer of GBS-NN on a PBS subphase (a). Compression modulus (C_S^{-1}) as a function of surface pressure (b). The arrow indicates the phase transition. The curves represent averages of three experiments.

RESULTS AND DISCUSSION

Native State of GBS-NN. First, we investigated the solution properties of the native state of GBS-NN. The protein consists of 346 amino acids (Supplementary Data, Figure S1), and it has a predicted molecular weight of 38 257.9 Da. The theoretical pI is 4.82, the experimentally determined pI is 4.8 (results not shown), and the estimated charge at pH 7.0 is −12.5. The fusion protein contains no cysteine residues. On the basis of reducing SDS-PAGE (results not shown), GBS-NN appeared as a single band. The mass and identity of GBS-NN were verified using electrospray ionization MS (ESI-MS, results not shown). The MS analyses showed the presence of one species with a molecular weight of 38 257.67 Da, which corresponds closely to the theoretical mass of unmodified, full-length GBS-NN predicted from the sequence of the single polypeptide. The native structure of the protein in solution in the 0.55−8.64 mg/mL concentration range was further characterized using SAXS (Figure 1a). Increasing the concentration did not lead to any interparticle effects, evident from the absence of any distortion in the scattering curve (semilog plot) in the low s region (Figure 1b). When normalized for the concentration values, data perfectly overlapped (Figure 1b), showing that increasing the concentration does not induce any measurable changes in the interactions experienced by the single particle with its neighbors. Moreover, a linear Guinier region $[\ln(I)$ vs $s^2]$ could be isolated for all samples in the low s range (Figure 1c), and for all the curves, a radius of gyration in the range of 3.8−3.9 nm was estimated using PRIMUS.[19,20] This confirms that there is no apparent self-association or intermolecule

interference at increasing protein concentration for the native state of the protein at 25 °C. This is an important observation from a vaccine development point of view, because the native monomeric state of the protein is physically stable in solution during storage (at 4 °C) and handling (at 25 °C). The size of the protein was compared to the R_h measured by using DLS, suggesting an R_h of 3.65 nm (results not shown), which correspond closely to the value measured by using SAXS. This relatively large radius of gyration, as compared to the radius of gyration for globular proteins of the same molecular weight, might be explained by the elongated shape deduced from the crystal structure of NtACP.[14]

The secondary structure of the native protein in solution was investigated by far-UV CD and FTIR spectroscopy. The presence of specific peaks at 208 and 220 nm in the CD spectrum suggests that the main secondary structural element is the α-helical structure (Figure 2a), which was confirmed by FTIR, where a main band at approximately 1650 cm^{-1} shows the presence of an α-helix (Figure 2b). In addition, the protein contains β-sheet structure, evident from the presence of peaks in the amide I frequency regions of 1625−1640 and 1675−1695 cm^{-1}. These results are well in agreement with the previously published crystallographic analysis of NtACP.[14] Near-UV CD showed that the protein has a well-defined tertiary structure in solution (data not shown).

Strong Amphiphilic Properties of GBS-NN. The ACP plays an important role for the entry of GBS into host cells and its subsequent cellular membrane translocation. We therefore hypothesized that GBS-NN might be of amphiphilic nature to facilitate membrane passage. The amphiphilic properties of
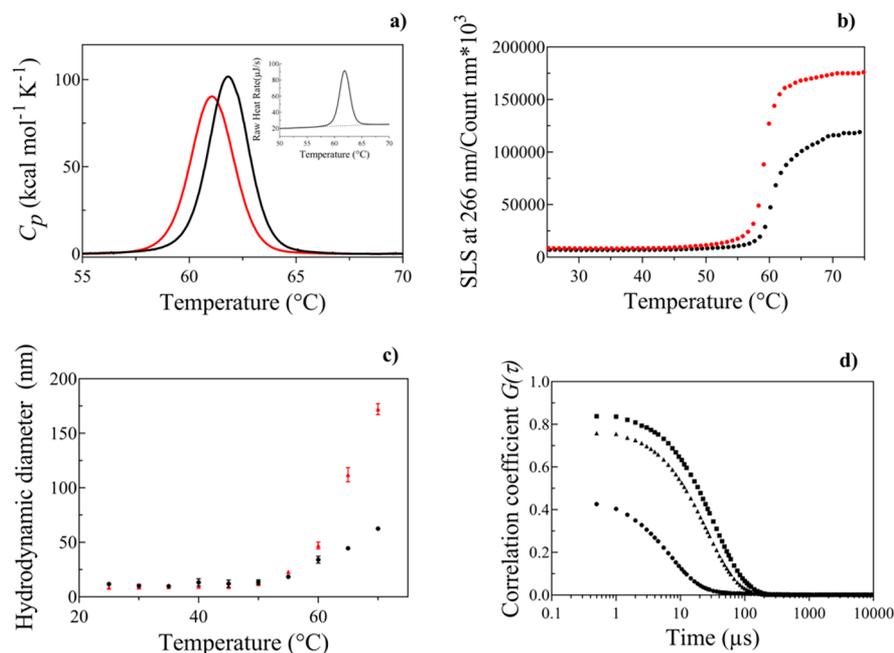
Figure 4. Representative DSC thermograms of GBS-NN at 1.2 mg/mL (black line) and 4.8 mg/mL (red line) in PBS (a). Inset: Raw thermogram of GBS-NN at 1.2 mg/mL before buffer subtraction. The dashed line represents the buffer scan. Static light scattering of GBS-NN in PBS as a function of temperature at 1.2 mg/mL (black dots) and 4.8 mg/mL (red dots) (b). Hydrodynamic diameter of GBS-NN in PBS as a function of temperature at 1.2 mg/mL (black dots) and 4.8 mg/mL (red triangles) measured by using dynamic light scattering (c). Representative correlation functions of data from (c) measured at 50 °C (black dots), 65 °C (black triangles), and 75 °C (black squares) at 1.2 mg/mL (d).

GBS-NN were studied further by examining the adsorption of the protein to the air−water interface by using the Langmuir monolayer technique.[21−23] GBS-NN appeared to possess strong amphiphilic properties, enabling adsorption of the protein to the air−water interface. This was evident from a change in the surface pressure (π), eventually resulting in a characteristic π−A isotherm (Figure 3a), suggesting that the protein has surfactant properties. A plateau was observed at a π-value of approximately 13.7 ± 2.3 mN/m and an A-value of 6.9 ± 0.7 nm^2 (Figure 3a), resulting in a characteristic minimum for the C_S^{-1} versus π dependency for the monolayer (Figure 3b). This plateau might correspond to a transition from a horizontal orientation to a more vertical orientation of the protein at the air−water interface, as previously reported for helical peptides at the air−water interface.[24] The limiting mean molecular area, estimated by extrapolating the isotherm from the plateau (13.7 mN/m) to 0 mN/m was approximately 10.7 nm^2/molecule. If the NtACP is oriented vertically to the interface, the molecular area is approximately 3.4 × 2.7 nm = 9.18 nm^2 according to previously published crystallographic data.[14] Above the surface pressure of the plateau (below approximately 6.9 nm^2/molecule), the isotherm exhibited a steep increase, implying that the phase transition accompanying the change in the molecular orientation occurred at the plateau. As control, the isotherm for the protein ovalbumin measured using similar experimental conditions (data not shown) was comparable to data previously published in the literature.[25,26]

GBS-NN Shows an Unusual Thermal Behavior, Characterized by a Well-Defined Transition Temperature at Approximately 61 °C. To study further the conformational stability of the GBS-NN, we investigated the thermal behavior by DSC (Figure 4a). A transition at 62.1 ± 0.3 °C ($n = 4$) was observed with a ΔH_{cal} of 1047 ± 25 kJ/mol ($n = 4$) at a concentration of 1.2 mg/mL. Moreover, the C_p as a function of temperature did not show any shift in the baseline to a higher value upon completion of the transition (Figure 4a, inset), as usually observed for the denaturation process of proteins, indicating that the ΔC_p is 0. This suggests that the protein does not undergo a classical unfolding process, going from a folded to a denatured protein. The transition was not fully reversible, as two consecutive scans were not identical (Supplementary Data, Figure S2). Interestingly, a small decrease in the T_m from 62.1 ± 0.3 °C ($n = 4$) to 61.1 °C ($n = 1$) was observed when the protein concentration was increased from 1.2 to 4.8 mg/mL (Figure 4a). The thermogram could not be fitted to a simple two-state transition commonly observed for monomeric proteins,[27] which indicates that the transition does not only involve an unfolding process but also additional cooperative processes, e.g., aggregation.[28]

Higher-Order Structures Are Formed during the Thermal Transition. In order to evaluate further such additional cooperative processes, the protein was examined by SLS as a function of temperature (Figure 4b). A sharp increase in the scattering intensity was observed at approximately 57.1 °C for the protein at a concentration of 1.2 mg/mL and at 55.9 °C at a concentration of 4.8 mg/mL. The self-association was confirmed by performing a temperature ramp combined with DLS measurements (Figure 4c). The sample was heated from 25 to 70 °C with 5 °C steps and 10 min of equilibration time in between each step. At 55 °C and above, a clear increase in size (up to 50 nm at 1.2 mg/mL and up to 150 nm at 4.8 mg/mL) was observed. The aggregates had a rather low PDI (0.2), and the corresponding correlograms were very well-defined, suggesting that relatively monodisperse and large aggregates were formed (Figure 4d). When the sample was cooled to 25 °C, clear aggregation was observed (results not shown), confirming the irreversible nature of the transition measured by DSC.
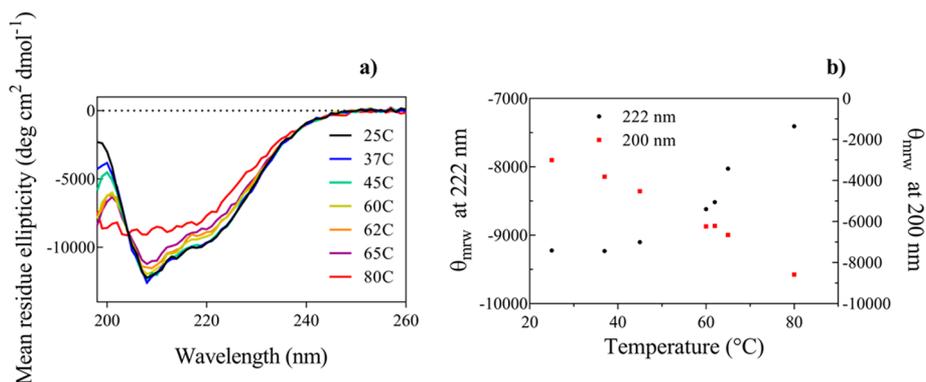
Figure 5. Temperature effect on GBS-NN in PBS (pH 7.2). Secondary structure as a function of temperature for GBS-NN at 1.2 mg/mL in PBS measured by CD. Representative CD spectra of GBS-NN measured at different temperatures (a). CD signals monitored at 200 and 222 nm as a function of temperature (b).
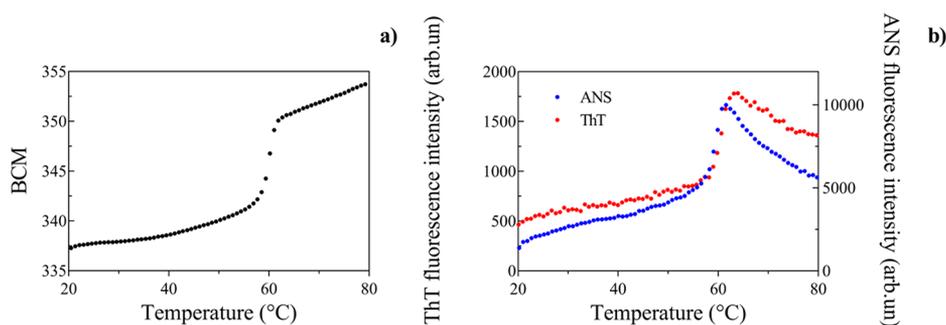
Figure 6. Tertiary structure as a function of temperature for GBS-NN at 1.2 mg/mL in PBS measured by intrinsic fluorescence (a) and extrinsic fluorescence of the dyes ANS and ThT (b).
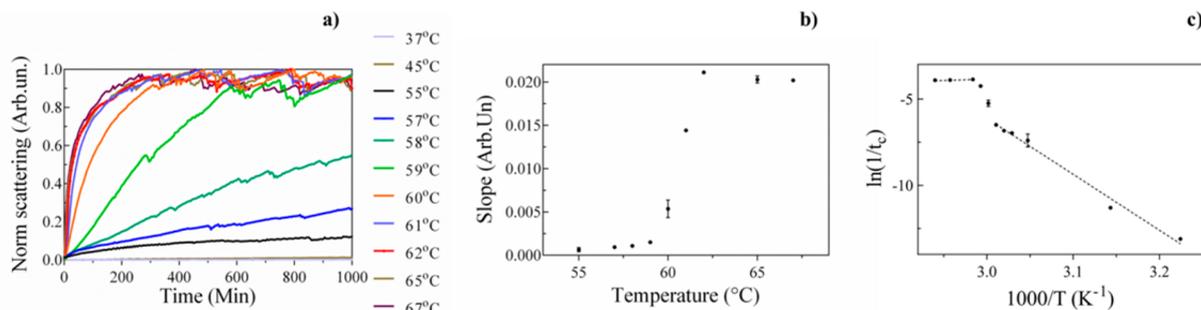
Figure 7. Kinetics of the normalized scattering intensity of GBS-NN (1.2 mg/mL in PBS) at different temperatures measured by DLS (a). Initial slope of the scattering curves as a function of temperature (b). The initial slope of the different curves was estimated by linear regression of the first part of the curve from $t = 0$ to $t = 20$ min, except for the curves generated at 37 and 45 °C, for which the linear regression was performed using the full curve. Arrhenius plot showing the natural log of the slope plotted as a function of the inverse of the temperature (c).

Formation of Higher-Order Structures Is Accompanied by a Partial Change in the Secondary Structure and a Pronounced Change in the Tertiary Structure. The secondary and tertiary structures of the protein were examined as a function of temperature to study further the formation and properties of these higher-order structures. CD measurements showed a clear decrease of the mean residue ellipticity at 222 nm as a function of temperature (Figure 5a,b), which was most pronounced around the transition temperature (approximately 60 °C). However, the main secondary structure remained intact at 65 °C (which is slightly above the main transition measured by using other techniques), and no shift of the wavelength of the maximum absorbance was observed above the transition at 65 °C. In addition, the secondary structure was preserved even at high temperature (80 °C) suggesting that protein unfolding

does not take place. We hypothesize that this might be due to the formation of higher-order conformational structures that stabilize the secondary structure of the protein against heat-induced denaturation. The effect of temperature on the tertiary structure was monitored further by intrinsic fluorescence measurements. The primary structure of GBS-NN contains 2 tryptophan residues and 13 tyrosine residues (Supplementary Data, Figure S1) that fluoresce upon excitation at 266 nm. The BCM was calculated to get the maximum intensity that reflects best any changes in the tertiary structure. At 57 °C and above, a clear shift and increase of the BCM was observed with a T_m of 60.2 ± 0.2 °C at a concentration of 1.2 mg/mL and at 59.3 ± 0.1 °C at a concentration of 4.8 mg/mL (Figure 6a). This suggests an increased hydration of the aromatic amino acid residues. This change in tertiary structure and the accompany-
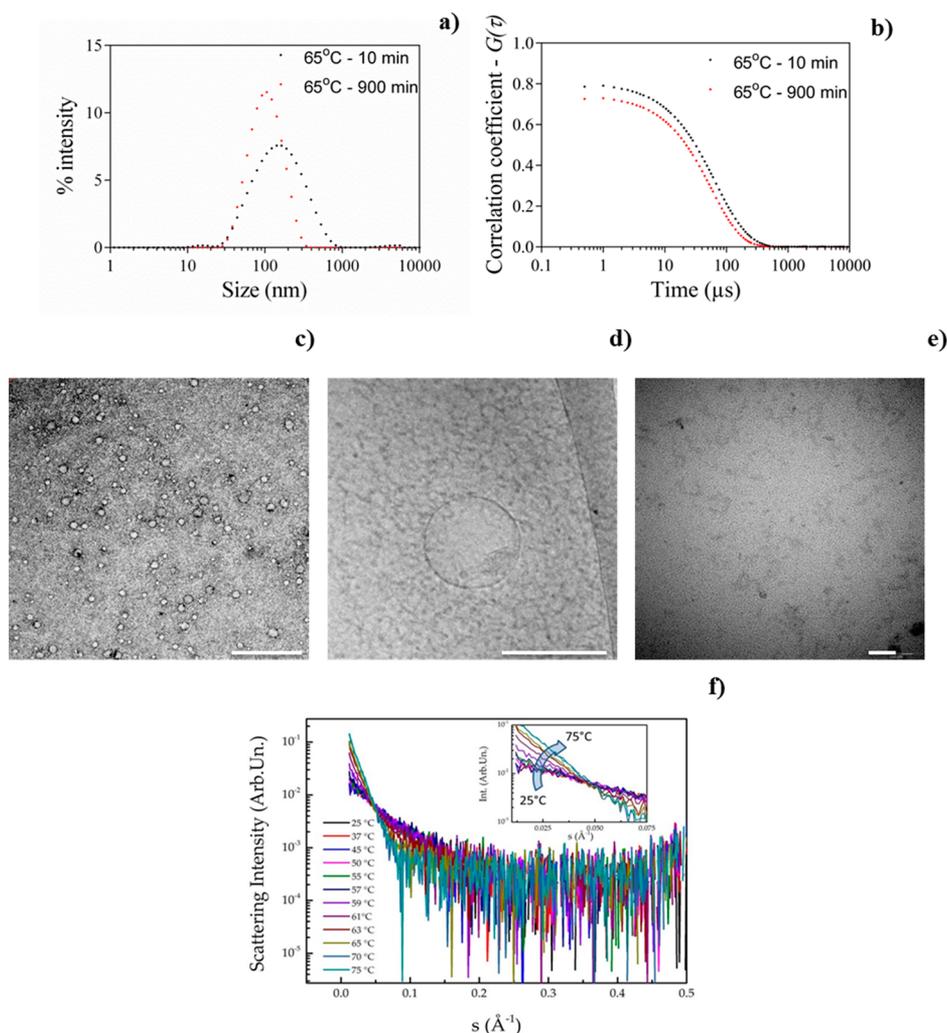
Figure 8. Characterization of GBS-NN self-assemblies by DLS (a,b), TEM/cryoTEM (c−e), and SAXS (f). Representative intensity-based size distribution of GBS-NN preincubated at 65 °C for 10 (black dots) or 900 (red dots) min. Correlation functions of size distributions presented in a (b). Representative TEM (c) and cryo-TEM (d) images of GBS-NN preincubated at 65 °C for 900 min, and representative cryo-TEM image of native GBS-NN (e). Scale bar = 5 μm (c), scale bar = 100 nm (d,e). SAXS curve of GBS-NN as a function of temperature (f).

ing increase in hydrophobicity were also monitored by extrinsic fluorescence measurements by using two specific fluorescent dyes (ANS and ThT, Figure 6b). A pronounced increase of the fluorescence intensity of both dyes was observed at temperatures of 60.3 ± 0.7 and 59.6 ± 0.1 °C, respectively, which can be explained by a higher degree of binding of the dyes to the protein due to increased exposure of hydrophobic regions. These results are therefore well in accordance with the intrinsic fluorescence measurements.

The aggregation kinetics were measured using DLS by recording the scattering intensity at different temperatures (Figure 7a). The kinetic profiles showed a clear change of the kinetics around the transition temperature (approximately 60.5 °C) toward a faster and accelerated aggregation process. Above 62 °C, the aggregation rate appeared constant. An Arrhenius plot was constructed by plotting the initial slopes of the curves in Figure 7a as a function of temperature (Figure 7b,c). A pronounced and unusual nonlinear behavior was observed, and the aggregation rate appeared clearly temperature-dependent, suggesting that the aggregation process takes place via different pathways below and above the T_m. The aggregation process is characterized by three distinct phases, depending on the

temperature (Figure 7c): one phase below 59 °C, one phase in the temperature range of 60−61 °C, and one phase above 62 °C. The curve reaches a plateau after the transition, where a more energetically favorable process takes place with a very low energy of activation. The energy of activation was calculated below and above the transition: below the transition, the activation energy was approximately 270 kJ, and after the transition, the activation energy approached 15 kJ.

Temperature-Induced Aggregation Results in 100 nm Vesicular Self-Assemblies of Ring-Like Structures with a Hollow Aqueous Interior. The aggregated state of GBS-NN was characterized further by DLS by measuring the size of GBS-NN preincubated at 65 °C for 10 and 900 min, respectively (Figure 8a,b). The size distribution of the aggregates was characterized by a single and narrow peak, which was further narrowed by increasing the preincubation time at 65 °C from 10 to 900 min. The average particle diameter was approximately 100 nm after 900 min of incubation. This suggests that GBS-NN monomers self-assemble and stabilize into colloidal structures of a relatively narrow size distribution of approximately 100 nm in diameter at this specific temperature. These structures were visualized using

cryo-TEM and TEM (Figure 8c–e) and appeared as ring-like spherical structures with a hollow, aqueous interior, as previously described for other proteins.[29] Finally, preliminary SAXS analyses confirmed the self-assembly behavior: SAXS curves recorded at different temperatures showed a clear transition toward larger species at approximately 55–57 °C (Figure 8f). This further corroborates our results obtained by DLS as well as intrinsic and extrinsic fluorescence measurements, indicating that the main transition temperature for aggregation is close to 60 °C. However, further advanced investigations of the kinetics of the self-assembly process, as well as the structural characteristics of the self-assembled structures using a high-brilliance synchrotron X-ray source, are pending.

GBS-NN contains multiple hydrophobic domains that are exposed during the transition, which might allow for protein–protein interactions. In addition, the protein has charged domains that might mediate self-assembly via electrostatic interactions. Further studies are needed to identify the specific domains involved in the self-assembly process.

CONCLUSIONS

Our studies show that it is important to address the physical protein stability during manufacture and storage and to define optimal storage conditions for vaccine antigens to meet the quality regulations. It is evident from the current studies that prolonged storage at high temperatures will induce self-assembly of GBS-NN. In addition, exposure of the protein to hydrophobic surfaces and interfaces might cause conformational changes in the protein that can accelerate the self-assembly process. This is important knowledge for the design of high-quality, robust, and stable vaccine dosage forms suitable for long-term storage without change or loss of immunogenicity and compromised safety. The present good manufacturing practice (GMP) batch of GBS-NN appears to be a very well-defined batch suitable for clinical studies, and the planned storage conditions for GBS-NN during manufacture are well below 60 °C (−20 °C for long-term storage of bulk protein and 4 °C for working solutions). Finally, the ability of the N-terminal domains to expose hydrophobic surfaces may explain the apparent ability of the alpha-proteins to facility entry of GBS into epithelial cells, which may likely be important for bacterial pathogenesis.

ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.molpharmaceut.8b00101.

> Amino acid sequence of NN-GBS and second scan of DSC thermogram (raw data) of GBS-NN at 1.2 mg/mL (PDF)

AUTHOR INFORMATION

Corresponding Author
*Phone/Fax: +45 35 33 64 02/+45 35 33 60 01; E-mail: camilla.foged@sund.ku.dk (C.F.)

ORCID Ⓘ
Camilla Foged: 0000-0003-2812-5588

Author Contributions
C.F. conceived the idea for the project, coordinated the study, analyzed the results, and wrote the paper with F.R. F.R. and S.R. conducted most of the experiments and analyzed the results. K.H. and P.F. designed and constructed vectors for expression of GBS-NN. M.F. and K.H. conducted the purification experiments. F.R., S.H.H., and J.J.K.K. designed, performed, and analyzed the SAXS experiments. All authors reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGMENTS

ABBREVIATIONS

ACP, alpha C protein; ANS, 8-anilino-1-naphthalenesulfonate; BCM, barycentric mean; CPS, capsular polysaccharide; cryo-TEM, cryogenic transmission EM; D, domain; DSC, differential scanning calorimetry; DLS, dynamic light scattering; ESI-MS, electrospray ionization MS; FTIR, Fourier transform infrared spectroscopy; GAG, glycosaminoglycan; GBS, group B streptococcus; GMP, good manufacturing practice; Nt, N-terminal; PDI, polydispersity index; R_h, intensity-weighted hydrodynamic radius; rt, room temperature; SAXS, small-angle X-ray scattering; SEC, size exclusion; SLS, static light scattering; TEM, transmission EM

■ REFERENCES

(1) Schuchat, A. Epidemiology of group B streptococcal disease in the United States: Shifting paradigms. *Clin. Microbiol. Rev.* **1998**, *11*, 497−513.

(2) Hansen, S. M.; Uldbjerg, N.; Kilian, M.; Sorensen, U. B. S. Dynamics of Streptococcus agalactiae colonization in women during and after pregnancy and in their infants. *J. Clin. Microbiol.* **2004**, *42*, 83−89.

(3) Johri, A. K.; Paoletti, L. C.; Glaser, P.; Dua, M.; Sharma, P. K.; Grandi, G.; Rappuoli, R. Group B Streptococcus: global incidence and vaccine development. *Nat. Rev. Microbiol.* **2006**, *4*, 932−942.

(4) Kogan, G.; Uhrin, D.; Brisson, J. R.; Paoletti, L. C.; Blodgett, A. E.; Kasper, D. L.; Jennings, H. J. Structural and immunochemical characterization of the type VIII group B Streptococcus capsular polysaccharide. *J. Biol. Chem.* **1996**, *271*, 8786−8790.

(5) Lindahl, G.; Stalhammar-Carlemalm, M.; Areschoug, T. Surface proteins of Streptococcus agalactiae and related proteins in other bacterial pathogens. *Clin. Microbiol. Rev.* **2005**, *18*, 102−127.

(6) Maeland, J. A.; Afset, J. E.; Lyng, R. V.; Radtke, A. Survey of immunological features of the alpha-like-proteins (Alp) of Strepto-coccus agalactiae. *Clin. Vaccine Immunol.* **2015**, *22*, 153−159.

(7) Stalhammar-Carlemalm, M.; Stenberg, L.; Lindahl, G. Protein Rib - A Novel Group-B Streptococcal Cell-Surface Protein That Confers Protective Immunity and Is Expressed by Most Strains Causing Invasive Infections. *J. Exp. Med.* **1993**, *177*, 1593−1603.

(8) Wastfelt, M.; Stalhammar-Carlemalm, M.; Delisse, A. M.; Cabezon, T.; Lindahl, G. Identification of a family of streptococcal surface proteins with extremely repetitive structure. *J. Biol. Chem.* **1996**, *271*, 18892−18897.

(9) Michel, J. L.; Madoff, L. C.; Olson, K.; Kling, D. E.; Kasper, D. L.; Ausubel, F. M. Large, identical, tandem repeating units in the C protein alpha antigen gene, bca, of group B streptococci. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 10060−10064.

(10) Larsson, C.; Stalhammar-Carlemalm, M.; Lindahl, G. Exper-imental vaccination against group B streptococcus, an encapsulated bacterium, with highly purified preparations of cell surface proteins rib and alpha. *Infect. Immun.* **1996**, *64*, 3518−3523.

(11) Larsson, C.; Stalhammar-Carlemalm, M.; Lindahl, G. Protection against experimental infection with group B Streptococcus by immunization with a bivalent protein vaccine. *Vaccine* **1999**, *17*, 454−458.

(12) Baron, M. J.; Filman, D. J.; Prophete, G. A.; Hogle, J. M.; Madoff, L. C. Identification of a glycosaminoglycan binding region of the alpha c protein that mediates entry of group B streptococci into host cells. *J. Biol. Chem.* **2007**, *282*, 10526−10536.

(13) Bolduc, G. R.; Madoff, L. C. The group B streptococcal alpha C protein binds alpha(1)beta(1)-integrin through a novel KTD motif that promotes internalization of GBS within human epithelial cells. *Microbiology* **2007**, *153*, 4039−4049.

(14) Auperin, T. C.; Bolduc, G. R.; Baron, M. J.; Heroux, A.; Filman, D. J.; Madoff, L. C.; Hogle, J. M. Crystal structure of the n-terminal domain of the group B Streptococcus alpha C protein. *J. Biol. Chem.* **2005**, *280*, 18245−18252.

(15) Stalhammar-Carlemalm, M.; Waldemarsson, J.; Johnsson, E.; Areschoug, T.; Lindahl, G. Nonimmunodominant regions are effective as building blocks in a Streptococcal fusion protein vaccine. *Cell Host Microbe* **2007**, *2*, 427−434.

(16) Dong, A.; Huang, P.; Caughey, W. S. Protein secondary structures in water from second-derivative amide I infrared spectra. *Biochemistry* **1990**, *29*, 3303−3308.

(17) Kendrick, B. S.; Dong, A. C.; Allison, S. D.; Manning, M. C.; Carpenter, J. F. Quantitation of the area of overlap between second-derivative amide I infrared spectra to determine the structural similarity of a protein in different states. *J. Pharm. Sci.* **1996**, *85*, 155−158.

(18) Vetri, V.; Leone, M.; Morozova-Roche, L. A.; Vestergaard, B.; Fodera, V. Unlocked concanavalin A forms amyloid-like fibrils from coagulation of long-lived ″crinkled″ intermediates. *PLoS One* **2013**, *8*, e68912.

(19) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2003**, *36*, 1277−1282.

(20) Mertens, H. D.; Svergun, D. I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **2010**, *172*, 128−141.

(21) Norde, W.; Lyklema, J. Why proteins prefer interfaces. *J. Biomater. Sci., Polym. Ed.* **1991**, *2*, 183−202.

(22) Makievski, A. V.; Fainerman, V. B.; Bree, M.; Wüstneck, R.; Krägel, J.; Miller, R. Adsorption of proteins at the liquid/air interface. *J. Phys. Chem. B* **1998**, *102*, 417−425.

(23) Heitz, F.; Van Mau, N. Protein structural changes induced by their uptake at interfaces. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **2002**, *1597*, 1−11.

(24) Kato, N.; Sasaki, T.; Mukai, Y. Partially induced transition from horizontal to vertical orientation of helical peptides at the air-water interface and the structure of their monolayers transferred on the solid substrates. *Biochim. Biophys. Acta, Biomembr.* **2015**, *1848*, 967−975.

(25) Kamilya, T.; Pal, P.; Talapatra, G. B. Interaction of ovalbumin with phospholipids Langmuir-Blodgett film. *J. Phys. Chem. B* **2007**, *111*, 1199−1205.

(26) Pal, P.; Mahato, M.; Kamilya, T.; Tah, B.; Sarkar, R.; Talapatra, G. B. Fibrillation of egg white ovalbumin: a pathway via biomineralization. *J. Phys. Chem. B* **2011**, *115*, 4259−4265.

(27) Freire, E.; Murphy, K. P. Molecular basis of co-operativity in protein folding. *J. Mol. Biol.* **1991**, *222*, 687−698.

(28) Luque, I.; Leavitt, S. A.; Freire, E. The linkage between protein folding and functional cooperativity: two sides of the same coin? *Annu. Rev. Biophys. Biomol. Struct.* **2002**, *31*, 235−256.

(29) Matsuuura, K. Rational design of self-assembled proteins and peptides for nano- and micro-sized architectures. *RSC Adv.* **2014**, *4*, 2942−2953.