Conformational and Aggregational Properties of the Gene 9 Minor Coat Protein of Bacteriophage M13 in Membrane-Mimicking Systems†

M. Chantal Houbiers,* Ruud B. Spruijt, Cor J. A. M. Wolfs, and Marcus A. Hemminga

Department of Biomolecular Sciences, Laboratory of Molecular Physics, Wageningen University and Research Center, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

Received May 18, 1998; Revised Manuscript Received September 14, 1998

ABSTRACT: The membrane-bound state of the gene 9 minor coat protein of bacteriophage M13 was studied in various membrane-mimicking systems, including organic solvents, detergent micelles, and phospholipid bilayers. For this purpose we determined the conformational and aggregational properties of the chemically synthesized protein by CD, FTIR, and HPSEC. The protein appears to be in a monomeric or small oligomeric α-helical state in TFE but adopts a mixture of α-helical and random structure after subsequent incorporation into SDS or DOPG. When solubilized by sodium cholate, however, the protein undergoes a transition in time into large aggregates, which contain mainly *β*-sheet conformation. The rate of this *β*-polymerization process was decreased at lower temperature and higher concentrations of sodium cholate. This aggregation was reversed only upon addition of high concentrations of the strong detergent SDS. By reconstitution of the cholate-solubilized protein into DOPG, it was found that the state of the protein, whether initially α-helical monomeric/oligomeric or *β*-sheet aggregate, did not change. On the basis of our results, we propose that the principal conformational state of membrane-bound gene 9 protein in vivo is α-helical.

M13 is a filamentous bacteriophage which consists of a circular single stranded DNA molecule surrounded by a cylinder of coat proteins. The particles are 880 nm in length and 6 to 7 nm in diameter and infect strains of *Escherichia coli* bearing F pili. The molecular mass of a particle is about 1.6×10^7 Da, of which 88% is protein and 12% is DNA (*1*). There are about 2700 copies present of the major coat protein, the product of gene 8. The minor coat proteins, present in 3−5 copies each, are located at the ends: One end contains the gene 7 protein (3630 Da) and gene 9 protein (3681 Da). The other end contains the gene 3 and gene 6 proteins (for reviews see refs *2*−*6*).

During the membrane-bound phage assembly process, the coat proteins as well as three phage encoded nonstructural proteins are involved, the proteins encoded by genes 1, 4, and 11. Bacterial proteins are also involved, one of which is thioredoxin. During the process of assembly the phage DNA sheds the gene 5 proteins and picks up the coat proteins while extruding through the cellular envelope membranes (for reviews see refs *2*−*8*).

Whereas a lot of information has become available about the structure of the major coat protein in the phage particle (*9*−*11*) and in its membrane-bound state (*12*−*19*), the information about the minor coat proteins encoded by genes 7 and 9 is scarce and incomplete. The primary structures were deduced from the nucleotide sequence of the M13 DNA genome (*20*). Grant (*21*) suggested that these proteins are located at one end of the phage particle. This end appears to be the protein part that leaves the cell first (*22*).

On the basis of the finding that the gene 7 and gene 9 proteins retain their amino-terminal formyl group (*23*), they were proposed to be inserted into the membrane during or directly after synthesis. This suggests that the gene 7 and gene 9 proteins are incorporated in phage as primary translational products. These proteins, therefore, behave differently from the gene 8 protein, which is synthesized as a preprotein (*24*) and is finally processed to yield the membrane-bound coat protein (*25*). It was shown by subcellular location experiments that both gene 7 and gene 9 proteins are fractionated with the *E. coli* inner membrane (*26*).

The finding that the gene 7 and gene 9 proteins are located at the end of the phage particle that emerges first (*22*) shows that they must play a role early in assembly. It was found that deletions in the packaging signal could be compensated by mutations in gene 7, gene 9, and gene 1 (*27*). This suggests an important role for the products of these genes in initiating the assembly process by interacting with the packaging signal of the phage genome. Furthermore, absence of gene 7 or gene 9 proteins almost completely abolishes the production of phage (*2*). A complete description of assembly requires a detailed knowledge of both cellular localization and ultimate location of the coat proteins in the phage and functioning of all components during the assembly process. Until now, an assembly model was based on assumptions regarding the minor coat proteins (*2, 7, 28*).

In this study we will focus on the properties of the gene 9 protein. The primary structure of the gene 9 protein (*20*) and its hydropathy pattern, shown in Figure 1, indicates the

† This research was supported by the Life Sciences Foundation (SLW) with financial aid of The Netherlands Organization for Scientific Research (NWO).

* Corresponding author. Telephone: 31-317-484326. Fax: 31-317-482725. E-mail: chantal.houbiers@virus.mf.wau.nl.

10.1021/bi981149e CCC: $18.00
Published on Web 12/29/1998

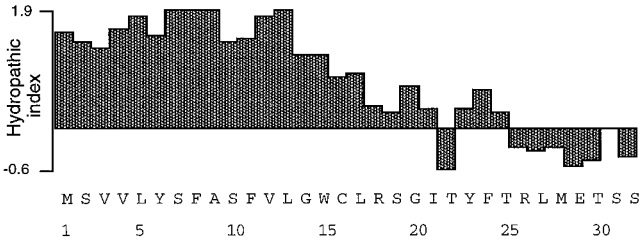

FIGURE 1: Hydropathy analysis of the gene 9 amino acid sequence according to the hydropathy scale of Kyte and Doolittle (*55*), using a sliding window of 9. The N-terminal methionine is assumed to be formylated (*23*).

presence of a hydrophobic stretch, long enough to span the membrane in an α-helical way. The gene 9 protein has no charged residue at the N-terminus, whereas the C-terminus contains mainly charged and polar residues.

Our work, therefore, is aimed to characterize the membrane-bound state of the gene 9 protein. We studied the conformational and aggregational properties of the protein in several membrane-mimicking systems, proceeding from organic solvents to detergent micelles and phospholipid bilayers.

MATERIALS AND METHODS

Materials. Fmoc-amino acids[1] and the solid support *p*-alkoxybenzyl alcohol resin (Wang resin) were purchased from Bachem (Bubendorf, Switzerland). The compounds 1-hydroxybenzatriazole, diisopropylcarbodiimide, and ethanedithiol (EDT) were obtained from Fluka (Buchs, Switzerland). Dichloromethane (DCM), dimethylformamide (DMF), isopropyl alcohol, piperidine, trifluoroacetic acid (TFA), and sodium dodecyl sulfate (SDS) were obtained from Merck (Darmstad, Germany), and 2,2,2-trifluoroethanol (TFE) was obtained from Acros (Pittsburgh, PA). 1,2-Dioleoyl-*sn*-glycero-3-phosphoglycerol (DOPG) was obtained from Avanti Polar Lipids (Alabaster, AL). Sodium cholate was from Sigma (St. Louis, MO). The Superose 12 and the Superdex 75 columns were obtained from Pharmacia (Uppsala, Sweden), and the P6-DG column was from BioRad (Hercules, CA).

Chemical Synthesis and Purification of the Protein. The gene 9 protein was synthesized by Mr. P. J. H. M. Adams at the Peptide Department of the Laboratory for Organic Chemistry of the University of Nijmegen. The cysteine residues in the protein were protected as disulfides with the *S-tert*-butyl group. The protein was prepared on a Wang resin using a solid-phase technique with a semiautomatic peptide synthesizer (Labortec SP640) according to the Fmoc protocol. Apart from the carrier, the steps in each cycle were performed as described previously (*29*). The removal of α-amino protective groups and the efficiency of each new acylation was controlled with the Kaiser test (*30*); resin grains were visually judged by observing them after the test against a white background. Even a slight coloration could be detected in this way, and the pertinent cycle was repeated, if coloration was observed. Each cycle, irrespective of the result of the Kaiser test, was terminated by treatment with a capping mixture (Ac_2O/diisopropylamine). After completion of the required number of acylations, each protein−resin adduct was suspended in the mixture TFA−EDT−water (38:1:1) for 4 h and filtered. The filtrate was added to the 10-fold volume excess of absolute ether to precipitate the crude S-protected peptide trifluoroacetate. It should be noted that disulfides are not attacked by ethanedithiol in acidic media (*31*). The product was subjected to amino acid analysis, confirming that all amino acids were present. The protein was further checked by gel electrophoresis, HPLC, and mass spectrometry.

Solubilization of Protein in Organic Solvent. The protein was first dissolved in a small volume of TFA (10−20 μL/mg of protein) and dried under a stream of nitrogen. Next, the protein was washed with TFE, followed by evaporation of the TFE under a stream of nitrogen, to remove residual TFA. Finally the protein was dissolved in a known volume of TFE. The protein concentration was determined by the procedure of Peterson (*32*) with bovine serum albumin as a standard.

Solubilization of Protein in Detergent. To solubilize the gene 9 protein in detergent (sodium cholate or SDS), the method described by Killian (*33*) was used. Briefly, the desired amount of gene 9 protein dissolved in TFE was added to an equal volume of water, containing the desired amount of detergent. The mixture was vortexed for a few seconds. Next, this mixture was diluted to yield a 16:1 ratio of water to TFE by volume. The samples were mixed by vortexing for 2 s and lyophilized. The sample was rehydrated in the desired volume.

Determination of the Aggregation State of the Protein. The aggregational state of the gene 9 protein was determined using high-performance size exclusion chromatography (HPSEC) on a Superose 12 column (1.0 × 30 cm) or a Superdex 75 column (1.0 × 30 cm). A flow of 0.4 or 0.5 mL/min was applied, eluting in 50 mM cholate, 10 mM sodium phosphate, pH 8.0, for all samples but the SDS sample. The latter was eluted with 25 mM SDS, 10 mM sodium phosphate, pH 8.0. On-line detection was done by measuring the absorption and fluorescence. The absorption was measured at 280 nm and emission at 340 nm, with an excitation wavelength of 280 nm. For molecular weight calibration of the columns, alcohol dehydrogenase (150 000 Da), bovine serum albumin (67 000 Da), myoglobine (18 800 Da), A-protinin (6500 Da), and bacitracin (1400 Da) were used. The aggregational state of the gene 9 protein in TFE was checked by passage over a nanosep filter, with a molecular weight cutoff of 3 kDa, from Pall Filtron (Northborough, MA). The filtrate and residue were tested for the presence of protein by tricine−sodium dodecyl sulfate−polyacrylamide gel electrophoresis (*34*).

Fourier Transform Infrared Spectroscopy. For the FTIR experiments the gene 9 protein stock solution in TFE was treated by an additional step to remove traces of TFA, since TFA contributes to the signal in the amide I region (*35*). This was done by adding a 2 mM HCl solution to the TFE stock in a 1/1 (v/v) ratio, to yield a 1 mM HCl concentration in the protein solution. The solution was subsequently lyophilized, redissolved into TFE, and used for further sample preparation. Samples for FTIR typically contained 0.1−0.2 mg protein at a cholate/protein molar ratio of 60, or a lipid/

[1] Abbreviations: CD, circular dichroism; DOPG, 1,2-dioleoyl-*sn*-glycero-3-phosphoglycerol; EDT, ethanedithiol; Fmoc, 9-fluorenylmethoxycarbonyl; FTIR, Fourier transform infrared; HPSEC, high-performance size exclusion chromatography; L/P, lipid to protein molar ratio; RT, room temperature; SDS, sodium dodecyl sulfate; TFA, trifluoroacetic acid; TFE, trifluoroethanol.

protein ratio of 50. Dehydrated films were prepared on a CaF_2 window (13 mm diameter) in a cabin continuously purged with dry air. FTIR spectra were recorded on a Perkin-Elmer 1750 FTIR spectrometer at room temperature. The optical bench was purged with dry air (Balston, Maidstone Kent, U.K.) at a flow rate of 25 L min^{-1}. The acquisition parameters were the following: 4 cm^{-1} resolution; 128 coadded interferograms; 3500−900 cm^{-1} wavenumber range. For the sake of comparison of the samples with cholate, the spectral intensities were normalized with the cholate peak intensity at 1560 cm^{-1}.

Circular Dichroism. CD spectra were recorded from 190 to 260 nm on a JASCO 715 spectrometer at 25 °C, using a 1-mm path length cell, 1-nm bandwidth, 0.1-nm resolution, and 1s response time. Spectra were corrected for the background with an equally prepared sample without the protein. Protein concentrations were within the range of 25 to 46 μM and are mentioned in the figure legends.

Reconstitution of Protein into Lipid Vesicles. Reconstitution of the protein into lipid bilayers was performed by two different methods: (1) using cholate and subsequent dialysis; (2) via cosolubilization in TFE/water. In both cases the desired amount of DOPG was taken from a chloroform solution. The chloroform was evaporated under a stream of nitrogen, and the remaining lipid film was kept overnight under vacuum to remove all traces of chloroform. In the first method, the lipid film was dissolved in 100 mM cholate and mixed with cholate solubilized protein (prepared according to the method described above under Solubilization of Protein in Detergent). Reconstitution was carried out by removal of cholate by dialysis for 48 h (4 × 12 h) at room temperature, according to Spruijt et al. (*36*), using a dialysis membrane with a molecular weight cutoff of 12−14 kDa. The second method is equal to the method of solubilization of protein into detergent, as described above, except that, instead of detergent, now DOPG is used. After rehydration the vesicles were sonicated on ice for 3 min with a Branson B15 sonifier to minimize scattering by the vesicles in the CD measurements. Titanium particles of the sonifier tip were removed by centrifugation.

RESULTS

Characterization of Gene 9 Protein in TFE, Detergents, and Lipids. The gene 9 protein could be dissolved in TFE and a mixture of TFE/water (1/1, v/v), but the protein was insoluble in water and strictly hydrophobic solvent, such as chloroform (as observed from visual turbidity, due to the presence of large precipitates). Completely dissolved gene 9 protein in TFE was used to be solubilized in detergents or phospholipids via cosolubilization in TFE/water (1/1, v/v) and subsequent dilution, lyophilization, and rehydration (*33*).

In Figure 2, CD spectra are presented of gene 9 protein in TFE, sodium cholate, SDS, and DOPG. Spectra were recorded immediately after sample preparation via the cosolubilization method. The CD spectrum of protein in TFE/water (1/1, v/v) was equal to the spectrum in pure TFE (data not shown). The spectrum in TFE features two minima at 208 and 222 nm, indicating the presence of α-helical structure (*37*). Also the spectra in SDS, cholate, and DOPG are characterized by these two minima. However, the minima are not so well-pronounced, slightly merging into a single broad band. Furthermore, the intensity is much lower than

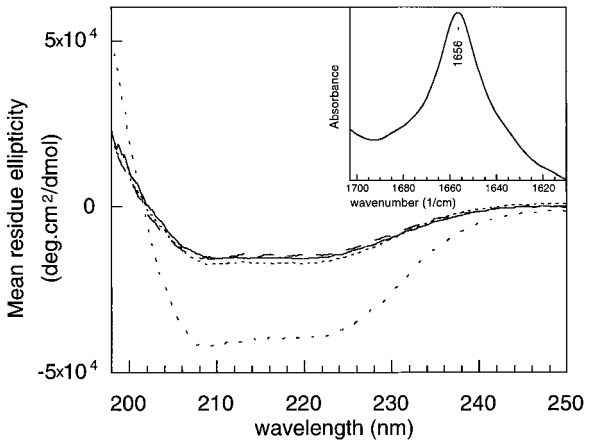

FIGURE 2: CD spectra of the gene 9 protein in TFE (- - -) and incorporated via the cosolubilization method in TFE/H_2O into cholate (•••), SDS (— —), and DOPG (solid line), directly after sample preparation. Protein concentrations were 46 μM in pure TFE, 46 μM in 50 mM SDS/10 mM phosphate/pH 8.0, 35 μM in 26 mM cholate, and 24 μM in 12 mM DOPG. Insert: FTIR absorbance of dehydrated film of gene 9 protein incorporated into DOPG, at an L/P ratio of 50.

in TFE. Additionally, the FTIR spectrum of gene 9 protein in DOPG was recorded. The amide I band, shown in the insert of Figure 2, contains a single band at 1656 cm^{-1}. A band around this position is generally found for α-helical structures (*38*), but also random structure gives rise to a band around this position. The decreased CD intensity in detergents and phospholipids suggests that the amount of α-helix is lowered in these systems, as compared to TFE. Probably part of the α-helical conformation, present in TFE, has converted into another structure. However, from the FTIR spectrum in DOPG it is clear that no β-sheet structure is present. Therefore, it is likely that the remaining structure is random structure, in agreement with the observed FTIR spectrum. Thus, the conformation of gene 9 protein in detergents or phospholipids is a mixture of α-helix and random coil. Assuming that in TFE the protein is 100% α-helical, we roughly estimated the amount of α-helix in these systems, resulting in 50−60% ± 10% α-helix in SDS, cholate, and DOPG. The remaining part is supposed to be random coil. Such a large contribution of random coil is not surprising, since in such a small protein a relatively large part of the amino acids is close to one of the two ends, and terminal amino acid residues are generally less ordered.

The aggregational states of the protein in the samples, shown in Figure 2, were checked by HPSEC or ultrafiltration. The protein dissolved in TFE was observed to pass a 3 kDa filter, leaving no residue on the filter, indicating the predominant presence of protein monomers. Gene 9 protein solubilized by SDS and analyzed using HPSEC in 25 mM SDS showed a single peak, corresponding to monomeric protein (data not shown). Gene 9 protein solubilized by cholate and reconstituted into DOPG, subjected to HPSEC in 50 mM cholate, eluted as a single broadened peak at a position indicating an average size of trimers. Since estimation of the protein size by this method is not sufficiently accurate to discriminate between monomers or small oligomers, we will refer to this state as "monomeric/oligomeric state", bearing in mind that the protein may be present as monomer or small oligomers.

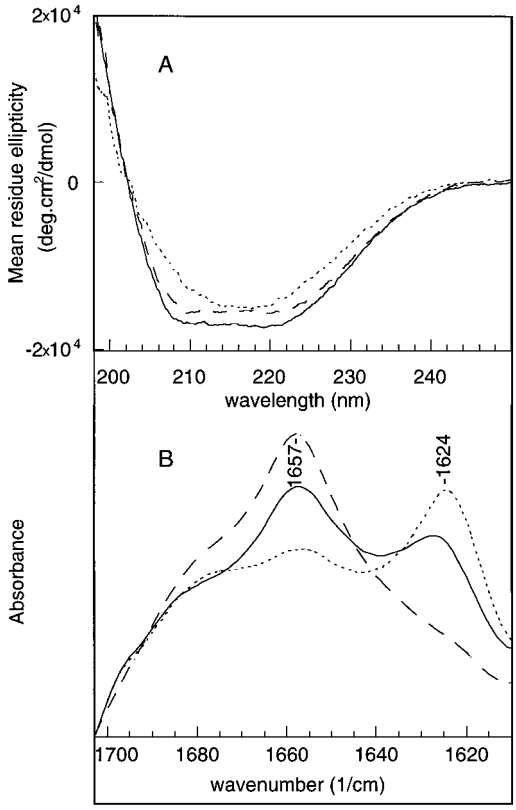

FIGURE 3: (A) CD spectra of gene 9 protein in cholate at times zero (— —) and 3 days after sample preparation (- - -) and after addition of SDS to the sample at time 3 days (solid line). These samples were made from the (more concentrated) FTIR samples of Figure 3B (protein concentration 1.4 mM, cholate 83 mM), but diluted with a cholate solution just before the CD measurement, resulting in 35 μM protein in 26 mM cholate. The FTIR sample with 50 mM SDS added was diluted with an SDS solution, resulting in 26 mM SDS. (B) FTIR spectra of dehydrated film of gene 9 protein in cholate at times zero (— —) and 3 days (- - -) after sample preparation and after addition of SDS to the sample at time 3 days (solid line). The cholate/protein ratio was 60; the SDS/protein ratio was 36. The spectra were not corrected for background of cholate, giving rise to a broad band underneath protein peaks.

Change in Conformational and Aggregational State of the Cholate-Solubilized Protein. After 3 days, the conformational and aggregational states of the gene 9 protein in TFE, cholate, SDS, and DOPG were checked again by CD and HPSEC, to determine the stability of the protein. For the protein in TFE, SDS, and DOPG the results were similar (data not shown), indicating a stable state of the protein in these systems. It should be mentioned that the CD spectra of the protein in TFE or TFE/water (1/1, v/v) did not even change after weeks (data not shown). In contrast, the CD spectrum of the cholate-solubilized gene 9 protein was significantly changed after 3 days (Figure 3A). Instead of two minima at 208 and 222 nm, a single minimum around 218 nm was observed, while the zero-crossing and intensity hardly changed. To analyze the protein conformation in more detail, also FTIR spectra of the cholate-solubilized gene 9 protein were recorded. In Figure 3B, the amide I bands of the corresponding samples are shown. Directly after preparation the most intense absorbance is at 1657 cm^{-1}, suggesting a predominant mixture of α-helical and random structure (*38*), in agreement with the CD result. There is also some intensity present around 1680 cm^{-1}, possibly originating from turnlike structures or distorted α-helices (*38*). The spectral changes after 3 days are shown in Figure 3B: The absorbance at 1657 cm^{-1} has diminished, in favor of a new peak at 1624 cm^{-1}. Also a new small shoulder at 1697 cm^{-1} is present. In the literature the two peaks around 1624 and 1697 cm^{-1} have been assigned to antiparallel β-sheet structure (*38*−*40*).

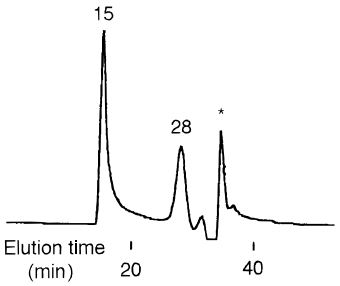

FIGURE 4: Typical HPSEC elution pattern of gene 9 protein, solubilized in cholate. The void volume at 15 min and the monomeric/small-oligomeric peak at 28 min are indicated. The peak indicated by an asterix is due to background of the buffer. The column was Superose 12 (1.5 × 30 cm), at a flow of 0.5 mL/min, with fluorescence detection at 340 nm, exciting at 280 nm, and the elution buffer was 50 mM cholate, 10 mM phosphate, pH 8.0.

The aggregational state of the protein in cholate was checked by HPSEC (data not shown). Whereas the protein was found to be monomeric/oligomeric directly after sample preparation, after 3 days over 90% of the protein eluted in the void volume, indicating the predominant presence of large aggregates. Thus, the conformational change, as observed by CD and FTIR, is concomitant with a state of increased aggregation of the gene 9 protein.

Reversibility of the Aggregated State of the Gene 9 Protein. Apparently, the protein can exist in two states: as monomers/oligomers containing α-helical and random structure and as large aggregates containing β-sheet structure. By addition of SDS, we tested whether the aggregated state in cholate was reversible. The CD and FTIR spectra of aggregated protein in cholate, to which SDS has been added, are shown in Figure 3A,B, respectively. By comparison of the spectra before and after addition of SDS, it is clear that SDS completely restores the α-helical conformation as measured by CD but that the FTIR absorbance at 1657 cm^{-1} is only partly restored. It should be noted that the final SDS/protein ratio in the FTIR experiment was much smaller as compared to CD (see legend Figure 3), explaining the partial disruption of the aggregates.

Since the aggregation of the gene 9 protein is reversible, it should be possible to influence the aggregation state, as determined by HPSEC, to some extent by variation of conditions. A typical HPSEC chromatogram is presented in Figure 4, showing the void volume peak at 15 min and a peak at 28 min, representing aggregated and monomeric/oligomeric protein, respectively. As is shown in Table 1, the aggregational state is influenced by several parameters. From lines 1−4 in Table 1 it is clear that the protein aggregates with time and that it is a slow process (hours to days). Lines 5−9 show that the aggregation rate depends on the temperature. Lowering the temperature to 4 °C slows down the aggregation process, whereas storage at −20 °C completely stops aggregation. Increasing the temperature to 75 °C results in a very fast aggregation (seconds). Lines 10

Table 1: Relative Percentages of Aggregated Gene 9 Protein after Incubation at Several Conditions, as Measured by HPSEC

sample[a]	final concn (mM) after addition of:	temp (°C)[c]	incubation time	relative area of void vol (%)[d]
1		RT	none	0
2		RT	5 h	44
3		RT	1 day	88
4		RT	2 weeks	100
5		4	1 day	20
6		4	2 weeks	94
7		−20	1 day	0
8		−20	2 weeks	0
9		75	20 s	57
10[b]	cholate, 250 mM	RT	none	79
11[b]	cholate, 250 mM	RT	5 h	38
12[b]	cholate, 250 mM	RT	>1 day	>38
13[b]	SDS, 250 mM	RT	<1 h	3

[a] Sample contains 0.56 mM gene 9 protein and 50 mM cholate in sodium phosphate pH 8.0. [b] Already aggregated protein, as preincubated at RT during 1 day (=sample 3). Incubation times mentioned are after addition in column 2 to preincubated sample. [c] RT = room temperature. [d] Values are ±5%.

and 11 show that an increase of the cholate concentration in an already aggregated sample can disrupt the aggregates slowly to some extent, but eventually aggregation will proceed (line 12). The strong detergent SDS can disrupt aggregates almost completely in a short time, as is clear from line 13.

Effect of Lipids on the Initial State of Cholate-Solubilized Protein. Cholate-solubilized protein in the monomeric/oligomeric α-helix containing state or in the aggregated β-sheet state, as characterized in Figure 3, was reconstituted into DOPG by cholate dialysis. The aggregational state of the protein after transfer to DOPG was determined using HPSEC. The initially monomeric/oligomeric protein retained this state upon reconstitution. Also when initially aggregated protein was reconstituted into DOPG, the aggregated state was retained (data not shown). In both cases the initial aggregational state of the protein appeared to be maintained as the elution patterns remained similar during several days. Apparently, the aggregational state of the protein is not affected by reconstitution or upon subsequent aging.

The corresponding CD spectra of gene 9 protein after transfer to DOPG are shown in Figure 5. When α-helix-containing monomeric/oligomeric protein in cholate was transferred to DOPG by the cholate dialysis protocol, a similar CD spectrum was observed in DOPG. This CD spectrum was comparable to the spectrum in DOPG as shown in Figure 2, in which the protein was directly transferred from TFE to DOPG, and identical to the CD spectrum of the α-helix-containing protein in cholate prior to reconstitution (Figure 3A). In contrast, when initially aggregated protein was reconstituted into DOPG, the CD spectrum still showed a single minimum at 219 nm but showed, as compared to the CD spectrum of the aggregated protein in cholate (Figure 3A), an additional red-shifted zero-crossing and a lower intensity. The CD spectrum of the aggregated protein in DOPG appeared to be indicative for the presence of β-sheet structures (*37*). The corresponding infrared absorbance by the amide I band of the same aggregated sample in DOPG is shown in the insert of Figure 5. Three peaks are visible: a major band at 1626 cm^{-1}, a peak at 1655 cm^{-1}, and a small peak at 1697 cm^{-1}. The 1626 and

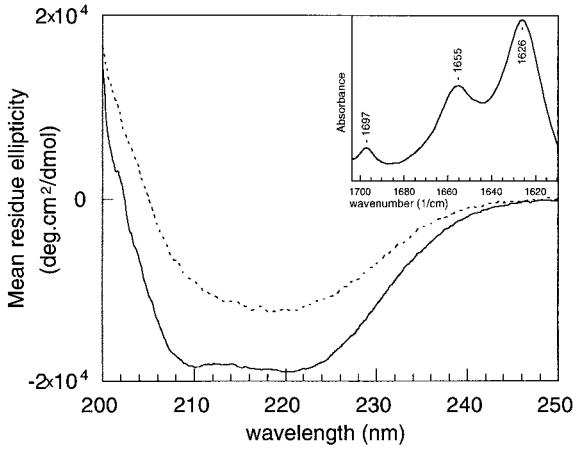

FIGURE 5: CD spectra of gene 9 protein containing vesicles, reconstituted via cholate dialysis. The dashed line represents reconstitution started from monomeric/small-oligomeric protein in cholate; the solid line represents reconstitution started from aggregated protein in cholate. The protein concentration was 30 μM and the concentration DOPG was 1.7 mM, resulting in an L/P of 57. Insert: FTIR absorbance of dehydrated film of gene 9 protein containing DOPG vesicles, reconstituted from aggregated protein in cholate, at an L/P ratio of 50.

1697 cm^{-1} peaks are in accordance with the presence of antiparallel β-sheet structure (*38*, *39*). The 1655 cm^{-1} band, however, corresponding to α-helix or random coil, indicates that also these secondary structures, as observed in the initial state of the aggregated protein in cholate (Figure 3B), were preserved.

DISCUSSION

From the literature the gene 9 protein is known to be located in the bacterial inner membrane prior to assembly (*26*). However, no detailed information is available about the nature of the interaction with the membrane and about the conformation of the protein in the membrane-bound state prior to assembly. The hydropathy plot (Figure 1) suggests the possibility for the protein to span the membrane in an α-helical way, which requires the presence of a hydrophobic segment of about 20 amino acids. Since the protein has not been characterized before, in this work the conformational and aggregational properties of the protein in various model systems are investigated, moving toward reconstitution and determination of the state of the protein in model phospholipid membranes.

Characterization of Gene 9 Protein in TFE, Detergents, and Lipids. After the protein was dissolved in TFE, it was transferred to cholate, SDS, or DOPG by cosolubilization in a TFE/water mixture and subsequent dilution, lyophilization, and rehydration as described by Killian et al. (*33*). In all systems, including TFE, we found the protein to be in an α-helix-containing monomeric or small-oligomeric state, with no large aggregates present.

For CD measurements of gene 9 protein in pure TFE two minima at 208 and 222 nm were observed, typical for an α-helical conformation (*37*, *41*). These two minima were also present in the other systems but less distinct and less intense. In pure TFE and TFE/water (1/1, v/v), the negative CD bands are more intense and the minima are better defined, as compared to the gene 9 protein in cholate, SDS, and DOPG. Two effects could explain the intensity increase in TFE. First,

the helix in TFE may be elongated, compared to the helix in the other systems. In cholate, SDS, and DOPG, possibly not the full peptide but only the hydrophobic stretch of about 20 residues (Figure 1) is immersed in a hydrophobic environment. The remaining charged part, likely to be exposed to the aqueous phase in micelles or vesicles, may contain other secondary structures, like random structure. This part may also become α-helical in TFE, since TFE induces nativelike helices and segments that have an α-helix propensity (*42*−*44*), thus increasing the length of the α-helical strand. According to the literature, the intensity of the CD band for the helix is dependent on the average number of residues per strand (*45*). Second, the dielectric constant of the medium could be different in TFE, affecting the transition probabilities (*46*), but this cannot account for the large intensity difference.

The protein in DOPG, measured by FTIR, resulted in a single amide I band at 1656 cm^{-1}, which may be another indication for α-helix structure (*38*, *47*). However, a band at 1656 cm^{-1} can also be interpreted as random structure (*38*, *48*). Combining this with the CD data, we propose that the conformation in DOPG is a mixture of α-helix and random structure. FTIR clearly indicates that no β-structure is present, so the decreased CD intensity is probably due to a conversion into random structure, which also agrees with the observed FTIR spectrum. From the observations we roughly estimated the percentage of α-helix to be 50−60% ± 10%, and the remaining part is supposed to be random coil. The protein in SDS shows a similar CD spectrum, also indicating the presence of this mixture of α-helical and random structure. The FTIR spectrum of the protein in cholate (Figure 3B) shows a low-intensity band around 1680 cm^{-1}. Often peaks in this area are assigned to turns (*38*). However, definitive assignment of turns and discrimination from unordered structures is difficult (*49*), since data on turns are scarce. In addition, several types of turns exist, giving rise to a broad distribution of frequencies in the 1695−1660 cm^{-1} region. Therefore, the gene 9 protein in cholate might contain turnlike structures or distorted helical parts, but the main part is a mixture of α-helical and random structure.

From the literature it is known that close proximity of aromatic amino acid residues can give rise to a significant contribution in the CD spectrum (*50*, *51*). This effect could make secondary structure estimations unreliable. In the case of the gene 9 protein in lipid systems, we do not expect the aromatic residues of the protein to be strongly immobilized or clustered. Moreover, we have investigated the near-UV region (250−300 nm) of the CD spectrum of the protein in lipid bilayers (data not shown), where we could find no detectable contribution of aromatic side chains. This observation also suggests that the contribution of the aromatic amino acid residues in the far-UV region (190−250 nm) is negligible.

The aggregational state of the protein, determined by a cutoff filter or HPSEC, in TFE or detergents was determined to be between monomers and small oligomers. We also used HPSEC analysis in cholate elution buffer to determine the aggregational state of the gene 9 protein in DOPG vesicles. When the sample is applied to the cholate-eluted column, the conditions are actually changed, because the sample is diluted and mixed with cholate. This may result in a change of the aggregation state. However, with the protein concentrations used, it was found that the aggregational state hardly changed, since both elution patterns with almost all protein in the void volume peak, as well as patterns with all protein in the monomeric/oligomeric peak, were observed. This suggests that during the time of elution no disruption of aggregates nor formation of aggregates occurs.

Stability of the Protein in TFE, Detergent Micelles, and Lipids. In TFE and TFE/water (1/1, v/v) the protein can be stored for prolonged times, without any change in conformational or aggregational state. Also in SDS and DOPG, the CD spectra and HPSEC elution profiles did not change during several days, indicating conformational and aggregational stability upon aging. In cholate, however, a conformational change was observed with both CD and FTIR. The CD spectrum after 3 days showed a single minimum at 218 nm (instead of the two minima at 208 and 222 nm directly after preparation), although the zero-crossing and intensity hardly changed (Figure 3A). This is not a typical β-sheet spectrum as deduced from the CD spectra of poly-L-lysine in the β-sheet form (*37*), which has a red-shifted zero-crossing and a decreased intensity compared to the α-helical form. However, the appearance of a main band at 1624 cm^{-1} and a small accompanying shoulder at 1697 cm^{-1} in the FTIR spectrum (Figure 3B) clearly indicate the presence of antiparallel β-sheet structure (*38*−*40*).

If the FTIR and CD data are combined, the altered CD spectrum of cholate-solubilized gene 9 protein is supposed to represent mainly antiparallel β-sheet structure, with some additional α-helix, random, and turn structure. More recent literature shows that the CD spectrum of β-sheet structure may result in a considerable variation of spectra, all with a single minimum between 210 and 225 nm (*41*). Spectra with a single minimum and zero-crossing around 200−203 nm have for example been observed for the all-β protein prealbumin, the α/β protein subtilisin, and the $\alpha + \beta$ protein ribonuclease (*52*). (See for definition of these protein classes ref *53*). The spectral changes, indicating a conformational change, are concomitant with the presence of large aggregates, since all protein is eluted in the void volume in the HPSEC analysis. Therefore, we conclude that in cholate the protein can exist as α-helix-containing monomers/oligomers or as mainly β-sheet-containing aggregates. The cholate solubilized protein undergoes a transition toward the aggregated state in time. Apparently, cholate is not suitable to solubilize the gene 9 protein.

Reversibility of the Aggregation. Addition of the strong detergent SDS to the aggregated β-sheet protein resulted in restoration of the α-helix containing conformation, as shown by CD and FTIR, and in a concomitant disruption of the large aggregates, as shown by the disappearance of the void volume peak and return of the monomeric/oligomeric peak with HPSEC. Thus, aggregation of the cholate-solubilized gene 9 protein is a reversible process in SDS, affected by several parameters mentioned in Table 1.

Effect of Lipids on the Initial Aggregational State of Cholate-Solubilized Protein. The cholate-solubilized protein can be reconstituted into vesicles of DOPG. The state of the protein, whether α-helix-containing monomeric/oligomeric or β-sheet aggregate, does not change upon incorporation into phospholipids. Monomeric/oligomeric protein, solubilized in cholate, was incorporated into DOPG using the cholate dialysis protocol. The CD spectrum in DOPG was

α-helical (*37*) and remained identical during several days. The aggregational state, as determined by HPSEC, was monomeric/oligomeric and stable for days as well. When the states of the protein before reconstitution (Figure 3) and after reconstitution (Figure 5) are compared, CD, as well as HPSEC, yields similar results. This indicates that transfer from cholate to DOPG does not change the conformational and aggregational state of the protein. After incorporation in DOPG, the state of the protein is stable, whereas it is not stable in cholate. The state of the protein was similar to directly transferred protein into DOPG from TFE, as shown with CD in Figure 2.

Also large aggregates, solubilized in cholate, were incorporated into DOPG using the cholate dialysis protocol. The CD spectrum in DOPG was a typical β-sheet spectrum (*37*), with a red-shifted zero-crossing, decreased intensity, and a single minimum at 219 nm. The presence of β-sheet in this sample was additionally confirmed by FTIR, showing bands at 1624 and 1697 cm^{-1}, characteristic for antiparallel β-sheet (*38*−*40*). Also some intensity at 1655 cm^{-1} was observed, showing that still some α-helix or random coil was present.

Small differences in the zero-crossing and intensity of the CD spectra before (Figure 3A) and after reconstitution (Figure 5) suggest, however, that the protein might be organized somewhat differently in the planar DOPG bilayers, as compared to the more random and spherical organization in cholate micelles. This might influence the homogeneity of the structures and protein−protein interactions. Also, in cholate some turn structure might be present, as concluded from the presence of a peak at 1680 cm^{-1} (Figure 3B), which is not present in DOPG. Despite these minor differences, the aggregational and conformational state of the protein remains very similar before and after reconstitution. The protein remains present as large aggregates, maintaining a predominantly β-sheet conformation, which cannot be disrupted by DOPG. From these observations it follows that when using the cholate-dialysis protocol as a method for reconstitution, care has to be taken to avoid aggregation during cholate exposure, since the aggregational state is preserved in DOPG afterward. Thus in both cases, whether the protein was reconstituted as monomeric/oligomeric protein or as large aggregates, a conserving effect of the lipids was observed.

Conformational and Aggregational Properties of the Gene 9 Protein in Membrane-Mimicking Systems. Summarizing the conclusions, we found that the gene 9 protein is predominantly α-helical in TFE and TFE/water (1/1, v/v) and can be incorporated in detergent micelles and phospholipid vesicles, the system most close to the in vivo situation, by the method of cosolubilization via TFE. An alternative method of reconstitution via cholate dialysis requires precaution, because cholate is not sufficiently strong to prevent the protein from aggregating. It is surprising that this aggregation can be reversed by changing conditions. This in contrast with the gene 8 protein of bacteriophage M13, which forms irreversible aggregates (*36*).

Apparently, the tendency to aggregate is a property of the protein. However, it is not conceivable that this feature is important in vivo, since the aggregation-related conformational change appeared not to be reversible with phospholipids. In addition, the observed aggregation occurred on a time scale of days, which is very slow for a process to be relevant in vivo. The β-structures present in detergent are antiparallel strands, as shown by FTIR. If antiparallel strands are present in vivo, then one expects the gene 9 proteins to be oriented in two directions before conversion. This is not very likely with two positive charges on the C-terminal end. Another indication is that, similar to the gene 8 protein, also gene 9 and gene 7 protein of infecting bacteriophages may enter the membrane and can be reused for synthesis of new particles (*2*). This means that they must be inserted in the membrane in the same way as newly synthesized proteins. It is questionable that this could be done with recycled proteins in a β-conformation. Thus, the observed aggregation probably is a nonnative artifact. This conclusion means that it is important to control the aggregation of the protein during reconstitution or purification.

The gene 9 protein adopts an α-helical conformation in TFE and in TFE/water (1/1, v/v). TFE, although said to disrupt tertiary interactions (*54*), can be a good solvent for single spanning membrane peptides. TFE is known to induce α-helical structure in proteins or peptides, which have a propensity to form α-helices in the native state. However, at very high percentages of TFE also other secondary structures become α-helical (*54*). From these results, TFE seems a suitable solvent from which to incorporate the protein into more membranelike systems such as micelles and vesicles.

Although the α-helical content is lowered upon solubilization by detergents or reconstitution into DOPG, the rough estimate of 50−60% α-helical conformation in these systems is an indication for the ability of the protein to form an α-helix in membrane-mimicking systems. SDS is often used as a model system to determine membrane protein structures, although it has been argued recently that it does not mimic well the typical nature of a phospholipid membrane structure (*17, 18*). For the gene 8 protein of M13 and the related fd bacteriophage the conformation has been studied in SDS by high-resolution NMR (*13*−*16*), and it appears very similar to the conformation in phospholipids as studied by solid-state NMR by Opella (*9*). Moreover, the transmembrane domain determined from cysteine scanning in phospholipid membranes (*17*) followed exactly the determined transmembrane helix in SDS.

Biological Implications. The observation of an α-helix-containing structure of the gene 9 protein in different hydrophobic environments (TFE, detergents, phospholipids) suggests that this state is the functional state in vivo and effective in the assembly process. This is consistent with the finding that the protein is located in the inner membrane prior to assembly (*26*). In general, transmembrane domains of inner membrane proteins are expected to be α-helical. Up to now the conformation of the gene 9 protein was unknown. Previously a model was described in which the products of gene 7 and 9 were assumed to be α-helices, interacting with each other by their hydrophobic domains to cover the end of the phage particle (*28*). Thus, an α-helical conformation present in the gene 9 protein may be suitable in different stages of the life cycle of the bacteriophage, where it must be able to exist in various environments. First it must insert rapidly from the cytoplasm into the membrane phase, second it must recognize and associate with the assembly site, and third it might be essential for initializing the phage assembly process (*27*). Finally, it is part of the

viral particle. The finding that the α-helix-containing protein can be reconstituted in lipids in an α-helical conformation may be biologically relevant and opens the way for more detailed studies of gene 9 protein in lipid bilayers in the future.

ACKNOWLEDGMENT

We thank Cees Versluijs and Anita Oosterlaken-Buijs for their kind help with the mass spectrometry.

REFERENCES

1. Berkowitz, S., and Day, L. (1976) *J. Mol. Biol. 102*, 531−547.
2. Webster, R. E., and Lopez, J. (1985) in *Virus structure and assembly* (Casjens, S., Ed.) pp 235−267, Jones and Bartelett Publishers Inc., Boston MA.
3. Rasched, I., and Oberer, E. (1986) *Microbiol. Rev. 50*, 401−427.
4. Model, P., and Russel, M. (1988) in *The Bacteriophages* (Calender, R., Ed.) pp 375−456, Plenum Press, New York.
5. Webster, R. E. (1996) in *Phage Display of peptides and proteins* (Kay, B., Winter, J., and McCafferty, J., Eds.) pp 1−20, Academic Press, San Diego, CA.
6. Marvin, D. (1998) *Curr. Opinion in Struct. Biol. 8*, 150−158.
7. Russel, M. (1991) *Mol. Microbiol. 5*, 1607−1613.
8. Russel, M. (1995) *Trends Microbiol. 3*, 223−228.
9. Opella, S. J., and McDonnell, P. A. (1993) in *NMR of Proteins* (Clore, G. M., and Gronenborn, A. M., Eds.) pp 160−189, The MacMillan Press Ltd, London, GB.
10. Glucksman, M. J., Bhattacharjee, S., and Makowski, L. (1992) *J. Mol. Biol. 226*, 455−470.
11. Marvin, D. A., Hale, R. D., Nave, C., and Citterich, M. H. (1994) *J. Mol. Biol. 235*, 260−286.
12. Spruijt, R. B., and Hemminga, M. A. (1991) *Biochemistry 30*, 11147−11154.
13. Henry, G. D., and Sykes, B. D. (1992) *Biochemistry 31*, 5284−5297.
14. McDonnell, P. A., and Opella, S. J. (1993) *J. Magn. Reson., Ser. B 102*, 120−125.
15. Van de Ven, F. J. M., Van Os, J. W. M., Aelen, J. M. A., Wymenga, S. S., Remerowski, M. L., Konings, R. N. H., and Hilbers, C. W. (1993) *Biochemistry 32*, 8322−8328.
16. Papavoine, C. H. M., Aelen, J. M. A., Konings, R. N. H., Hilbers, C. W., and Van de ven, F. J. M. (1995) *Eur. J. Biochem. 232*, 490−500.
17. Spruijt, R. B., Wolfs, C. J. A. M., Verver, J. W. G., and Hemminga, M. A. (1996) *Biochemistry 35*, 10383−10391.
18. Stopar, D., Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (1996) *Biochemistry 35*, 15467−15473.
19. Wolkers, W. F., Haris, P. I., Pistorius, A. M. A., Chapman, D., and Hemminga, M. A. (1995) *Biochemistry 34*, 7825−7833.
20. Van Wezenbeek, P., Hulsebos, T., and Schoenmakers, J. (1980) *Gene 11*, 129−148.
21. Grant, R. A., Lin, T.-C., Webster, R. E., and Konigsberg, W. (1981) *Prog. Clin. Biol. Res. 64*, 413−428.
22. Lopez, J., and Webster, R. E. (1983) *Virology 127*, 177−193.
23. Simons, G. F. M., Beintema, J., Duisterwinkel, F. J., Konings, R. N. H., and Schoenmakers, J. G. G. (1981) *Prog. Clin. Biol. Res. 64*, 401−411.
24. Sugimoto, K., Sugisaki, H., Okamoto, T., and Takanami, M. (1977) *J. Mol. Biol. 111*, 487−507.
25. Wickner, W. (1983) *Trends Biochem. Sci. 8*, 90−94.
26. Endemann, H., and Model, P. (1995) *J. Mol. Biol. 250*, 496−506.
27. Russel, M., and Model, P. (1989) *J. Virol. 63*, 3284−3295.
28. Makowski, L. (1992) *J. Mol. Biol. 228*, 885−892.
29. Rietman, B., Smulders, R., Eggen, I., van Vliet, A., van de Werken, G., and Tesser, G. (1994) *Int. J. Pept. Protein Res. 44*, 199−206.
30. Kaiser, B., Collescot, R., Bossinger, C., and Cook, P. (1970) *Anal. Biochem. 34*, 595−598.
31. Seidel, C., Klein, C., Empl, B., Bayer, H., Lin, M., and Batz, H. (1991) *Pept. Proc. 21st EPS*, 236−237.
32. Peterson, G. L. (1977) *Anal. Biochem. 83*, 346−356.
33. Killian, J. A., Trouard, T. P., Greathouse, D. V., Chupin, V., and Lindblom, G. (1994) *FEBS Lett. 348*, 161−165.
34. Schägger, H., and Von Jagow, G. (1987) *Anal. Biochem. 166*, 368−379.
35. Haris, P. I., and Chapman, D. (1995) *Biopolymers 37*, 251−263.
36. Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (1989) *Biochemistry 28*, 9158−9165.
37. Greenfield, N. J., and Fasman, G. D. (1969) *Biochemistry 8*, 4108−4116.
38. Goormaghtigh, E., Cabiaux, V., and Ruysschaert, J.-M. (1994) in *Physicochemical Methods in the Study of Biomembranes* (Hilderson, H. J., and Ralston, G. B., Eds.) pp 329−450, Plenum Press, New York.
39. Chirgadze, Y. N., Shestopalov, B. V., and Venyaminov, S. Y. (1973) *Biopolymers 12*, 1337−1351.
40. Chirgadze, Y. N., and Nevskaya, N. A. (1976) *Biopolymers 15*, 607−625.
41. Venyaminov, S. Y., and Yang, J. (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Fasman, G., Ed.) pp 69−105, Plenum Press, New York.
42. Segawa, S., Fukuno, T., Fujiwara, K., and Noda, Y. (1991) *Biopolymers 31*, 497−509.
43. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., and Wright, P. E. (1992) *J. Mol. Biol. 226*, 795−817.
44. Kemmink, J., and Creighton, T. E. (1995) *Biochemistry 34*, 12630−12635.
45. Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
46. Fasman, G. D. (1995) *Biopolymers 37*, 339−362.
47. Nevskaya, N. A., and Chirgadze, Y. N. (1976) *Biopolymers 15*, 637−648.
48. Miura, T., and Thomas, G. J., Jr. (1995) in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (Glasel, J. A., and Deutscher, M. P., Eds.) pp 261−315, Academic Press, San Diego, CA.
49. Torii, H., and Tasumi, M. (1992) *J. Chem. Phys. 96*, 3379−3387.
50. Arnold, G. E., Day, L. A., and Dunker, A. K. (1992) *Biochemistry 31*, 7948−7956.
51. Manning, M. C., and Woody, R. W. (1989) *Biochemistry 28*, 8609−8613.
52. Hennessey, J. P., Jr., and Johnson, W. C., Jr. (1981) *Biochemistry 20*, 1085−1094.
53. Levitt, M., and Chothia, C. (1976) *Nature 261*, 552−558.
54. Sivaraman, T., Kumar, T. K. S., and Yu, C. (1996) *Int. J. Biol. Macromol. 19*, 235−239.
55. Kyte, J., and Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105−132.

BI981149E