FT-IR Spectroscopy of the Major Coat Protein of M13 and Pfl in the Phage and Reconstituted into Phospholipid Systems†

Win F. Wolkers,† Parvez I. Haris, Arthur M. A. Pistorius, Dennis Chapman, and Marcus A. Hemminga*‡

Department of Molecular Physics, Agricultural University, Dreijenlaan 3, NL-6703 HA Wageningen, The Netherlands, Department of Protein and Molecular Biology, Royal Free Hospital School of Medicine, Rowland Hill Street, London NW3 2PF, U.K., and Department of Biochemistry, University of Nijmegen, P.O. Box 9101, NL-6500 HB, Nijmegen, The Netherlands

Received August 26, 1994; Revised Manuscript Received April 7, 1995

ABSTRACT: FT-IR spectroscopy has been applied to study the secondary structure of the major coat protein of Pfl and M13 as present in the phage and reconstituted in DOPG and mixed DOPC/DOPG (4/1) bilayers. Infrared absorbance spectra of the samples were examined in dehydrated films and in suspensions of D2O and H2O. The secondary structure of the coat protein is investigated by second-derivative analysis, Fourier self-deconvolution, and curve fitting of the infrared bands in the amide I region (1600−1700 cm⁻¹). It is found that, in dehydrated films of Pfl and M13 phage, the amide I region contains three bands located at about 1633, 1657, and 1683 cm⁻¹, that are assigned to hydrogen-bonded turn, α-helix/random coil, and non-hydrogen-bonded turn, respectively. From a comparison of the infrared spectra in dehydrated film with those in aqueous suspension, the percentages of secondary structure were found with an accuracy of about ±5%. For the coat protein of Pfl phage, the FT-IR quantification gives 69% α-helix conformation, 19% turn structure, and 12% random coil structure. For Pfl coat protein in the membrane-embedded state, the amount of α-helix is 57%, whereas 42% is in a turn structure and 1% in a random coil structure. The same assignment strategy was used for the analysis of the data obtained for M13 coat protein reconstitution into phospholipid systems. For M13 coat protein in the phage, this gives 75% α-helix conformation, 21% turn structure, and 4% random coil structure. For reconstituted M13 coat protein, it can be calculated that the amount of α-helix is 50% and the amount of random coil structure is 4%. The remaining structure (46%) is a turnlike structure. These results show that for both the Pfl and M13 coat protein in the phage environment contains a high amount of α-helical structure, which decreases in a lipidic environment in favor of turnlike structures. It is concluded that, when embedded in a membrane, the C- and N-terminal regions of the coat proteins contain a large amount of secondary structure, resembling a partly "broken" and discontinuous helix structure. This suggests that these regions probably undergo an increased molecular motion.

The filamentous bacteriophages M13 and Pfl are long (800−2000 nm) and narrow (6−7 nm) filaments constructed from a circular single-stranded DNA molecule, which encodes the viral genes, several thousand copies of a major protein consisting of about 45−50 amino acids, and a few copies of one or more minor proteins located at the filament ends. During infection of bacterial cells, the major coat protein is stripped off from the DNA and inserted into the cytoplasmic membrane of the host. Newly synthesized coat proteins are inserted into the cytoplasmic membrane as well prior to being used in the assembly process. Several studies with emphasis on the assembly process have been carried out recently (Webster & Lopez, 1985; Rasched & Oberer, 1986; Model & Russel, 1988; Russel, 1991); however, the lipid−protein interactions and the molecular structure of the assembly site remain to be elucidated.

The amino acid sequences of the major coat proteins of M13 and Pfl share a central hydrophobic region and short hydrophilic N- and C-terminal regions. In spite of a lack in sequence homology, X-ray and neutron diffraction studies have revealed that the coat proteins in these phages have a similar secondary and tertiary structure (Nambudripad et al., 1991; Glucksman et al., 1992). This structure basically consists of two colinear α-helical regions interrupted by a loop structure. A study of the secondary structure of the M13 coat protein in phospholipid systems has been carried out using circular dichroism in combination with Raman and FT-IR spectroscopy (Sanders et al., 1993). As a result of the high amount of α-helical structure observed, it was suggested that the coat protein when embedded in lipid bilayers is predominantly in an α-helical conformation. A more detailed structure of M13 coat protein in a lipidic environment follows from high-resolution NMR studies of the protein solubilized in sodium dodecyl sulfate (Henry & Sykes, 1992; Van de Ven et al., 1993). From this work it follows that micellar-bound M13 coat protein consists of two

† This research was supported by the Netherlands Foundation of Chemical Research with financial aid of the Netherlands Organization for Scientific Research (NWO) and by the commission of the European Communities, contract ST2J-0088. P.I.H. and D.C. wish to thank the IRC in Biomedical Materials for financial support.
‡ Author to whom correspondence should be addressed at the Department of Molecular Physics, Agricultural University, P.O. Box 8128, 6700 ET Wageningen, The Netherlands.
§ Agricultural University.
¶ Royal Free Hospital School of Medicine.
© University of Nijmegen.

0006-2960/95/0434-7825$09.00/0 © 1995 American Chemical Society

Abbreviations: DOPC, 1,2-dioleoyl-sn-glycero-3-phosphocholine; DOPE, 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine; DOPG, 1,2-dioleoyl-sn-glycero-3-phosphoglycerol; HSR, electron spin resonance; FT-IR, Fourier transform infrared; L/P, lipid to protein molar ratio; NMR, nuclear magnetic resonance; CD, circular dichroism; DMPG, 1,2-dimyristoyl-sn-glycero-3-phosphoglycerol; D2O, H2O.
α-helical domains linked by a short region of uncertain conformation. In contrast, Thiaudière et al. (1993) found from FT-IR spectroscopy an average helix content of 60% for M13 coat protein in phospholipid bilayers. This low value suggests a membrane-bound structure to be much more disordered and consequently more flexible.

The secondary structure of Pf1 coat protein in lipid systems has been studied using NMR spectroscopy (Bogusky et al., 1985; Schiksnis et al., 1987, 1988; Shon et al., 1991). From this work it is found that the secondary structure of the coat protein consists of a long hydrophobic helix that spans the bilayer and a short amphipathic N-terminal helix. FT-IR studies of Pf1 coat protein in sodium dodecyl sulfate micelles and in lipid membrane systems indicate that the protein has a predominantly α-helical structure (Azpiazu et al., 1993). About 50–60% of the amide protons of membrane-bound coat protein exchange rapidly, while the rest is in slow exchange with the solvent. This is consistent with a part of the protein being exposed to the solvent and the other being embedded in the lipid bilayer and supports the conclusions from the NMR experiments that the protein is a single helix spanning the membrane.

Among the spectroscopic techniques, FT-IR spectroscopy has proved to be a very useful method to study the conformation of membrane proteins (Haris & Chapman, 1992; Sanders et al., 1993). The application of FT-IR spectroscopy to proteins is based on the assessment of the amide bands that arise from vibrations of the peptide backbone. The so-called amide I band and the amide II band are widely used for protein structural studies. The amide I absorption band is principally due to an in plane C=O stretching vibration, whereas the amide II band is associated with in plane N—H bending (Susi, 1969). The amide I region of the infrared spectrum between 1600 and 1700 cm⁻¹ can be used for determining the secondary structures of proteins. The C=O stretching frequency is very sensitive to changes in the nature of the hydrogen bonds arising from the different types of secondary structure. This leads to a characteristic set of infrared absorbance bands for each type of secondary structure (Susi et al., 1967).

In the present study we have recorded FT-IR spectra of the major coat protein of Pf1 and M13, in the phage and reconstituted in phospholipid bilayers. New information about the effect of the lipids on the secondary structure protein of the coat protein in both environments is obtained by mathematical data treatments such as second-derivative analysis, Fourier self-deconvolution, and curve fitting to resolve the bands in the amide I region (1600–1700 cm⁻¹) that are associated with the C=O stretching vibrations. The results show that the coat protein in the phage environment contains a high amount of α-helical structure, which decreases in a lipidic environment in favor of turnlike structures.

MATERIALS AND METHODS

Phage Preparation and Coat Protein Purification. Bacteriophages Pf1 and M13 were grown and purified as described before (Spruijt et al., 1989). The cholate-isolated coat protein of M13 was prepared by a modification of the procedure of Makino et al. (1975). Typically, 4 mL of bacteriophage solution (concentration 12 mg/mL) was mixed with 8 mL of 100 mM cholate in 150 mM NaCl, 0.2 mM EDTA, 10 mM Tris-HCl, pH 8.0, and a few drops of chloroform. NaCl was added to prevent irreversible protein aggregation. The suspension was incubated at 37 °C with occasional mixing until a clear nonopalescent solution was obtained. DNA and protein of the disrupted phage were separated on a Sephacryl S-300 column, using 10 mM cholate in 150 mM NaCl, 10 mM Tris-HCl, and 0.2 mM EDTA, pH 8.0, as elution buffer. Fractions with an absorbance ratio A₂₆₀/A₅₀₀ greater than 1.5 were collected and stored at 4 °C. The purification of the major coat protein of Pf1 was performed as described above, however using a cholate concentration of 96 mM. Disruption of the phage was carried out by sonication for 3 min with a Branson Sonifier B15 (power setting: 5 (max 350 W at setting 10), duty cycle 45%).

Lipid—Protein Reconstitution. DOPC and DOPG were obtained from Avanti Polar Lipids (Birmingham, AL), dissolved in chloroform (20 mg/mL). From the desired amounts of lipid solution, the chloroform was removed with nitrogen gas. To remove remaining chloroform, all samples were lyophilized for 3 h. The lipids were solubilized in a 50 mM cholate solution (containing 150 mM NaCl, 10 mM Tris-HCl, and 0.2 mM EDTA adjusted to pH 8.0) by sonication with a Branson Sonifier B15 for 1 min (power setting: 3 (max 350 W at setting 10), duty cycle 40%). To this solution was added the desired amount of protein to obtain L/P ratios in the range of 10—100. To remove cholate from the samples, the solutions were dialyzed at room temperature against a 100-fold excess of dialysis buffer (150 mM NaCl, 10 mM Tris-HCl, 0.2 mM EDTA, pH 8.0) for a total of 48 h, changing the buffer every 12 h. After the dialysis procedure, the samples were lyophilized, solubilized in the same amount of water, and centrifuged for 2 h in a Beckman centrifuge at 30 000 rpm. The aggregation and conformational state were checked using high performance size exclusion chromatography and circular dichroism as described by Spruijt et al. (1989). The protein concentration was determined with a bicinchoninic acid assay (Smith et al., 1985). The lipid concentration was determined using a phosphorus assay (Bartlett, 1959).

Fourier Transform Infrared Spectroscopy. Samples of Pf1 and M13 coat protein in the phage and reconstituted in phospholipid bilayers were recorded in a temperature controlled Specac cell with either a 6 μm tin spacer for the studies in H₂O or a 50 μm Teflon spacer for the studies in D₂O. The samples in D₂O were lyophilized and reconstituted in phospholipid bilayers were incubated for 1 h. Dehydrated membrane films, suitable for FT-IR analysis, were prepared on AgCl windows (16 mm diameter, Fischer Scientific, USA) using the isopotential spin dry procedure (Clark et al., 1980). The membrane films typically contained 40 nmol of protein. FT-IR spectra of samples in H₂O and D₂O were obtained on a Perkin-Elmer 1750 FT-IR spectrometer at 25 °C as described elsewhere (Sanders et al., 1993). For the samples in H₂O and D₂O, 400 and 160 scans were averaged, respectively. FT-IR spectra of the dehydrated membrane films were recorded at 25 °C on a Cygnus 100 spectrometer (Mattson, Madison, WI), equipped with a liquid nitrogen cooled, narrow band mercury/cadmium/telluride detector (Mattson, Madison, WI) and interfaced to a microcomputer (Motorola Model 2334 with UNIX operating system). The optical bench was purged with dry nitrogen gas at a flow rate of 20 L/min. The
acquisition parameters were as follows: 2 cm$^{-1}$ resolution; 256 co-added interferograms, 1.27 cm/s moving mirror speed, 4000–750 cm$^{-1}$ wavenumber range, triangle apodization function. The time needed for acquisition and processing of a spectrum was 4.5 min. Signal to noise ratios between 2200 and 2000 cm$^{-1}$ were found to be better than 4 x 10$^3$.

Data Analysis. Spectral display and analysis were carried out using the EXPERT-IR analytical software (Mattson, Madison, WI). The parameters for the Fourier self-deconvolution procedure were as follows: $K$-factor 1.6 (this results in 10 cm$^{-1}$ linewidth in the deconvoluted spectrum, assuming 16 cm$^{-1}$ linewidth for the original spectrum), apodization function given by $\sin^2(\pi x/\lambda)$, and Lorentzian lineshape function (Kauppinen et al., 1981). The linewidth in the deconvoluted spectrum was carefully chosen, to avoid the introduction of erroneous bands (Surewicz et al., 1993). Difference spectra were generated using an interactive routine (IDIFF, Perkin-Elmer) to subtract the spectrum of the buffer from the corresponding spectrum containing the protein. The buffer spectra to be subtracted were collected under the same conditions as the reconstituted samples. Criteria for correctness of subtraction were (1) removal of the band near 2200 cm$^{-1}$ and (2) flat base line between 1800 and 2000 cm$^{-1}$ for samples in H$_2$O and elimination of the strong band at 1209 cm$^{-1}$ for samples in D$_2$O, avoiding negative side lobes. Curve fitting of the original absorbance spectra was performed between 1800 and 1490 cm$^{-1}$. Second derivative and deconvoluted spectra were used to determine the number and position of the bands as starting parameters for the curve fitting procedure, assuming Lorentzian lineshapes. Second derivative spectra were smoothed over 23 data points. Initially, peak positions and linewidths were manually adjusted, guided by the difference between the measured and the generated spectrum, whereas the peak heights were calculated by the computer. In the final optimization procedure, during which all parameters were free for adjustment, a computer fitting was carried out by taking into account a convolution of Lorentzian and Gaussian lineshapes. The quality of the fits was always judged by comparing the deconvoluted generated spectrum with the deconvoluted measured spectrum, using the same parameters.

The deconvoluted spectra were used to estimate the structure fractions in the viral coat proteins. First, each component was assigned to a structure type, and then the integrated areas of these components were expressed as a percentage of the total amide I area and were used to determine the secondary structure elements. The curve-fitted spectrum was deconvoluted, and the deconvoluted spectrum was compared with the deconvoluted spectrum of the original absorbance spectrum. This procedure is described in detail in the Results section.

RESULTS

Protein Studies. M13 coat protein when reconstituted in phospholipid systems is able to adopt different aggregation states depending on its lipidic environment and history of preparation (Spruijt et al., 1989; Spruijt & Hemminga, 1991). Therefore, the aggregation state of the protein was checked in the various reconstituted lipid–protein systems. High performance size exclusion chromatography demonstrated the absence of large aggregates in the reconstituted systems. This is indicative for an oligomeric state of the protein (Spruijt et al., 1989; Spruijt & Hemminga, 1991). After the spectroscopic experiments carried out in this paper, the coat protein was checked again using high performance size exclusion chromatography. No changes could be detected, showing that the oligomeric state of the protein in the model membrane systems was stable at all L/P ratios used. A similar result was found for the PFI coat protein.

PFI Coat Protein in the Phage. Figure 1A shows the infrared absorbance spectrum of a dehydrated film of PFI phage particles. Two main bands are seen in the spectral region between 1500 and 1800 cm$^{-1}$ that can be identified as the amide I band centered at 1657 cm$^{-1}$ and the amide II band located at 1543 cm$^{-1}$. Figure 1A also gives the component spectra that are used to fit the actual spectrum. From the curve fitting analysis, it is found that the amide I band is composed of three bands at 1633, 1657, and 1683 cm$^{-1}$ (Spruijt et al., 1989; Spruijt & Hemminga, 1991). The assignment for these bands is 15.11 amide I band centered at 1657 cm$^{-1}$, and the amide II band centered at 1633 cm$^{-1}$.

![Figure 1A](image-url)
The amide II band can be resolved into two bands at 1535 and 1543 cm\textsuperscript{-1} (see Table 1). For comparison, the best fit of the curve fitting procedure is shown in Figure 1B; the difference spectrum between the original spectrum and best fit is presented in Figure 1C, showing the high quality of the fit. Residual error intensities, calculated over the full range studied, varied from less than 0.006 to 0.009. The full assignment of the peaks in the 1500–1800 cm\textsuperscript{-1} region including tentative assignments to Tyr, Asp, and Lys of Pfl coat protein in various systems are presented in Table 2.

A comparison of the infrared absorbance spectra of Pfl phage particles in dehydrated film and in suspensions of D\textsubscript{2}O and H\textsubscript{2}O is shown in Figure 2. The corresponding curve fitting results are given in Table 2. When going from the phage in dehydrated film to the phage in H\textsubscript{2}O suspension, the linewidths decrease and the positions of the peaks shift to lower wavenumbers. In Table 2 it is observed, however, that the relative area of the three bands at 1633, 1657, and 1683 cm\textsuperscript{-1} is almost identical in H\textsubscript{2}O suspension and dehydrated film. Therefore, it is supposed that the spectral changes in Figure 2 are caused by changes in the hydration of the phage with the aqueous solvent, and not by a structural change of Pfl coat protein. By drying an aqueous solution of the phage, hydration water is removed, reducing the hydrogen bonds of the phage to the water molecules. This will lead to the observed spectral changes. The increase of the linewidths is probably related to the change from a homogeneous environment for the hydrogen bonds to the C\textequiv\textO groups in H\textsubscript{2}O suspension, to an inhomogeneous environment in a dehydrated system, in which molecular averaging processes are strongly reduced.

During the course of our experiments, we found it more difficult to apply curve fitting to the FT-IR spectra in H\textsubscript{2}O, as compared to the spectra in D\textsubscript{2}O suspension and dehydrated film. This arises because the spectra in H\textsubscript{2}O suspension need to be corrected for the high-intensity water band. Especially the end point of the water subtraction is difficult to determine, and the signal-to-noise ratio in these spectra is generally worse, leading to higher residual error intensities (about 0.017). In the following, therefore, we have used the FT-IR spectra in D\textsubscript{2}O suspension and dehydrated film for further analysis of secondary structure, assuming that for all systems investigated the secondary structure of the phage coat proteins is not affected by a change in solvent (H\textsubscript{2}O – D\textsubscript{2}O) or by dehydration.

Since the amide I band is most useful for the determination of secondary structure, the curve fitting results for this band of Pfl coat protein in various systems are presented in Table 3. For intact Pfl phage in a dehydrated film, the major amide I component at 1657 cm\textsuperscript{-1} can be assigned to \(\alpha\)-helical structure. It should be noted that this band might also contain a contribution of unordered (or less ordered) structures (random coil), since random coil and \(\alpha\)-helical absorptions strongly overlap in H\textsubscript{2}O (Susi, 1969; Azpiroz et al., 1993). The relatively narrow bandwidth, however, suggests a major contribution of \(\alpha\)-helical absorption.

The ratio of the weak bands in the amide I region at 1633 and 1683 cm\textsuperscript{-1} can be used to distinguish between a \(\beta\)-sheet structure and a turn structure (Byler & Susi, 1986). We observe a ratio of the band at 1633 cm\textsuperscript{-1} to the band at 1683 cm\textsuperscript{-1} of less than 1 in the dehydrated film spectra. In case of a \(\beta\)-sheet structure this ratio should have been about 10. Therefore, these bands are assigned to turn structure. Following the work of Mantsch et al. (1993) on \(\beta\)-turn-forming cyclic hexapeptides, the low wavenumber at 1633 cm\textsuperscript{-1} indicates hydrogen-bonded C\textequiv\textO groups, whereas the band at 1683 cm\textsuperscript{-1} can be assigned to non-hydrogen-bonded C\textequiv\textO groups present in a turn structure (Bandekar, 1992;
Mantzsch et al., 1993). On the basis of this assignment, the total amount of turn structure then amounts to 19%.

It is known that the band position of α-helical structures in the C=O amide I region is hardly affected by H2O/D2O exchange, in contrast to random coil and turn structures. This is due to the inaccessibility of the intramolecular H-bonds within the α-helix structure. FT-IR spectra of the Pf1 phage in D2O suspension show a shift of the main band from 1657 cm\(^{-1}\) in the dehydrated film to 1653 cm\(^{-1}\) in D2O suspension (Table 3). Along with this effect, the bandwidth decreases from 19 to 15 cm\(^{-1}\) and the relative percentage of the band decreases from 81% to 60%. A concomitant increase of intensity and bandwidth is found at 1634 cm\(^{-1}\). Therefore, the transfer of intensity from the 1657 cm\(^{-1}\) position in dehydrated film to the 1634 cm\(^{-1}\) position in D2O is presumed to arise from random coil structure, and the intensity remaining at 1653 cm\(^{-1}\) in D2O is assigned to α-helical structure.

In addition to this effect, the band at 1683 cm\(^{-1}\) in dehydrated film shifts to 1676 cm\(^{-1}\) in D2O, and the intensity is reduced by about a factor of 2. In principle, the intensity can transfer to the bands at 1653 and 1634 cm\(^{-1}\). It is unlikely that the intensity is added to the α-helix band at 1634 cm\(^{-1}\), because this band becomes sharper. The band at 1634 cm\(^{-1}\) broadens. Therefore, it is suggested that the intensity is transferred to the band at 1634 cm\(^{-1}\). This implies that in D2O, non-hydrogen-bonded C=O groups undergo a large shift from a 1683 to 1634 cm\(^{-1}\) in the FT-IR spectrum. A similar intensity change has been observed by Thiaudiere et al. (1993) for Pf3 coat protein incorporated in D2O-hydrated membrane systems.

**M13 Coat Protein in the Phage.** In Table 4, the curve fitting results of the amide I region of M13 coat protein in various systems are shown. For M13 phage particles in dehydrated film and D2O a similar band assignment as for Pf1 phage can be applied. M13 in D2O suspension has a slightly higher α-helix content (67%) than the comparable Pf1 system (60%).

**Membrane-Bound Pf1 Coat Protein.** In Figure 3A the infrared spectrum of the coat protein of Pf1 reconstituted in DOPG bilayers at L/P 20 in a dehydrated film is shown. It can be seen that the amide I band is sharper as compared to the protein in the phage environment under comparable conditions. This suggests a more homogeneous structural composition of the protein situated in a lipidic environment. Table 3 shows the curve fitting results of Pf1 coat protein reconstituted in the DOPG bilayers in dehydrated films as well as D2O suspensions. Curve fitting of the infrared spectra of dehydrated film samples shows that the amide I band is mainly composed of a narrow band at 1658 cm\(^{-1}\), which can be assigned to α-helical structure. The component at 1673 cm\(^{-1}\) reflects 30% of the amide I region in dehydrated film. This band is assigned to non-hydrogen-bonded C=O groups present in a turn structure. The small band at 1644 cm\(^{-1}\) in the dehydrated film spectra reflects a hydrogen-bonded turn structure (Bandekar, 1992; Mantzsch et al., 1993).

For the samples of Pf1 coat protein reconstituted in DOPG bilayers in D2O suspension, the band at 1657 cm\(^{-1}\) is assigned to an α-helical structure. As compared to the coat protein in the phage, this band is shifted to a higher wavenumber and the bandwidth is decreased from 15 to 10 cm\(^{-1}\). The percentage of α-helical structure obtained from the infrared spectrum slightly decreases from 58% in dehydrated films to 53% in D2O suspension. The band at 1634 cm\(^{-1}\) of Pf1 phage in D2O suspension, arising from random coil structure and hydrogen-bonded turn structure, shifts to 1640 cm\(^{-1}\) when going to the DOPG system. The band area increases from 33% to 43%. The band at 1676 cm\(^{-1}\) only shows a slight reduction in area, from 7% to 4%.

The results obtained for Pf1 coat protein reconstituted in mixed DOPC/DOPG (4/1) bilayers, mimicking the bacterial membrane, are similar to those obtained for pure DOPG bilayers, showing that partly substituting DOPC for DOPG does not influence the protein conformation (data not shown).

**Membrane-Bound M13 Coat Protein.** In Figure 3B the infrared spectrum is given of the coat protein of M13 reconstituted in DOPG bilayers in a dehydrated film. In Table 4 the curve fitting results and assignments of the amide I region are presented. The results for M13 and Pf1 coat protein are similar. Again, the band around 1656–1657 cm\(^{-1}\) in dehydrated film as well as D2O suspension is assigned to absorption of mainly α-helical structures. The
area of this band reduces strongly when going from the phage to the DOPG system. The band at 1643 cm$^{-1}$ of the reconstituted protein indicates a relatively high amount (21%) of hydrogen-bonded turn structure. The weak bands at 1629 and 1691 cm$^{-1}$ in the dehydrated system are indicative for a small amount (less than 5%) of the protein being in an aggregated $\beta$-polymeric state (Sanders et al., 1993).

Reconstituted M13 coat protein in D$_2$O suspension has a lower $\alpha$-helix content (42%) than the comparable Pf1 system (53%). Further differences between reconstituted M13 and Pf1 coat protein can be observed at the broad band (18 cm$^{-1}$) at 1640 cm$^{-1}$. This band indicates the presence of a significant amount (56%) of random coil structure and other solvent accessible structures.

DISCUSSION

In this paper a comparison of the secondary structure of the Pf1 and M13 coat protein in the phage and reconstituted in phospholipids is carried out using FT-IR spectroscopy. The aim of this investigation is to obtain a quantitative determination of the appearance of the various structural arrangements of the major coat proteins of Pf1 and M13 in the phage and reconstituted in phospholipid bilayers. In lipid model systems it has been found that M13 coat protein can adopt two different conformations with different aggregation behavior (Hemminga et al., 1992, 1993). M13 coat protein in a predominant $\alpha$-helix conformation can form reversible small aggregates ($\alpha$-oligomeric protein). The coat protein in a predominant $\beta$-sheet conformation appears to be aggregated in an irreversible way ($\beta$-polymeric protein). The presence of either one of these M13 coat protein forms depends on the protein purification procedure and reconstitution conditions (Spruijt et al., 1989; Spruijt & Hemminga, 1991). M13 coat protein in the $\alpha$-oligomeric form is assumed to be the biological active form in membrane-associated processes (Spruijt & Hemminga, 1991; Hemminga et al., 1993). The $\beta$-polymeric form is an artificial state that occurs in reconstituted lipid-protein systems. For the coat proteins reconstituted in phospholipid systems, care was taken to avoid the presence of protein in the artificial aggregated $\beta$-polymeric form. It is found that reconstituted Pf1 coat protein is more stable than reconstituted M13 coat protein, which gives rise to the formation of $\beta$-polymers, as deduced from the infrared bands around 1629 and 1691 cm$^{-1}$ (see Table 4). The presence of a small amount (generally <10%) of $\beta$-polymers was found to be unavoidable. However, since the corresponding FT-IR bands can be easily identified, the presence of $\beta$-polymers does not interfere with the conclusions about the nonaggregating coat protein.

The most useful characteristic vibrational mode for protein secondary structure determination is the C=O stretching vibration, located between 1600 and 1700 cm$^{-1}$ (the amide I band region). The C=O stretching frequency is very sensitive to changes in the nature of the hydrogen bonds arising from the different types of protein secondary structure. This leads to a characteristic set of infrared absorption bands for each type of secondary structure (Susi et al., 1967; Byler & Susi, 1986; Haris et al., 1986; Krimm & Bandekar, 1986). The application of FT-IR spectroscopy to proteins is based on the assessment of the amide I band frequencies that arise from vibrations of the peptide backbone.

Previously, a study of secondary structure of M13 coat protein in phospholipids was carried out using a factor analysis method based on pattern recognition, which compares measured infrared spectra with those of a reference set (Sanders et al., 1993). In this method, the amide I region is compared with a reference set of 20 proteins (Lee et al., 1990). This gave an unusual result of over 100% $\alpha$-helix for M13 coat protein reconstituted in lipid bilayers. It was suggested that this effect could be due to the fact that the spectral properties of M13 coat protein are outside the range of properties defined by the calibration set. Furthermore, it was argued that the calibration set was composed of large, water-soluble proteins, whereas M13 coat protein is a small membrane protein. For this reason we have chosen here to use the alternative method of curve fitting to analyze our spectra and obtain a quantification of structural elements. This method was also applied to viral coat proteins by Thiaudière et al. (1993). The advantage of this method, as compared to the factor analysis method, is that it is an independent method, since the spectra are not compared with reference sets. However, it should be mentioned that in this method it is assumed that the extinction coefficients for the different types of secondary structure are the same, so that

![Figure 3: FT-IR absorption spectra and deconvolved spectra of (A) Pf1 and (B) M13 coat protein reconstituted in DOPG bilayers at L/P 20 in a dehydrated film at 25 °C. The peak at 1740 cm$^{-1}$ arises from the carboxyl groups of the lipids. Left panel: absorption spectra; right panel: deconvoluted spectra.](image-url)
FT-IR Spectroscopy of Viral Coat Proteins

Biochemistry, Vol. 34, No. 24, 1995 7831

the relative band area in the amide I region reflects the amount of secondary structure present. Recently, a review has been published that discusses the advantages and disadvantages associated with the different techniques available for quantitative analysis of protein secondary structure from their infrared spectra (Surewicz et al., 1993).

In the process of analyzing our results, we have chosen to use the infrared spectra in dehydrated films and D$_2$O suspension for the curve fitting analysis. As compared to the spectra of H$_2$O suspensions, there is a slight disadvantage in using the spectra in dehydrated films, because of the increased bandwidths. However, the curve fitting results of the spectra in dehydrated films are more accurate, because the signal-to-noise ratio in the spectra is better and no subtraction of the water band is needed. By using the curve fitting procedure, no internal inconsistencies were observed.

In the curve fitting analysis a convolution of Lorentzian and Gaussian lineshapes was taken into account for optimal fitting of the infrared spectra. For the infrared spectra of Pfl coat protein in the phage, for example, it was observed that the band at 1657 cm$^{-1}$ had a Lorentzian shape, whereas the band at 1683 cm$^{-1}$ had a strong Gaussian character (about 75%). This difference in bands shape character is probably related to the homogeneity of the secondary structure that gives rise to the corresponding band. Because of the lack of literature about the lineshape character of infrared bands, these results are not further used in our analysis.

**Pfl Coat Protein in the Phage.** Analysis of the amide I regions of the infrared spectra of Pfl coat protein in the phage (Table 3) shows that it consists of three main bands in dehydrated films in H$_2$O solutions as well as in D$_2$O suspensions. However, the relative contribution of the bands is different in dehydrated films and in D$_2$O suspensions. The assignment of the peak at 1657 cm$^{-1}$ together with a peak at 1543 cm$^{-1}$ in the amide II region to α-helical structure is in agreement with other work (Byler & Susi, 1986; Haris et al., 1986; Krimm & Bandekar, 1986). The assignment of the bands at 1683 and 1633 cm$^{-1}$ to turn structure is confirmed by infrared spectra obtained of hemoglobin, where similar bands are observed at 1675 and 1638 cm$^{-1}$, assigned to turn structure (Byler & Susi, 1986), by FT-IR experiments carried out on dehydrated films of cytochrome c (A. M. A. Pistorius, unpublished results), and rhodopsin in bovine rod outer segment membranes (Haris et al., 1989).

In FT-IR spectra of several proteins in H$_2$O suspension or in dehydrated film, it is found that there is an overlap of the α-helical absorption with that of random coil (Byler & Susi, 1986; Krimm & Bandekar, 1986; Surewicz & Mantsch, 1988). However, these two types of structures can be distinguished by carrying out experiments in D$_2$O. Hydrogen–deuterium exchange of the peptide NH groups in random coil structure results in a shift of its amide I band to around 1640 cm$^{-1}$ (Byler & Susi, 1986; Haris et al., 1986; Surewicz & Mantsch, 1988). This can be observed in our infrared spectra (see Table 3), where a broad band at 1634 cm$^{-1}$ appears in D$_2$O suspension. This band is on the lower edge of frequencies possible for random coil structure. The band at 1634 cm$^{-1}$ also contains a contribution from turn structure. By comparing dehydrated films obtained from H$_2$O suspensions, with D$_2$O suspensions, the α-helix band shifts from 1657 to 1653 cm$^{-1}$. This shift arises from hydrogen (deuterium) bonding of α-helices with the solvent (Thiaudière et al., 1993). In addition to this effect, the linewidth of the α-helix band decreases from 19 to 15 cm$^{-1}$, which is in agreement with the conclusion that the band representing the random coil structure moves and separates from this band. The percentage of α-helical structure found in the dehydrated film (81%) sets an upper limit, and in D$_2$O suspension (60%) a lower limit to the amount of α-helical structure present in Pfl coat protein in the phage. The band arising from non-hydrogen-bonded C=O groups present in a turn structure shifts from 1683 cm$^{-1}$ in dehydrated films to 1676 cm$^{-1}$ in D$_2$O suspensions, and decreases in intensity. This change probably arises from the fact that non-hydrogen-bonded C=O groups are likely to be sensitive to hydrogen–deuterium exchange.

Neutron diffraction experiments (Nambudripad et al., 1991) have revealed that the 46 amino acids of Pfl coat protein in the phage consist of two α-helical regions of 23 and 11 amino acid residues, respectively, separated by a loop structure of 7 amino acids. The remaining amino acid residues are in a random coil configuration. Summing up both α-helical regions, this means that the protein contains 74% α-helical structure. The lower α-helical percentage found in D$_2$O suspension (60%) can be explained by assuming that the terminal parts of an α-helical region are accessible for H–D exchange. This effect has been observed in H–D exchange experiments in two-dimensional high-resolution NMR studies on helical regions in several proteins (Wüthrich, 1986). It may be expected that these slightly disordered α-helical regions will shift to lower wavenumbers, similar to the random coil structure upon changing the solvent from H$_2$O to D$_2$O. It is therefore likely that in D$_2$O suspension the peak at 1634 cm$^{-1}$ is composed of random coil, disordered α-helix, and turn structure.

It is reasonable to assume that one amino acid residue at each end of both α-helical regions is accessible for H–D exchange. It can then be calculated that 9% of the amino acids are affected by exchange. The amount of α-helix conformation is then estimated to be 69%. This is in good agreement with the α-helical percentage deduced from neutron diffraction experiments (74%). The decrease of intensity of the band around 1655 cm$^{-1}$ when going from dehydrated film samples to D$_2$O suspensions (21%) can then calculated to be the sum of the amount of disordered α-helical regions (9%) and random coil structure (12%). This leaves 19% of the protein in a turn structure. This is in good accordance with the result obtained from neutron diffraction, giving 11% and 15% for the random coil and loop conformations, respectively. This result indicates that the loop structure from amino acid residues 13–19 most likely contributes to the infrared absorption around 1635 cm$^{-1}$ assigned to turn structure.

Regarding the assignments of the band around 1640 and 1680 cm$^{-1}$, there is some uncertainty. It has been recently reported for myoglobin, on the basis of theoretical calculations, that bands below 1640 cm$^{-1}$ arise from helical vibrations (Torii & Tasumi, 1992). Also in this region, turns and β-sheets can overlap. However, the good comparison of the secondary structure deduced from FT-IR with the neutron diffraction results confirms our analysis strategy and the assignments presented in Tables 3 and 4.

**M13 Coat Protein in the Phage.** Clearly, the assumption that one amino acid residue on each end of an α-helix is accessible for H–D exchange enables us to fit the experimental FT-IR results of Pfl coat protein to the known neutron
diffraction structural results. However, it should be noted that no perfect agreement can be obtained, giving rise to errors of about 4–5% for the secondary structure determinations. Nevertheless, the same strategy will be used for the analysis of the data obtained for M13 coat protein in the phage and for Pfl1 and M13 coat proteins reconstituted in phospholipid systems. Thus, the analysis of the data obtained for M13 coat protein in the phage (Table 4) follows the same arguments as those developed for the coat protein of Pfl1. The percentage of \( \alpha \)-helical structure found in the dehydrated film (79%) sets an upper limit, and in \( \text{D}_2\text{O} \) suspension (67%) a lower limit to the amount of \( \alpha \)-helical structure present in M13 coat protein in the phage. If it is assumed that the coat protein contains two helices and that one amino acid residue at each end of both \( \alpha \)-helical regions is accessible for H–D exchange, it can be calculated that 8% is due to this effect. The amount of \( \alpha \)-helix conformation is then estimated to be 75%. The decrease of intensity of the band around 1650–1656 cm\(^{-1} \) when going from dry film samples to \( \text{D}_2\text{O} \) suspensions (12%) can then be calculated to be the sum of the amount of disordered \( \alpha \)-helical regions (8%) and random coil structure (4%). This leaves 21% of the protein in a turn structure.

M13 coat protein when associated with DNA in the virion is almost completely \( \alpha \)-helical, as judged by X-ray diffraction techniques (Banner et al., 1981; Glucksman et al., 1992), circular dichroism (Fodor et al., 1981; Clack & Gray, 1989; Arnold et al., 1992), and laser Raman spectroscopy (Thomas et al., 1983; Williams et al., 1984). The results found here for the coat protein in M13 phage correspond well with these studies, indicating that for the M13 coat protein in the phage, similar to Pfl1, two \( \alpha \)-helical regions are observed separated by a loop structure. However, the high percentage of turn structure (21%) indicates that the loop region between the helices involves more amino acid residues as compared to Pfl1 coat protein.

Membrane-Bound Pfl1 Coat Protein. As compared to Pfl1 coat protein in the phage, the linewidth of the \( \alpha \)-helical band at 1657 cm\(^{-1} \) is significantly decreased in the membrane-embedded state (Table 3), suggesting that the protein–protein contacts are much reduced and that the phospholipid provides a homogeneous environment for the coat protein. This is reflected in the infrared spectra of the coat protein of Pfl1 reconstituted in DOPG bilayers by a sharp \( \alpha \)-helical band around 1658 cm\(^{-1} \) (Figure 3A). The band position and linewidth are not sensitive for replacement of \( \text{H}_2\text{O} \) by \( \text{D}_2\text{O} \), indicating that the \( \alpha \)-helix is shielded from the aqueous solution. This suggests that the band arises from a transmembrane \( \alpha \)-helix embedded in a lipidic environment, in agreement with other observations (Pistorius & De Grip, 1994). This insensitivity also makes it unlikely that the band at 1673 cm\(^{-1} \) shifts and overlaps with the \( \alpha \)-helix band at 1657 cm\(^{-1} \), when substituting \( \text{H}_2\text{O} \) by \( \text{D}_2\text{O} \). This supports our previous interpretation that the band at 1673 cm\(^{-1} \) in \( \text{H}_2\text{O} \) moves to 1640 cm\(^{-1} \) in \( \text{D}_2\text{O} \). In addition, the linewidth of the band at 1640 cm\(^{-1} \) increases. No effect is found by varying the L/P ratio from 10 to 100. This indicates that an increasing amount of protein does not change the secondary structure.

The band in the dehydrated film spectra that is assigned to hydrogen-bonded C=O groups present in a turn structure shifts from 1633 cm\(^{-1} \) in the phage to 1644 cm\(^{-1} \) in the reconstituted lipid–protein sample. Also the band arising from non-hydrogen-bonded turn structure shifts from 1683 to 1673 cm\(^{-1} \). These shifts cannot be explained specifically, but are assumed to arise from a general effect of changing the environment from phage (mainly protein/water) to the membrane (mainly lipid/water) of the C=O groups involved.

To estimate the amounts of secondary structure of M13 coat protein, reconstituted in DOPG, the same calculation is followed as carried out for the coat protein in Pfl1 phage. From our results and other work (Azpiazu et al., 1993; Thiaudière et al., 1993), it follows that Pfl1 coat protein in the lipidic state contains only one transmembrane \( \alpha \)-helix. By assuming that one amino acid residue at each end of the helix is accessible for H–D exchange, it can be calculated that 4% is due to this effect. This means that, after taking into account this correction for the results in \( \text{D}_2\text{O} \) suspension, the amount of \( \alpha \)-helix is 57%. This indicates that the amount of random coil structure is 1%. The remaining structure (42%) is a turnlike structure. The high amount of turn structure and low amount of random coil structure differ from that found in other work (Sanders et al., 1993; Van de Ven et al., 1993).

It should be noted that the calculated amounts of secondary structure are estimates, based on a number of assumptions. The accuracy will therefore be about ±5%. Even so, our results suggest that 57% of the structure is \( \alpha \)-helix, which corresponds well with a transmembrane \( \alpha \)-helix of 21 amino acid residues. The remaining part of the protein is mainly in a turnlike structure. If a turn structure is visualized as a partly “broken” and discontinuous helix structure, the model that arises from our FT-IR results is that of a single-stranded protein, embedded in the membrane with the hydrophobic region, whereas the C- and N-terminal regions stick out of the membrane as a partly “broken” and discontinuous helix structure.

Membrane-Bound M13 Coat Protein. It is possible to estimate the amounts of secondary structure of M13 coat protein, reconstituted in DOPG, in a similar manner as for the Pfl1 coat protein. After correcting for the presence of 8% of \( \beta \)-polymers in Table 4, and taking into account a value of 4% for exchangeable protons in a disordered \( \alpha \)-helical region, it can be calculated that the amount of \( \alpha \)-helix is 50% and the amount of random coil structure is 4%. The remaining structure (46%) is a turnlike structure.

When reconstituted in phospholipids, Pfl1 and M13 coat proteins adopt an \( \alpha \)-helical structure in the protein region that is embedded in the bilayer, whereas the remaining part has a high content of turn structure. On the basis of FT-IR results, values between 50% (Thiaudière et al., 1993) and almost 100% (Sanders et al., 1993) of \( \alpha \)-helical structure were previously assigned to membrane-bound M13 coat protein. A large part of this discrepancy is related to the lack of an adequate technique for quantitative analysis of membrane protein structures using FT-IR. The approach presented here provides a new approach that shows that the C- and N-terminal parts of these coat proteins contain a large amount of secondary structure that resembles a discontinuous \( \alpha \)-helix, indicating that these parts probably arising from an increased amount of molecular motion. This concept is in agreement with conclusions arising from NMR (Shon et al., 1991; Henry & Sykes, 1992; Van de Ven et al., 1993) and FT-IR studies (Thiaudière et al., 1993).
REFERENCES


B1942010G