Membrane-bound Conformation of M13 Major Coat Protein

A STRUCTURE VALIDATION THROUGH FRET- DERIVED CONSTRAINTS*

Received for publication, May 31, 2005, and in revised form, July 28, 2005. Published, JBC Papers in Press, September 8, 2005, DOI 10.1074/jbc.M505875200

Werner L. Vos, Rob B. M. Koehorst, Ruud B. Spruijt, and Marcus A. Hemminga

From the Laboratory of Biophysics, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

M13 major coat protein, a 50-amino-acid-long protein, was incorporated into DOPC/DOPG (80/20 molar ratio) unilamellar vesicles. Over 60% of all amino acid residues was replaced with cysteine residues, and the single cysteine mutants were labeled with the fluorescent label I-AEDANS. The coat protein has a single tryptophan residue that is used as a donor in fluorescence (or Förster) resonance energy transfer (FRET) experiments, using AEDANS-labeled cysteines as acceptors. Based on FRET-derived constraints, a straight α-helix is proposed as the membrane-bound conformation of the coat protein. Different models were tested to represent the molecular conformations of the donor and acceptor moieties. The best model was used to make a quantitative comparison of the FRET data to the structures of M13 coat protein and related coat proteins in the Protein Data Bank. This shows that the membrane-bound conformation of the coat protein is similar to the structure of the coat protein in the bacteriophage that was obtained from x-ray diffraction. Coat protein embedded in stacked, oriented bilayers and in micelles turns out to be strongly affected by the environmental stress of these membrane-mimicking environments. Our findings emphasize the need to study membrane proteins in a suitable environment, such as in fully hydrated unilamellar vesicles. Although larger proteins than M13 major coat protein may be able to handle environmental stress in a different way, any membrane protein with water exposed parts in the C or N termini and hydrophilic loop regions should be treated with care.

One of the most challenging problems in structural biology of the 21st century is the unraveling of the structure and function of membrane proteins. For the water-soluble proteins, x-ray diffraction and high field solution NMR spectroscopy are the most suitable tools for structure determination, but for membrane proteins that need to be embedded in an amphiphatic environment, there is not yet a well defined strategy for obtaining a protein structure (1). Among the spectroscopic methods, site-directed spectroscopic approaches are becoming increasingly important as alternative tools for structure determination of membrane proteins. These techniques are based on site-directed mutagenesis in combination with specific labeling and provide detailed information about the local environment of the labeled sites and membrane embedding (2–5).

Herein is described a further enhancement of the site-directed approach using site-directed labeling to determine distances within a membrane protein from fluorescence (or Förster) resonance energy transfer (FRET).2

In the case of the M13 major coat protein, >60% of the amino acid residues were replaced by a cysteine and used in the FRET analysis. Different structural models for the membrane-bound M13 coat protein and related proteins have been proposed in the literature. Some structures are “L-shaped” (10) or “U-shaped” (11), whereas it was also proposed that the conformation of membrane-bound coat protein is in fact “I-shaped,” similar to the conformation in the phage particle (12). On the basis of our FRET data, we showed that the conformation of coat protein in DOPC/DOPG vesicles is best described as I-shaped, in good agreement with the phage assembly model that was proposed earlier (13).

Our FRET results are quantitatively compared with the existing literature on the structure determination of the M13 coat protein in detergent micelles and stacked, oriented bilayers by high resolution (11) and solid state NMR spectroscopy (14), respectively. In both micelles and in stacked, oriented bilayers, the structure of the coat protein turns out to be strongly affected by the stress of the membrane-mimicking environment.

The advantage of the fluorescence approach applied in this study is that the experiments are carried out in unilamellar vesicles providing information about the structure of the coat protein in a natural non-stressed state. As such, this paper serves as a caution to experimentalists who study membrane proteins in systems other than vesicles, as the conformation can depend strongly on the environment.

EXPERIMENTAL PROCEDURES

Sample Preparation—A number of 31 single cysteine mutants of the coat protein was prepared following the protocol described earlier (15): A7C, A9C, N12C, S13C, Q15C, A16C, S17C, A18C, T19C, Y21C, I22C, G23C, Y24C, A25C, A27C, M28C, V29C, V30C, V31C, I32C, G34C, A35C, T36C, G38C, I39C, L41C, F42C, K43C, K44C, T46C, and A49C.

AEDANS, N-(acetylaminoethyl)-1-naphthylamine-5-sulfonic acid; DOPC, 1,2-dioleoyl-sn-glycerol-3-phosphocholine; DOPG, 1,2-dioleoyl-sn-glycerol-3-(phospho-rac-(1-glycerol)).

* This work was funded from the European Commission under Contract QLG-CT-2000-01801. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

To whom correspondence should be addressed: Laboratory of Biophysics, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands. Tel.: 31-317-482044; Fax: 31-317-482725; E-mail: marcus.hemminga@wur.nl.

2 The abbreviations used are: FRET, fluorescence (or Förster) resonance energy transfer;
For the purpose of calculating the spectral overlap integral, the mutant Y21A/Y24A/W26A/G23C, with the aromatic residues replaced for alanines and a glycine replaced by a cysteine at position 23, was prepared as well. The mutants were purified and labeled with 1,5-1-AEDANS (purchased from Molecular Probes) as described previously (15). Labeled coat protein was reconstituted in DOPC/DOPG (80/20 molar ratio, purchased from Avanti Polar Lipids) using the cholate dialysis method (16). To avoid intermolecular energy transfer, the lipid to protein molar ratio was kept high for all samples (~1500). For mutants A9C and G38C, it was necessary to use a slightly lower lipid to protein molar ratio (~400) because of a low fluorescence signal, due to low AEDANS labeling efficiency of these mutants. The final concentration-labeled mutant protein typically was ~1 μM in all samples.

**Fluorescence Measurements**—Absorption spectra were recorded on a Varian Cary 5E UV-Vis-NIR spectrophotometer. Fluorescence spectra were recorded on a Fluorolog 3.22 spectrofluorimeter manufactured by Jobin Yvon-Spex. Emission was detected at the wavelength of maximum fluorescence. The band pass for excitation was 2.0 nm, the band pass for emission was 5.0 nm. Spectra were corrected for wavelength-dependent output of the lamp. The size of the vesicles was 41 nm, as determined by dynamic light scattering. All spectra were corrected for contributions of light scattering by subtracting the excitation spectrum of vesicles containing wild type protein in approximately the same concentration as the mutant proteins.

** Förster Radius**—The Förster radius was calculated (6) from the following.

\[ R_0 = 9.78 \times 10^3 \left( \kappa^2 n^4 Q_0(\lambda) \right)^{1/6} \quad \text{(Eq. 1)} \]

In this equation, \( \kappa^2 \) is the orientation factor, \( n \) is the refractive index of the medium, \( Q_0 \) is the quantum yield of the donor, and \( J(\lambda) \) is the overlap integral. The quantum yield of wild type protein was determined with a comparative method, using quinine sulfate in 0.05 M H₂SO₄ as a reference (17). The overlap integral was determined from the fluorescence emission spectrum of the wild type protein and from the absorption spectrum of the AEDANS-labeled Y21A/Y24A/W26A/G23C mutant (6), using the computer program IGOR Pro 3.13 (Wave-metrics, Lake Oswego, OR). Judging from the excitation spectra (see Fig. 1), the absorption spectrum did not depend on the position of the AEDANS label.

To calculate the Förster radius, a value \( \kappa^2 = 2/3 \) was used. This approximation is assumed to be generally valid for proteins (6) and in particular for the Trp-AEDANS donor-acceptor pair (18). In our case, this assumption is further supported by the fact that the Trp donor (19, 20) and AEDANS acceptor (21) have a high degree of motion. In addition, all our measurements were carried out at room temperature in a mobile lipid environment well above the gel to liquid crystalline transition temperature of the lipids used (\( T_{\text{gel}} \text{DOPC} = -20 \) °C and \( T_{\text{gel}} \text{DOPG} = -18 \) °C). This allows us to use \( \kappa^2 = 2/3 \) for the dynamic averaging of the relative donor-acceptor positions in the M13 coat protein system. The refractive index of the bilayer (1.5) was taken from the literature (22). The quantum yield (\( Q_0 \)) and overlap integral (\( J(\lambda) \)) were calculated to be 0.3 and 6.0 × 10⁻¹⁵ M⁻¹ cm³, respectively. The resulting Förster radius of 24 ± 1 Å is in good agreement with the value found previously (23, 24).

**Energy Transfer Efficiency**—In general, energy transfer efficiency depends on the distance \( r \) between the donor and acceptor (6),

\[ E_{\text{calc}} = R_0^6 (R_0^6 + r^6) \quad \text{(Eq. 2)} \]

where \( R_0 \) is the Förster radius as defined by Equation 1.

**FIGURE 1.** Excitation spectra of AEDANS-labeled V31C (red) and T46C (blue) mutants. The peak at ~280 nm is indicative of tryptophan to AEDANS energy transfer, and the peaks at ~260 nm and ~340 nm are arising from direct excitation of AEDANS. The emission is detected at the AEDANS emission maximum at 490 nm. AU, arbitrary units.

Two typical excitation spectra of an AEDANS-labeled mutant close to tryptophan (Fig. 1, red line) and of an AEDANS-labeled mutant far away from tryptophan (Fig. 1, blue line) are shown. In the absence of energy transfer, the shape of the excitation spectrum should resemble the shape of the absorption spectrum. At an excitation wavelength of ~290 nm, the shape of the excitation spectrum deviates from the AEDANS absorption spectrum. The shape of this particular peak resembles that of the tryptophan absorption peak, which has a maximum at ~280 nm and a small characteristic shoulder at ~290 nm. Tryptophan itself does not emit at 490 nm, so the presence of a tryptophan peak in the AEDANS excitation spectrum indicates energy transfer from tryptophan to AEDANS.

Assuming that AEDANS is excited exclusively through direct excitation or tryptophan to AEDANS energy transfer, and that there is no intermolecular energy transfer, the fluorescence intensity depends on the wavelength of excitation through the following.

\[ F(\lambda) = (\epsilon_{\text{AEDANS}}^\lambda + \epsilon_{\text{tryp}}^\lambda) \eta \quad \text{(Eq. 3)} \]

Here, \( F \) is the fluorescence intensity, \( \epsilon_{\text{AEDANS}}^\lambda \) and \( \epsilon_{\text{tryp}}^\lambda \) are the molar extinction coefficients of AEDANS and tryptophan at wavelength \( \lambda \), \( \eta \) is the energy transfer efficiency, and \( \gamma \) is a constant that depends on the apparatus, energy acceptor, and concentration (25).

At 290 nm, AEDANS excitation can occur through both energy transfer and direct excitation, so Equation 3 becomes the following.

\[ F(290) = (\epsilon_{\text{AEDANS}}^{290} + \epsilon_{\text{tryp}}^{290}) \eta \quad \text{(Eq. 4)} \]

At 340 nm, tryptophan absorption is negligible, so Equation 4 reduces to the following.

\[ F(340) = (\epsilon_{\text{AEDANS}}^{340}) \gamma \quad \text{(Eq. 5)} \]

Rearrangement of Equations 4 and 5 directly gives the experimental energy transfer efficiency as follows.

\[ E_{\text{exp}} = \frac{F(290)}{F(340)} \frac{\epsilon_{\text{tryp}}^{290}}{\epsilon_{\text{AEDANS}}^{290}} \frac{\epsilon_{\text{AEDANS}}^{340}}{\epsilon_{\text{tryp}}^{340}} \quad \text{(Eq. 6)} \]

The values \( \epsilon_{\text{AEDANS}}^{290} \) and \( \epsilon_{\text{AEDANS}}^{340} \) were taken from the absorption spectrum of 1,5-AEDANS in phosphate buffer at pH 7.0 (26), being 1200 and 6000 M⁻¹ cm⁻¹, respectively. The value \( \epsilon_{\text{tryp}}^{290} \) was taken from the spectrum of N-acetyl-DL-tryptophan-amide in phosphate buffer at pH 6.5 (27), being 4800 M⁻¹ cm⁻¹.

The tyrosine absorption is small but not negligible at 290 nm (27). Through tyrosine to tryptophan energy transfer, additional excitation of tryptophan occurs. M13 major coat protein has two tyrosine residues.
Structure Validation through FRET-derived Constraints

FIGURE 2. Overview of the side chain dihedral angles $\chi_1$–$\chi_4$ of AEDANS-labeled cysteine. The dihedral angles $\chi_1$, $\chi_2$, $\chi_3$, and $\chi_4$ were allowed to adopt $p$, $t$, and $m$ conformations. The dihedral angles $\chi_5$ and $\chi_6$ were fixed to a $t$ conformation. Bond lengths and bonding angles were taken from the crystal structure of 1,5-IAENS and the AMBER force field.

The model of a straight ideal $\alpha$-helix was built from the primary structure with the computer program Swiss PdbViewer, giving a perfect right-handed $\alpha$-helix with a pitch of 5.4 Å and a rise/residue of 1.5 Å. AEDANS-labeled cysteine was built in MOLMOL. When possible, the bond lengths and bond angles were taken from the crystal structure of 1,5-IAENS. The dihedral angles in the naphthalene group and in the sulfonic acid group were taken from the crystal structure as well. Missing bond lengths and bond angles were taken from the AMBER force field.

To account for the flexibility, an atomic model of the AEDANS label was introduced that explicitly takes into account all possible rotamers. For each mutant, an atomic model of AEDANS-labeled cysteine was written in MOLMOL when possible. When possible, the bond lengths and bond angles were taken from the crystal structure as well. Missing bond lengths and bond angles were taken from the AMBER force field.

For each mutant, an atomic model of AEDANS-labeled cysteine was inserted in the structure using the computer program MOLMOL, preserving the backbone dihedral angles of the original residue. Side chain dihedral angles have a strong preference for three conformations: $p$, $t$, and $m$. For simplicity, it is assumed that these conformations correspond to discrete values of $-60$, $-60$, and $180^\circ$. The side chain dihedral angles $\chi_1$, $\chi_2$, $\chi_3$, $\chi_4$, $\chi_5$, and $\chi_6$ were allowed to adopt all three values, generating a large set of possible rotamers.

For each rotamer, a separate structure file was written. Two dihedral angles, $\chi_5$ and $\chi_6$, involve a nitrogen atom that is either part of an amide bond ($\chi_5$) or is bonded to the aromatic ring system ($\chi_6$). These dihedral angles were fixed to $180^\circ$, being the value found in the crystal structure of 1,5-IAENS, because the flexibility of these bonds is limited due to their double bond character.

For the tryptophan side chain, the dihedral angles $\chi_1$ and $\chi_2$ were fixed to $180^\circ$ and $90^\circ$, respectively, representing the most abundant rotamer of tryptophan in $\alpha$-helices. A distribution of conformations of tryptophan was not taken into account, because the spatial translation of the tryptophan amino acid residue upon rotation about the $\chi_1$ and $\chi_2$ angles is relatively small.

All resulting structures were screened for unfavorable AEDANS conformations, i.e., with the label folded onto the backbone, onto the tryptophan, or onto itself. Hereto, the number of atom pairs with van der Waals overlap was determined. To save computer time, all hydrogen atoms were removed first. Heavy atoms of side chains other than tryptophan were removed as well, because Swiss PdbViewer placed them in a rather arbitrary way. Of the remaining atoms, an atom pair with an interatomic distance $<0.8 \times$ the sum of the combined van der Waals radii was regarded as an overlapping atom pair. A maximum of two overlapping atom pairs/structure was allowed, otherwise the structure was discarded. Typically, ~80% of the structures was discarded based on van der Waals overlap. Assuming that all allowed conformations are equally probable, the energy transfer contribution for AEDANS conformation $i$ ($E_i$) is given by Equation 2. The calculated energy transfer taking into account all conformations is as follows,

$$E_{\text{theory}} = \frac{1}{N} \sum_{i=1}^{N} E_i \quad \text{(Eq. 7)}$$

where $N$ is the number of allowed conformations for a specific mutant. The corresponding distance in Equation 2, $r_i$, is calculated from the middle of the central bond of the AEDANS naphthalene group to the middle of the central bond of the tryptophan indole group.

**RESULTS**

The experimental energy transfer efficiencies are depicted in Fig. 3. Because of the high lipid to protein ratios used, we assumed that the energy transfer efficiencies were only due to intramolecular energy transfer. To justify this assumption, we performed some calculations based on a random distribution of wild type and mutant protein in the membrane. In this case, the contribution due to intermolecular energy transfer was small for lipid to protein ratio 400 (9% for mutant A9C, 13% for mutant G38C) and negligible (<1%) for lipid to protein ratio 1500. We therefore assumed that the contribution of intermolecular energy transfer could be ignored, although a small contribution could be still present in the case of lipid to protein ratio 400. The experimental extinction coefficients in Equation 6 are the main source for errors in the energy transfer efficiencies in Fig. 3. Therefore, error bars were calculated based on an uncertainty of ±200 in the extinction coefficients.

The energy transfer efficiencies become close to 1.0 for AEDANS-labeled mutants near the tryptophan at position 26, which is located in the center of the protein, and are almost zero for mutants in both termini. On top of this overall trend, the efficiencies show oscillations that are indicative for a helical conformation of the protein.

The efficiency is related to the distance between the donor and acceptor via Equation 2. The donor-acceptor distance is equal to the Förster radius when the efficiency is 0.5. As can be seen in Fig. 3, this is the case for labels around position 10 and 42. This implies that the distance from these residues to tryptophan at position 26 is ~24 Å, being the Förster radius of the tryptophan–AEDANS donor-acceptor pair. Both positions are 16 amino acid residues away from the tryptophan at position 26. A perfect $\alpha$-helix has a rise of 1.5 Å/residue (33), giving a translation of 24
Structure Validation through FRET-derived Constraints

FIGURE 3. Energy transfer efficiencies calculated with models A (black), B (green), C (blue), and D (red). The experimental efficiencies are depicted as red dots. The models are given in Fig. 4.

FIGURE 4. Different models to calculate the distance between the donor and acceptor, r, and their model quality parameters, Q. Model A, the distance between the C-α atoms is taken as r. Model B, the size of the donor and acceptor in a fully extended conformation is estimated based on a molecular model. Model C, the donor and acceptor size are taken from the literature (2). Model D, the flexibility of the acceptor is taken into account, and a separate energy transfer efficiency is calculated for each conformation of the acceptor. Subsequently, an average energy transfer efficiency is taken over all possible conformations.

FIGURE 5. Quality parameters Q of Protein Data Bank structures. 2CPB (1–25, blue) (11), 2CPA (26–50, purple) (11), 1MZT (51, yellow) (14), 1FJ (52, black) (12), and the straight helix (53, red) proposed in this work are shown. The first 50 structures were determined in DPC micelles (1–25) and SDS micelles (26–50) with high resolution solution NMR. Structure 51 was obtained in oriented bilayers with solid state NMR and structure 52 in the bacteriophage with x-ray diffraction.

Å for 16 residues. This calculation shows that the efficiencies can be explained qualitatively on the basis of an α-helical structure for the protein.

Model Construction—Based on this finding, a molecular model of the protein in an α-helical conformation is constructed, and the theoretical efficiencies are calculated using four increasingly sophisticated models (A–D) to represent the donor and acceptor (Fig. 4). In model A, the sizes of the donor and the acceptor are not taken into account. The trend of the experimental data is well reflected by this model (Fig. 3, black line), but the calculated values are in general too high. Moreover, the oscillations that are clearly visible in the experimental data are hardly present in the calculated curve. To account for this, the sizes of the donor and acceptor are taken into account explicitly in model B. In this model, these sizes were estimated based on molecular models, assuming that they are in a fully extended conformation (Fig. 4, model B). In this case, the efficiencies become too low and the oscillations become too strong, indicating that the size of the donor and/or acceptor is overestimated (Fig. 3, green line). To quantify the performance of the models, a model quality parameter Q is introduced.

\[
Q = \sum_{j=1}^{M} (E_{\text{experiment},j} - E_{\text{calc},j})^2
\]  

(Eq. 8)

Here, j runs over the mutants, M is the total number of mutants, and \( E_{\text{experiment}} \) and \( E_{\text{calc}} \) are the experimental and the calculated efficiencies. As can be seen in Fig. 4, model B performs not as well as model A.

Previously, the effective size of tryptophan- and AEDANS-labeled cysteine were estimated to be 6.5 and 8.0 Å, respectively (2). The effective size of the donor is in close agreement with the size of a fully extended tryptophan, which was estimated here to be 6.6 Å. The size of AEDANS-labeled cysteine (8.0 Å), however, is much smaller than the value of 16.0 Å estimated from the molecular model. This indicates that the acceptor is not in a fully extended conformation. When these sizes are used to calculate the theoretical efficiencies, they are in good agreement with the experimental ones (Fig. 3, blue line), and the value of Q reduces again (Fig. 4, model C).

Apparently, the AEDANS label is not in a fully extended conformation. Therefore, we proposed model D in which AEDANS is represented as a distribution of conformations, giving rise to a smaller effective size of the acceptor. It can be seen (Fig. 3, red line) that model D gives excellent theoretical efficiencies in both the N- and C-terminal parts of the protein, leading to a further reduction of the value of Q (Fig. 4, model D). This implies that the protein is well described by an α-helix when the conformational space of the acceptor is taken into account.

However, various conformations of membrane-bound coat protein have been proposed previously based on x-ray diffraction, solution NMR, and solid state NMR. Therefore, the same approach as described for model D was used to calculate the theoretical efficiencies for the known structures in the Protein Data Bank. The resulting quality parameters are depicted in Fig. 5. TABLE ONE gives an overview of the structures of the coat protein that were used in the comparison to the energy transfer data using model D.

DISCUSSION

Micellar Stress—The first fifty structures in Fig. 5, blue and purple bars, were calculated from solution NMR in detergent micelles (see Fig. 6A) (11). For the majority of these structures, the calculated energy transfer efficiencies of mutants in the C-terminal part of the protein are in good agreement with our energy transfer data. The main reason for
the high variation in \( Q \) values arises from the orientation of the N-terminal helix that is poorly defined with respect to the C-terminal helix. Highly curved micelles are known to induce stress on single \( \alpha \)-helices passing through a micelle. For instance, micelles were shown to cause considerable curvature in a 22-amino-acid-long micelle-bound peptide (37). Chou et al. (37) anticipated that for longer \( \alpha \)-helices, the environmental stress would be even larger. We propose that micelle-bound coat protein, which is 50 amino acids long, reduces this stress through partial unwinding of the helix around residues 20–24, allowing the protein to adapt to the highly curved micelle surface. Because of this structural disorder, fewer NOEs are observed in this region, hampering the determination of an unambiguous protein structure (11).

Dehydrated Bilayer Stress—Remarkably, the structure found with solid state NMR in stacked, oriented bilayers (14) has a particularly high \( Q \) value (see Fig. 5, yellow bar). This structure consists of two helices, one running from residue 7 to 20, the other from 21 to 44, and is best described as L-shaped (see Fig. 6B). The helices are connected via a short turn. In addition, the latter helix has a distinct kink near residue 39. In our energy transfer data, there is no indication of such a turn in the N-terminal part of the protein. We ascribed the difference in conformation of the protein in vesicles and in oriented bilayers to insufficient hydration of the lipid headgroups in oriented bilayers. Stacked, oriented bilayers are prepared by drying of a protein-lipid vesicle suspension on a glass plate and subsequent rehydration of the sample at high relative humidity. When the relative humidity is 100%, the water spacing between the stacks becomes ~18.0 Å (38), and the lipid leaflets are fully hydrated. However, when the relative humidity is slightly lower, the water spacing between the lipid leaflets decreases rapidly (to 3.6 Å at 97% relative humidity). The stacked, oriented bilayers that were used for solid state NMR in the case of the coat protein were rehydrated at an even lower relative humidity (93% (39)), suggesting a water layer <3.6 Å. Such a small water layer can impose a severe environmental stress on the N-terminal part of the protein, thereby changing its conformation. This stress is absent in unilamellar vesicles in which the lipid leaflets are fully hydrated, and the polar residues in the N-terminal part of the protein can interact with water molecules, allowing an L-shaped conformation as found in our work (Fig. 6D).

Phage-bound Coat Protein—The structure of phage-bound coat protein (12) has a relatively low \( Q \) value (Fig. 5, black bar). This structure is a continuous, slightly curved \( \alpha \)-helix (Fig. 6C). Throughout the entire structure, the calculated energy transfer efficiencies are in good agreement with the experimental ones, showing that the conformation in vesicles is not very different from the phage-bound conformation. This is consistent with the model of Marvin (13), who anticipated that the membrane-bound conformation of the coat protein could not be very different from the phage-bound conformation. This view is further supported by the fact that the \( Q \) value of a straight \( \alpha \)-helix based on FRET-derived distance constraints is also low.

\begin{table}
\centering
\caption{Overview of the structures of the coat protein that were used in the comparison to the energy transfer data using model D}
\begin{tabular}{|c|c|c|c|c|}
\hline
PDB entry & Number of structures & Experimental conditions & Method & Type of filamentous phage & Reference \\
\hline
2CPS & 25 & SDS micelles & Solution NMR & M13 & 11 \\
2CPB & 25 & DPC micelles & Solution NMR & M13 & 11 \\
1MZT & 1 & Oriented bilayers & Solid state NMR & Fd & 14 \\
1HIF & 1 & Intact phage & X-ray diffraction & Fd & 12 \\
\hline
\end{tabular}
\end{table}

* M13 coat protein has an asparagine residue at position 12, Fd coat protein (an aspartic acid residue). The rest of the protein is identical.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Different schematic models of the membrane-bound form of phage coat protein. A, set of 25 structures of the coat protein obtained in micelles (11). The backbone atoms of residues 26–40 were fitted to each other, and the putative transmembrane domain was placed to span the hydrophobic region of a bilayer, which is ~30 Å and is depicted in yellow. B, the protein in stacked, oriented bilayers with the N-terminal helix oriented perpendicular to the membrane normal (14). For reference, the tryptophan residue at position 26 is represented as a blue sphere, and the lysozyme residues at positions 40, 43, and 44 are represented as green spheres. In this model, the charged lysine residues, which have relatively long side chains, are able to interact with the membrane-water interface, even at position 40. C, the coat protein is represented as a slightly curved, continuous helix at an angle of ~39° to the membrane normal (13). The protein is oriented in such a way that the lysine residues are located at the membrane-water interface. D, protein model emerging from the FRET-derived constraints. The charged lysines in the C terminus are placed at the interface (40), and the protein adopts a tilt angle of 23° (2).}
\end{figure}

It should be noted, however, that the number of FRET-derived constraints in the N-terminal part of the protein may be insufficient to distinguish between some extended structures found in micelles and a continuous \( \alpha \)-helix. Furthermore, some dynamics and curvature of the structure might still be present, permitting charged residues in the N-terminal part of the protein to interact with the membrane-water interface, allowing a slight deviation from an ideal \( \alpha \)-helix. Nevertheless, the \( \alpha \)-helix model depicted in Fig. 6D is in good agreement with recent findings of our group describing the structure and membrane assembly of the coat protein in phospholipid membranes (2, 4).

CONCLUSIONS

Based on the data presented here, we conclude that M13 major coat protein bound to vesicles is well described by a continuous \( \alpha \)-helix,
resembling the l-shaped structure found in the bacteriophage. It should be noted that some extended structures of the coat protein in SDS micelles are also in good agreement with the energy transfer data presented here and that the number of FRET-derived distance constraints in the N-terminal domain is insufficient to distinguish between extended structures and a continuous α-helix. The presence of L-shaped and U-shaped structures in vesicles can be excluded based on our FRET-derived distance constraints.

In this work, we have shown that FRET offers a useful alternative for structure determination of membrane proteins, provided a sufficient number of single cysteine mutants can be obtained. In addition, the fluorescence technique is sensitive and thus not very demanding in terms of protein quantities needed.

Our findings emphasize the need to study membrane proteins in a suitable environment, such as fully hydrated vesicles. Although larger proteins than M13 major coat protein may be able to handle environmental stress in a different way, any membrane protein with water-exposed parts in the C or N termini and hydrophilic loop regions should be treated with care.

REFERENCES

1. Torres, J., Stevens, T. J., and Samso, M. (2003) Trends Biochem. Sci. 28, 137–144
2. Koehorst, R. B. M., Spruijt, R. B., Vergeldt, F. J., and Hemminga, M. A. (2004) Biophys. J. 87, 1445–1455
3. Narazov, P. V., Apanasovich, V. V., Lutkovski, V. M., Yatskou, M. M., Koehorst, R. B. M., and Hemminga, M. A. (2004) J. Chem. Inf. Comput. Sci. 44, 568–574
4. Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (2004) Biochemistry 43, 13972–13980
5. Štrancar, J., Koklic, T., Arsov, Z., Filipic, B., Stopar, D., and Hemminga, M. A. (2005) J. Chem. Inf. Model. 45, 394–406
6. Lakowicz, J. R. (1999) Principles of Fluorescence Spectroscopy, 2nd Ed., pp. 367–394, Kluwer Academic/Plenum Publishers, New York
7. Stopar, D., Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (2003) Biochem. Biophys. Acta 1611, 5–15
8. Hemminga, M. A., Sanders, J. C., and Spruijt, R. B. (1992) Prog. Lipid Res. 31, 301–333
9. Hemminga, M. A., Sanders, J. C., Wolfs, C. J. A. M., and Spruijt, R. B. (1993) in New Comprehensive Biochemistry (Watts, A., ed) pp. 191–212, Elsevier, Amsterdam
10. Shon, K. J., Kim, Y. G., Colnago, L. A., and Opella, S. J. (1991) Science 252, 1303–1304
11. Papavonie, C. H. M., Christaas, B. E. C., Folmer, R. H. A., Konings, R. N. H., and Filbers, C. W. (1998) J. Mol. Biol. 282, 401–419
12. Marvin, D. A., Hale, R. D., Nave, C., and Citterich, M. H. (1994) J. Mol. Biol. 235, 260–286
13. Marvin, D. A. (1998) Cur. Opin. Struct. Biol. 8, 150–158
14. Marassi, F. M., and Opella, S. J. (2003) Protein Sci. 12, 403–411
15. Spruijt, R. B., Meijer, A. B., Wolfs, C. J. A. M., and Hemminga, M. A. (2000) Biochim. Biophys. Acta 1509, 311–323
16. Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (1989) Biochemistry 28, 9158–9165
17. Eaton, D. F. (1988) Pure Appl. Chem. 60, 1107–1114
18. Lakshmikanth, G. S., Sridevi, K., Krishnamoorthy, G., and Udgaonkar, J. B. (2001) Nat. Struct. Biol. 8, 799–804
19. Datema, K. P., Visser, A. J. W. G., Van Hoek, A., Wolfs, C. J. A. M., Spruijt, R. B., and Hemminga, M. A. (1987) Biochemistry 26, 6145–6152
20. Johnson, I. D., and Hudson, B. S. (1989) Biochemistry 28, 6392–6400
21. Fernandes, F., Loura, L. M. S., Prieto, M., Koehorst, R. B. M., Spruijt, R. B., and Hemminga, M. A. (2003) Biophys. J. 85, 2430–2441
22. Salamon, Z., Lindblom, G., Rilfors, L., Linde, K., and Töllin, G. (2000) Biophys. J. 78, 1400–1412
23. Fisher, C. A., and Ryan, R. O. (1999) J. Lipid Res. 40, 93–99
24. Anderluh, G., Hong, Q., Boetzel, R., MacDonald, C., Moore, G. R., Virden, R., and Lakey, J. H. (2003) J. Biol. Chem. 278, 21860–21868
25. Fung, K. F., and Stryer, L. (1978) Biochemistry 17, 5241–5248
26. Hudson, E. N., and Weber, G. (1973) Biochemistry 12, 4154–4161
27. Edelhoch, H. (1967) Biochemistry 6, 1948–1954
28. Eisenhawer, M., Cattarinussi, S., Kuhn, A., and Vogel, H. (2001) Biochemistry 40, 12321–12328
29. Lindahl, E., Hess, B., and van der Spoel, D. (2001) J. Mol. Mod. 7, 306–317
30. Koradi, R., Billeter, M., and Wuthrich, K. (1996) J. Mol. Graph. 14, 51–55
31. Guex, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714–2723
32. Wall, L., Christiansen, T., and Schwartz, R. L. (1996) Programming Perl., 2nd Ed., O’Reilly Media, Sebastopol, CA
33. Creighton, T. E. (1983) Proteins, Freeman, New York
34. Baudet-Neuiler, S., Juliens, M., Crosio, M.-P., and Janin, J. (1993) Biochemistry 32, 8457–8464
35. Cornell, W. D., Cieplak, C. I., Bayly, I. R., Gould, K. M., Merz, D. M., Jr., Ferguson, D. C., Spellmeyer, T., Fox, J. W., Caldwell, J. W., and Kollman, P. A. (1995) J. Am. Chem. Soc. 117, 5179–5177
36. Lovell, S. C., Word, J. M., Richardson, J. S., and Richardson, D. C. (2000) Proteins Struct. Funct. Genet. 40, 389–408
37. Chou, J. J., Kaufman, J. D., Stahl, S. J., Wingfield, P. T., and Bax, A. (2001) J. Am. Chem. Soc. 123, 2450–2451
38. Tristram-Nagle, S., and Nagle, J. F. (2004) Chem. Phys. Lipids 127, 3–14
39. Marassi, F. M., Ramamoorthy, A., and Opella, S. J. (1997) Proc. Natl. Acad. Sci. 94, 8551–8556
40. Meijer, A. B., Spruijt, R. B., Wolfs, C. J. A. M., and Hemminga, M. A. (2001) Biochemistry 40, 8815–8820