Review

From 'I' to 'L' and back again: the odyssey of membrane-bound M13 protein

Werner L. Vos[1], Petr V. Nazarov[2], Rob B.M. Koehorst[3], Ruud B. Spruijt[3] and Marcus A. Hemminga[3]

[1] Department of Biology, National University of Ireland Maynooth, Co. Kildare, Ireland
[2] Microarray Center, CRP-Santé, 84, Val Fleuri, L-1526, Luxembourg
[3] Laboratory of Biophysics, Wageningen University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

The major coat protein of the filamentous bacteriophage M13 is a surprising protein because it exists both as a membrane protein and as part of the M13 phage coat during its life cycle. Early studies showed that the phage-bound structure of the coat protein was a continuous I-shaped α-helix. However, throughout the years various structural models, both I-shaped and L-shaped, have been proposed for the membrane-bound state of the coat protein. Recently, site-directed labelling approaches have enabled the study of the coat protein under conditions that more closely mimic the *in vivo* membrane-bound state. Interestingly, the structure that has emerged from this work is I-shaped and similar to the structure in the phage-bound state.

M13 coat protein: an overview

Filamentous phages (of the family Inoviridae) are small bacteriophages and differ from most other bacterial viruses in that they reproduce continuously upon infection without directing lysis of their host cells. Instead, infected cells continue to grow and divide, although at a lower rate than normal. Filamentous phages efficiently infect only host cells that express (sex) pili, which serve as receptor sites. Based on their pilus specificity, different classes of evolutionarily closely related filamentous phages can be distinguished. F-specific filamentous phages (the Ff group) only infect male *Escherichia coli* cells carrying sex pili, encoded by an F factor (i.e. a fertility or sex episome). Among these Ff bacteriophages, the nearly identical strains M13 and fd (see Glossary) are the most extensively studied and, together with their *E. coli* host strains, the best characterized from biochemical, genetics and biophysical points of view.

A single M13 viral particle has a diameter of ~6.5 nm and is 900 nm long. The filament contains a circular, single-stranded viral DNA genome that is protected by a long cylindrical protein coat, which is predominantly made up of ~2700 copies of the major coat protein (i.e. gp8, the protein coded by the viral gene 8). The major coat proteins, in an overlapping helical array, form a tube around the viral DNA, with the N terminus located at the outside of the coat and the lysine residues of the C terminus interacting with the phosphate groups of the DNA at the inside of the coat (Table 1). The hydrophobic domain of the major coat protein is located in the central section of the protein sequence, and it interlocks the coat protein with its neighbouring coat proteins. Together with a few minor coat proteins (the products of the viral genes 3, 6, 7 and 9), M13 phage is a very stable nucleoprotein particle.

The reproductive life cycle of bacteriophage M13 begins with the adsorption of the phage to the receptor site located at the F pilus of host *E. coli*. After destabilization, the phage particle can be disassembled. The coat proteins are stripped from the nucleoprotein particle and subsequently deposited in the *E. coli* inner (or cytoplasmic) membrane. During this process the phage DNA genome is injected into the cell cytoplasm. Infected *E. coli* cells facilitate viral DNA replication and viral protein production. The assembly of new viral particles occurs in the bacterial membrane at so-called assembly sites, which resemble adhesion zones between the inner and outer membranes and are formed by the viral proteins coded by the genes 1, 4 and 11. During the assembly of new progeny phage particles, the coat proteins fit together and wrap around the phage DNA, thereby forming the coat of the new assembling phage. After termination of phage assembly, the filamentous phage is released into the medium and the journey to another host cell can begin.

During the reproductive life cycle, the major coat protein is involved in various processes that take place

Glossary

14:1PC: 1,2-dimyristoleoyl-*sn*-glycero-3-phosphocholine. This phospholipid contains two short C14 acyl chains that form lipid bilayers with a hydrophobic thickness (thickness of hydrocarbon chain region) of 2.3 nm [27] flanked on both sides by a 1.5 nm head group region [26].
18:1PC: 1,2-dioleoyl-*sn*-glycero-3-phosphocholine. This phospholipid contains two medium-length C18 acyl chains that form lipid bilayers with a hydrophobic thickness of 3.0 nm [27] flanked on both sides by a 1.5 nm head group region [26].
22:1PC: 1,2-dierucoyl-*sn*-glycero-3-phosphocholine. This phospholipid contains two long C22 acyl chains that form lipid bilayers with a hydrophobic thickness of 3.7 nm [27] flanked on both sides by a 1.5 nm head group region [26].
SDL: site-directed labelling with a probe of choice to be used in fluorescence or electron spin resonance (ESR) spectroscopy (Box 2).
M13 and fd: related filamentous bacteriophages that belong to Inoviridae, a family of phage, which infect Gram-negative bacteria that contain pili (i.e. *Escherichia coli* and *Pseudomonas aeruginosa*).

Corresponding author: Hemminga, M.A. (marcus.hemminga@wur.nl).

0968-0004/$ – see front matter © 2009 Elsevier Ltd. All rights reserved. doi:10.1016/j.tibs.2009.01.007 Available online 8 April 2009

Table 1. Classification of important domains in the M13 major coat protein

Amino acid residue	Classification	Domain type	Refs
1–6	Acid	Membrane: hydrophilic anchor that contains negatively charged amino acid residues (i.e. Glu2, Asp4 and Asp5) that will extend into the aqueous phase. Phage: these negatively charged residues dictate bacteriophage solubility in aqueous solution.	[6,19]
7–20	Amphipathic	Membrane: membrane–water interface domain. Phage: covers and shields the hydrophobic surface of underlying protein units (the coat proteins pack like roof tiles).	[6,19]
21–39	Hydrophobic	Membrane: transmembrane domain. Phage: these residues promote tight protein–protein interactions and, thus, bacteriophage stability.	[6,19]
40–50	Basic	Membrane: membrane–water interface anchor by binding of Lys40, Lys43 and Lys44 to the phospholipid headgroup region (termed snorkelling; i.e. they can bury themselves with their aliphatic portion in the hydrophobic region of the lipid bilayer while positioning the charged amino group in the more polar interface) together with a firm interaction of Phe42 and Phe45 with the hydrocarbon chain region of the membrane (termed anti-snorkelling). Phage: the lysine residues are involved in DNA binding by charge neutralization; the phenylalanines participate in intercalating protein–protein interactions.	[6,19,35]

in different environments. First, during the infectious entry into the *E. coli* host cell, the major coat protein is stripped from the phage particle and undergoes deposition into the inner membrane. Second, new coat protein is synthesized as a procoat, a precursor of the coat protein containing an additional amino acid leader sequence necessary for insertion into the membrane. The procoat molecule is inserted into the membrane and subsequently the extra leader sequence is removed by a host cell leader peptidase. The resulting mature transmembrane coat protein is stored in the inner membrane before its use in the phage assembly process. Finally, when parental and newly synthesized coat proteins take part in the complex process of membrane-bound phage assembly and phage extrusion, the coat protein is transferred from a membrane environment to the protective coat that surrounds the viral genome [1].

Despite this remarkable versatility, the coat protein contains only 50 amino acid residues (Figure 1), and for years scientists have pondered how such a small protein can exist under such diverse chemical and physical conditions. To answer this question, the 3D structure of gp8 is of great importance, and several attempts have been undertaken to tackle this challenge. A variety of structural models for the protein in the membrane-bound state have been proposed, varying from I-shaped to L-shaped and even U-shaped structures (Figure 2). These models are discussed in this review, both in the historical context and in light of present-day literature. We also outline the techniques used to calculate the different models, together with their advantages and drawbacks, following the M13 major coat protein on its odyssey through the ever-changing landscape of the world of scientific ideas and techniques.

From 'I'...

Part of the confusion surrounding the structure of the M13 major coat protein and its closely related structural homologue fd coat protein was addressed in the early 1990s when a model of the phage-bound coat protein was calculated based on X-ray fibre diffraction experiments [2,3]. These experiments revealed that the phage-bound structure of the coat protein forms an almost perfect α-helix resembling the letter 'I'. In this model, 4–5 flexible unstructured amino acid residues in the N terminus protrude from the phage coat into the aqueous phase to maintain the phage particle in solution, whereas the lysine-rich C terminus interacts with the viral DNA phosphate groups [3] (Table 1). Recently this structure was further refined by cryo-electron microscopy [4] and solid-state nuclear magnetic resonance (NMR) approaches [5]. However, this model posed a puzzle because it was evident that the α-helix was too long to fit into the *E. coli* host membrane after infection and before phage assembly. In considering this discrepancy, it was pointed out that a slightly different membrane incorporation, in which the I-shaped protein adopts a tilt with respect to the membrane normal, would enable favourable embedding of a continuous α-helix into the membrane [6] (Figure 2a).

Although X-ray crystallography is arguably the most powerful technique when studying structure–function relations of membrane proteins, the approach comes with an inherent drawback: proteins can be studied only in a crystal lattice. Unfortunately, the coat protein could not be crystallized in a membrane-bound state, so alternative strategies were sought to determine its structure in membrane systems (Box 1).

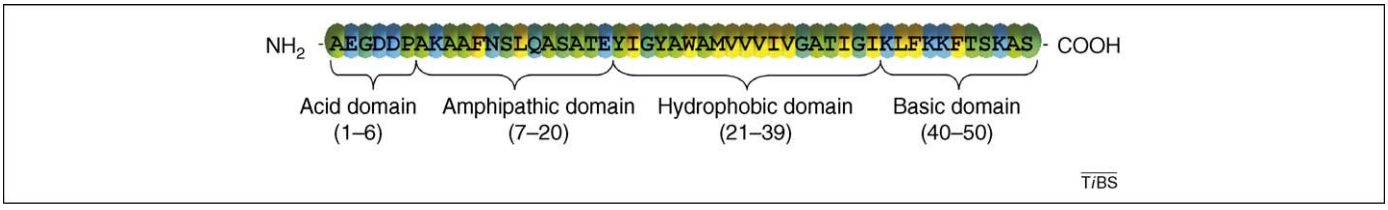

Figure 1. Primary structure of M13 major coat protein with the classification of the important domains. The colour coding is based on amino acid residue hydrophobicity scales [26] with the yellow colour corresponding to hydrophobic residues, green to neutral, and blue to charged residues.

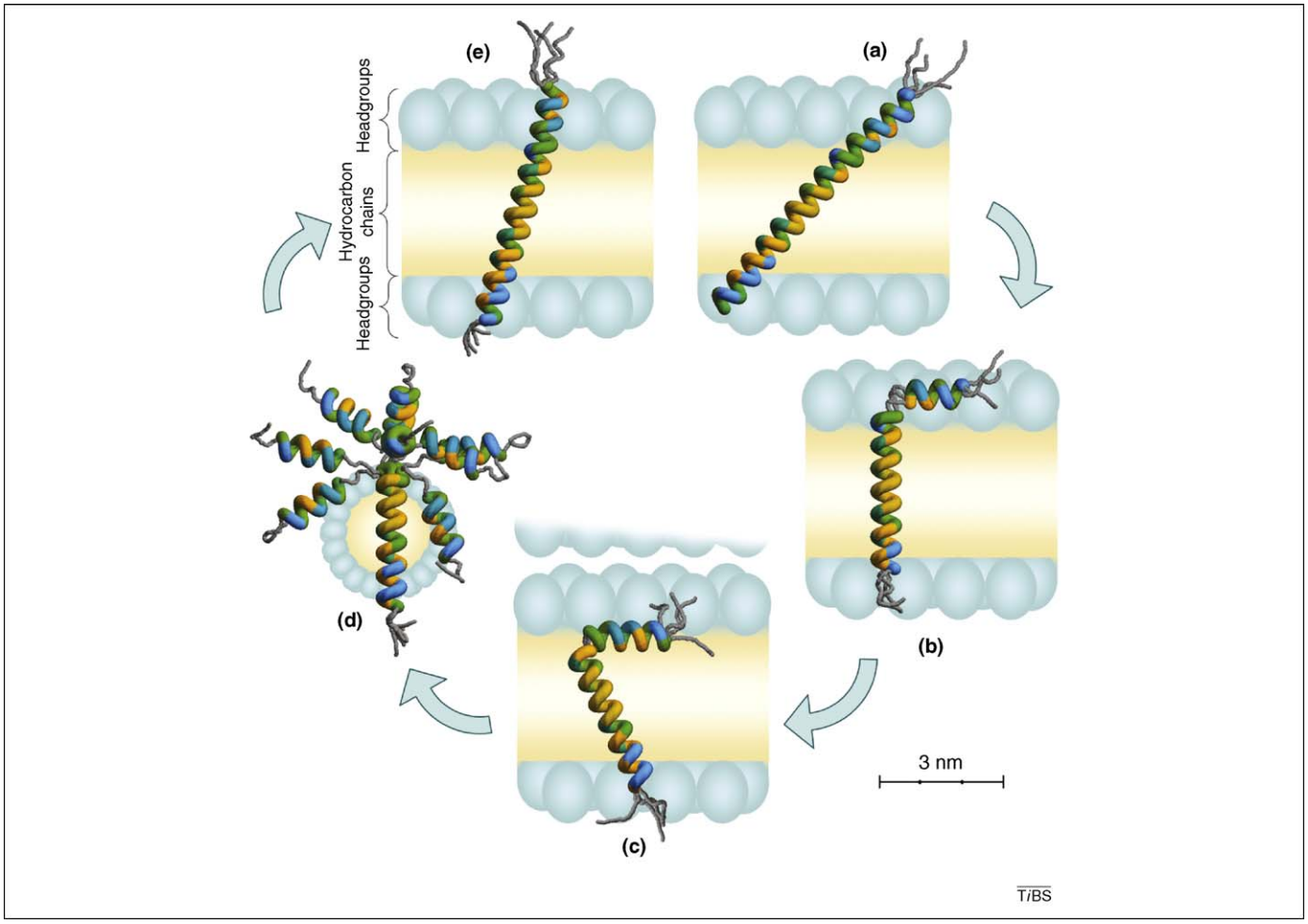

Figure 2. Odyssey of M13 coat protein through the landscape of structural concepts from 'I' to 'L' and back again. The colour coding of the amino acid residues is based on a hydrophobicity scale (see Figure 1). Unstructured protein regions are indicated in grey. **(a)** Original phage-bound I-shape model [3] in its membrane-bound state [6]. The size of the membrane regions is obtained from the literature, with the positions of the carbonyls serving as borders for the headgroup region [26,27]. The phospholipid headgroups are indicated with blue ellipsoids and the hydrocarbon chain region is coloured in yellow. **(b)** L-shape configuration as determined by NMR [7]. The protein has two α-helices with the amphipathic helix parallel to the plane of the bilayer and the hydrophobic helix (Table 1) perpendicular to the plane of the bilayer. **(c)** Tilted L-shape model determined by solid state NMR [10]. The side view shows a short amphipathic (Table 1) in-plane helix that rests on the membrane surface and the 26° tilt of the longer hydrophobic (Table 1) transmembrane helix. The experiments were performed in dehydrated bilayers, and the adjacent bilayer is schematically illustrated. **(d)** 'Tree' of possible two-helix structures as determined by solution NMR [13]. The size of the sodium dodecylsulphate micelle is approximated from its chemical structure. **(e)** Structure and membrane embedding of the M13 coat protein in fully hydrated vesicles of 18:1PC (and mixed phospholipid systems with C18 acyl chains), based upon recent SDL spectroscopy [17–21]. The protein forms a mainly α-helical conformation tilted by 18° with respect to the membrane normal. The first nine amino acid residues, which encompass the hydrophilic anchor (Table 1), are unstructured. The resulting membrane-embedded M13 coat protein structure (e) does not differ much from the native α-helical structure of the protein in bacteriophage M13 (a). The protein survives the membrane-bound state by a simple tilt mechanism and a subtle structural adjustment in the extreme end of the N terminus.

...To 'L'...

Seemingly, the conundrum was solved when a model was proposed based on NMR studies of the closely related major coat proteins of bacteriophages Pf1 and fd [7,8]. The experiments encompassed solution NMR on detergent-solubilized proteins in micelles and solid-state NMR experiments on the protein embedded in oriented lipid bilayers. The resulting model divided the protein into four specific domains: (i) an N-terminal acidic domain, (ii) an amphipathic domain, (iii) a hydrophobic domain, and (iv) a C-terminal basic domain (Table 1). The rationale behind this model was that the hydrophobic helix, together with the C-terminal domain, would span the lipid membrane, whereas the amphipathic helix would float on the membrane with the hydrophobic residues exposed to the lipids and the hydrophilic residues exposed to the water phase. In this model, the amphipathic helix is oriented almost perpendicular to the transmembrane helix; hence, it was termed 'L-shaped' (Figure 2b). The L-shaped model requires that a dramatic re-arrangement occurs upon incorporation of coat protein into the phage during its membrane-bound assembly, going from an L-shaped conformation to an almost perfect helix, resembling an I-shaped conformation.

The L-shaped model dominated the literature for years, although it was pointed out that a tilt of the transmembrane α-helical domain with respect to the membrane normal would also enable favourable membrane embedding [6]. Indeed, such a tilt was observed through the use of site-specific solid-state NMR experiments [9] in addition to other solid-state NMR approaches [10] (Figure 2c). Nevertheless, even in the most recent work, the L-shaped model prevailed [11], implying that the coat protein would require both a tilted transmembrane helix and an L-shaped conformation to be stably incorporated into a lipid bilayer.

Box 1. Determining membrane protein structure

One of the most challenging problems in structural biology is the solving of the structure and function of membrane proteins [33]. Various techniques are available for structure determination: X-ray crystallography, electron microscopy, high-resolution and solid-state NMR, site-specific infrared dichroism and site-directed spin and fluorescence labelling. X-ray crystallography and high-field solution NMR spectroscopy are the primary techniques used in determining the structures of water-soluble proteins, but in dealing with membrane proteins that must be embedded in an amphipathic environment there is not yet a well-defined strategy for obtaining a protein structure.

- Despite its ongoing success, X-ray crystallography of membrane proteins remains limited by difficulties in crystallizing membrane proteins from detergent solutions.
- Progress is being made in electron microscopic analysis of 2D crystals but, in practice, such crystals are often not ordered well enough to provide sufficient resolution. However, single-particle approaches have recently enabled the construction of a particularly highly ordered polymer, tobacco mosaic virus, at ~4.7 Å resolution (for a review, see Ref. [36]).
- High-resolution solution NMR is carried out in micellar solutions, which provide a membrane-like environment. However, the technique requires expensive isotope labelling (e.g. ^{2}H, ^{13}C or ^{15}N) and is limited to proteins with a low molecular weight (<40 kDa). As a compromise between micelles (used in solution NMR) and bilayers (used in solid-state NMR), bicelles (disk-shaped phospholipid-detergent aggregates) can provide an alternative membrane-like environment.
- Solid-state NMR can be used when crystallization is not possible, or when the size of the protein-detergent micelles exceeds the level that can be tackled with solution NMR. Solid-state NMR is used for isotope-labelled membrane proteins reconstituted in lipid bilayers, often packed into oriented multibilayers at low water content. Both NMR techniques can provide high-resolution atomic scale structural information.
- Site-specific infrared dichroism (SSID) requires specific incorporation of isotopes in the protein of interest, which usually poses a large challenge. This technique has been used for orientational analysis of transmembrane α-helices in aligned bilayers.
- SDL techniques, using ESR spin labels and fluorescent labels, are based on site-directed cysteine mutations of the membrane proteins and subsequent labelling of the protein.

When compared with the emerging models from solid-state NMR, the structures from solution NMR in detergent micelles provided a somewhat less clear-cut picture. The structures of M13 and fd coat protein in detergent micelles, as obtained by high-resolution solution NMR spectroscopy, showed two α-helical regions for the M13 coat protein comprising residues 8–16 and 25–45, respectively [7,12,13]. These studies resulted in a 'tree' of 50 structures, in which the two well-defined helices have vastly different orientations relative to each other throughout the individual structures. Some of the structures are extended, resembling an I-shape, whereas other structures are best described as L-shaped, or U-shaped [13] (Figure 2d). Curiously, the biologically most important experimental observation was the lack of close proton distances in the region connecting the two helices, which was ascribed to protein backbone dynamics. The term 'hinge region' was coined to indicate a flexible region spanning amino acid residues 17–24, which enables a conformational transition from an L-shaped or U-shaped protein to the I-shaped protein as found in the phage. Thus, the results agreed with existing ideas that proposed the existence of an L-shaped coat protein conformation when embedded in the lipid membrane before phage assembly and after infection [13].

At the same time, alternative models emerged based on site-directed labelling (SDL) techniques. Studies based on electron spin resonance (ESR) spin label and fluorescence spectroscopy indicated a dynamic exchange between the membrane-bound conformation and more extended structures [14,15], whereas a 'banana-shaped, but continuous helix' conformation was proposed based on additional fluorescent labelling studies [16]. In this latter model the proposed tilted transmembrane helix and amphipathic helix were bridged to one continuous, but slightly bent, helix.

...And back again

In the following years it became increasingly clear that the L-shaped model (based on NMR spectroscopy) could not be confirmed by SDL techniques (Box 2). The first indication of this difference came from an optical spectroscopic study that used fluorescent labels attached at different positions throughout the protein [17]. Unlike NMR experiments, which require the use of micelles or oriented bilayers,

Box 2. Site-directed labelling techniques for structure determination

- Site-directed labelling (SDL) techniques are new tools in structural biology that can provide intra- and intermolecular distances, in addition to local polarity information. Apart from their general features (Box 1), SDL methods have some key advantages over other available techniques.
- Site-directed fluorescence labelling (SDFL): fluorescence spectroscopy is a technique with a very high sensitivity that can be used at low molecular concentrations (~10 nM). This technique offers an advantage in determining the structure of membrane proteins because possible artefacts coming from un-natural membrane-mimicking environments (i.e. aggregation, dehydrated bilayers, micelles) can be minimized by working at low protein concentrations (high lipid-to-protein ratios) in unilamellar vesicles.
- Site-directed spin label ESR (SDSL-ESR): this type of spectroscopy is also tremendously useful for studying the structure of membrane proteins because the proteins can be examined in a native-like membrane bilayer environment. Experiments can be performed at room temperature under physiological conditions, thereby preserving native lipid–protein interactions. Additional advantages include a high sensitivity (molecular concentrations are ~10 μM as compared with NMR spectroscopy that requires at least 1 mM) and effectively no limit with regard to the molecular weight of the samples.
- The downside of SDL is the need to introduce non-natural labels in the protein, which could replace a crucial amino acid residue. Moreover, the presence of such a label could disturb the system itself. Indeed, it has been reported that certain spin labels can occasionally induce peptide orientations that differ from those adopted by the wild-type peptide [37]. For this reason, when using SDL approaches it is preferable to avoid replacing crucial amino acid residues, for instance by selecting solvent-exposed loop regions of membrane proteins where structural perturbations owing to the presence of labels is minimized (see, for instance, Ref. [38]). It is always preferable to use a range of sites for SDL via high-throughput mutagenic approaches. A more detailed analysis on the reliability of the SDL approach is given in a previous review [33].

these SDL experiments were carried out in fully hydrated vesicles, which more closely mimic *in vivo* conditions. The SDL approach confirmed that the transmembrane helical domain tilts by ~20° with respect to the membrane normal, in accordance with the existing models. In addition, the SDL approaches showed that the transmembrane helix, which displays a distinct kink in oriented bilayers as shown by solid-state NMR [10], is in fact a continuous α-helix in fully hydrated bilayers.

However, the strongest evidence against the L-shaped model in fully hydrated vesicles came from further SDL studies, in which fluorescence energy transfer was used. This work showed that the L-shaped model was in poor agreement with the fluorescence data, in contrast to the I-shaped model that fit the data well [18]. Subsequent work, combining both SDL approaches, culminated in a model in which residues 10–46 form a continuous α-helix leaving a few unstructured residues in the N and C-terminal domain [17–21] (Figure 2e). This model posits that minor structural adaptations in the N terminus (i.e. unfolding by fraying of the end [22]), combined with an adaptation of the tilt angle, are sufficient for the coat protein to be stably incorporated into the lipid membrane. Indeed, the structure of the coat protein in the membrane, before membrane-bound phage assembly, is not dramatically different from the structure in the bacteriophage, thereby enabling a fast and efficient incorporation of the protein into the bacteriophage with a low-energy cost [20]. This conclusion lies in apparent contradiction with the existing L-shaped model that, at the time, dominated the literature.

The lipid model system dictates the M13 coat protein structure

A key observation that was made during the odyssey of the M13 coat protein is that this protein, and the related coat proteins from the phages fd and Pf1, are single membrane-spanning proteins. Such proteins have no internal stability based on segment–segment interactions. The absence of such interactions signifies that there is no tertiary structure to hold the protein together [23]. Therefore, the structure of the protein will be strongly affected by the environment into which the protein is inserted, for example micelles, vesicles or oriented membranes [18]. As a result, extreme caution is warranted because the external molecular forces on the protein could disturb its structure. This consideration is most important for the N-terminal domain of the protein that protrudes from the membrane [23].

In addition to determining the M13 coat protein structure, it is important to understand the process of its membrane assembly, that is, the way in which the protein is embedded and where it is located in the membrane. Apart from specific membrane-anchoring determinants (Table 1), membrane assembly is basically determined by the length of the hydrophobic helical segment of the protein with respect to the hydrophobic thickness of the membrane. Both of these parameters have been varied experimentally in studies of the M13 coat protein [16,17,19,21,24]. A crucial element of the protein that was initially overlooked is its ability to tilt its transmembrane domain to enable embedding in phospholipid bilayers. The tilt of the protein with respect to the membrane normally depends on the thickness of the membrane. For example, in thick phospholipid bilayers (20:1PC) the tilt angle is 19° and in thin bilayers (14:1PC) it is 33° [17]. A tilt of the protein provides an additional mechanism to regulate its membrane embedding and to adjust to hydrophobic mismatch situations (i.e. the mismatch in length between the hydrophobic protein domain and the hydrophobic bilayer thickness) [25]. By changing the membrane thickness, only the length of the unstructured N-terminal domain is affected: it is 7 amino acid residues long in thick phospholipid bilayers (22:1PC) and 14 residues in thin membranes (14:1PC) [19,21]. Alternatively, site-directed mutagenesis enables the modulation of the numeric balance of hydrophilic and hydrophobic amino acid residues in the N-terminal arm [14], thereby altering the length of the hydrophobic helical domain of the protein. This change affects the hydrophobic mismatch situation and the tilt angle. In addition, the relatively large thickness of the bilayer headgroup region at the membrane–water phase

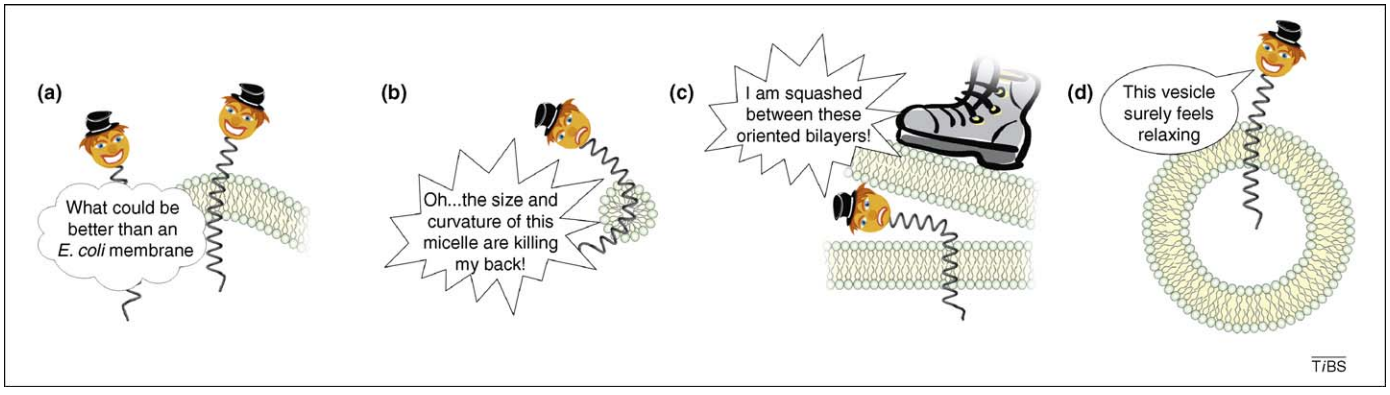

Figure 3. Cartoons illustrating the effect of the environment into which a single transmembrane protein is inserted. **(a)** Ideally, the protein should be placed in its natural *in vivo* environment; however, for most structural techniques this is not feasible. **(b)** Detergent micelles are very suitable for spectroscopic experiments, but because of their small size and strongly curved surface they do not provide a characteristic membrane bilayer environment. **(c)** Stacked bilayers are often used in solid-state NMR spectroscopy. However, between these stacks the space might be limited, not enabling full implantation of the termini of the protein, thereby forcing the protein in a bent shape. **(d)** Diluted vesicular membrane systems are closely related to the natural situation. Such systems can only be used in conjunction with high-sensitivity spectroscopic techniques for structure determination, such as SDFL (site-directed fluorescence labelling). As a result, extreme caution is warranted for determining the structure of a single transmembrane protein because the external molecular forces on the protein could disturb its structure.

was overlooked. In each bilayer leaflet, this region is half the thickness of the hydrocarbon chain region (hydrophobic core) of the membrane, which is 3.0 nm for 18:1PC phospholipid bilayers [26,27].

The available data support the idea that the structure of the M13 coat protein becomes strongly dependent on the employed lipid model system [18] (Figure 3). Ideally, the protein should be placed in its natural *in vivo* environment; however, for most structural techniques this is not feasible (Figure 3a). To utilize solution NMR the protein needs to be incorporated into detergent micelles that do not provide a characteristic bilayer environment (Figure 3b). Micelles have a curved surface, forcing the N-terminal domain of the protein to bend back on this surface, thus providing a variety of protein shapes including L- and U-shapes in addition to extended structures. This effect of micelle stress is also illustrated in studies of the HIV-1 Env peptide, which shows a micelle-induced curvature when solubilised in dihexanoyl phosphatidylcholine micelles [28].

The use of solid-state NMR requires incorporation of the protein into stacked bilayers. However, there is not sufficient space between these stacked bilayers to fully accommodate the N-terminal protein domain in an α-helical conformation (Figure 3c). Therefore, it must be parallel to the bilayer surface, forcing the protein into an L-shape.

Site-directed fluorescence labelling (SDFL) studies are performed at a low protein concentration in diluted membrane vesicle systems. These conditions provide a natural 'stress-free' environment for the protein, so that the resulting structure is obtained under natural membrane-embedding conditions (Figure 3d). Of course, one must be aware that mutagenesis, in combination with chemical modification by labelling, can have an effect on protein structure and membrane embedding.

Concluding remarks and future perspectives

After almost 20 years the problem surrounding the structure determination of membrane-bound M13 coat protein is 'straightened', that is, the protein is almost a straight α-helix. The protein survives the membrane-bound state by a simple tilt mechanism and a subtle structural adjustment in the extreme end of its N terminus. From an experimental point of view, it is essential to take into account the effects of environmental stress of membrane-mimicking environments on the structure of the protein. In this sense, SDL is a promising new biophysical methodology because it enables structure determination at biologically relevant conditions. However, although the membrane-bound coat protein is, on average, well described by a continuous α-helix, the data do not support the idea that the protein is perfectly stiff within the membrane. A small kink around residue 20 was identified in the helical protein model [20], which was interpreted to denote a maximal curvature of structure indicative of a smooth protein bend. In addition, a small deformation of the α-helical structure from residues 38–50 improved the performance of the helical model [21]. The ability of M13 coat protein to undergo helical deformations between residues 38–50 is thought to enable efficient incorporation into the phage particle when these residues are detached from the C-terminal membrane–water interface [21].

Single membrane-spanning proteins are delicate systems. There are no forces at work to stabilize the protein backbone by intramolecular interactions (as in multiple-spanning membrane proteins). Recent findings emphasize the need to study membrane proteins in a suitable environment, such as fully hydrated vesicles. As such, M13 coat protein studies provide an object lesson that mimicking the *in vivo* environment presents a challenge for *in vitro* biophysical and structural studies. In this sense, any membrane protein with water-exposed portions in the C or N termini and hydrophilic loop regions should be treated with great care.

X-ray crystallography and NMR spectroscopy are the primary techniques of choice for high-resolution structure determination of proteins at an atomic level. However, when studying membrane proteins, the protein must be incorporated in a membrane-like environment, which can create special conditions that could affect its structure. Therefore, these techniques should be complemented with 'low-resolution data', for example obtained via SDL techniques. In contrast to isotope labelling studies that are used in NMR spectroscopy, which are essentially non-disturbing, SDL can introduce a protein modification when molecular probes are covalently added; thus, this approach can bring about unwanted effects on the structure and function of the protein. However, SDL enables the use of a wide variety of probes, including environmental probes that enable monitoring further away from the protein backbone, structural probes for monitoring near the backbone, and fluorescent and spin probes. New labelling strategies based on expanding the genetic code are being developed that will enable a position-specific incorporation of 'un-natural' labelled amino acids [29–32]. Such an approach will permit the specific insertion of small tailor-made probes and will offer the maximum flexibility in the SDL methodology. In addition, NMR spectroscopy and emerging site-specific infrared dichroism (SSID) techniques [33,34], which require the specific incorporation of isotopes, will benefit from these new protein engineering methods.

References

1 Spruijt, R.B. *et al.* (1999) M13 phage. In *The Encyclopedia of Molecular Biology* (Creighton, T., ed.), pp. 1425–1429, New York, John Wiley

2 Glucksman, M.J. *et al.* (1992) Three-dimensional structure of a cloning vector: X-ray diffraction studies of filamentous bacteriophage M13 at 7 Å resolution. *J. Mol. Biol.* 226, 455–470

3 Marvin, D.A. *et al.* (1994) Molecular models and structural comparisons of native and mutant class-I filamentous bacteriophages Ff (fd, f1, M13), If1 and IKe. *J. Mol. Biol.* 235, 260–286

4 Wang, Y.A. *et al.* (2006) The structure of a filamentous bacteriophage. *J. Mol. Biol.* 361, 209–215

5 Zeri, A.C. *et al.* (2003) Structure of the coat protein in fd filamentous bacteriophage particles determined by solid-state NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 100, 6458–6463

6 Marvin, D.A. (1998) Filamentous phage structure, infection and assembly. *Curr. Opin. Struct. Biol.* 8, 150–158

7 Shon, K-J. *et al.* (1991) NMR studies of the structure and dynamics of membrane-bound bacteriophage Pf1 coat protein. *Science* 252, 1303–1305

8 McDonnell, P.A. *et al.* (1993) fd Coat protein structure in membrane environments. *J. Mol. Biol.* 233, 447–463

9 Glaubitz, C. *et al.* (2000) Structural and orientational information of the membrane embedded M13 coat protein by ^{13}C-MAS NMR spectroscopy. *Biochim. Biophys. Acta* 1463, 151–161

10 Marassi, F.M. and Opella, S.J. (2003) Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints. *Protein Sci.* 12, 403–411
11 Opella, S.J. *et al.* (2008) Structure, dynamics, and assembly of filamentous bacteriophages by nuclear magnetic resonance spectroscopy. *Annu. Rev. Phys. Chem.* 59, 635–657
12 Almeida, F.C.L. and Opella, S.J. (1997) Fd coat protein structure in membrane environments: structural dynamics of the loop between the hydrophobic trans-membrane helix and the amphipathic in-plane helix. *J. Mol. Biol.* 270, 481–495
13 Papavoine, C.H.M. *et al.* (1998) Solution structure of the M13 major coat protein in detergent micelles: a basis for a model of phage assembly involving specific residues. *J. Mol. Biol.* 282, 401–419
14 Spruijt, R.B. *et al.* (2000) Localization and rearrangement modulation of the N-terminal arm of the membrane-bound major coat protein of bacteriophage M13. *Biochim. Biophys. Acta* 1509, 311–323
15 Meijer, A.B. *et al.* (2001) Configurations of the N-terminal amphipathic domain of the membrane-bound M13 major coat protein. *Biochemistry* 40, 5081–5086
16 Spruijt, R.B. *et al.* (2004) Membrane assembly of M13 major coat protein: evidence for structural adaptation in the hinge region and a tilted transmembrane domain. *Biochemistry* 43, 13972–13980
17 Koehorst, R.B.M. *et al.* (2004) Lipid bilayer topology of the transmembrane α-helix of M13 major coat protein and bilayer polarity profile by site-directed fluorescence spectroscopy. *Biophys. J.* 87, 1445–1455
18 Vos, W.L. *et al.* (2005) Membrane-bound conformation of M13 major coat protein: a structure validations through FRET-derived constrains. *J. Biol. Chem.* 280, 38522–38527
19 Stopar, D. *et al.* (2006) Motional restrictions of membrane proteins: a site-directed spin labeling study. *Biophys. J.* 91, 3341–3348
20 Nazarov, P.V. *et al.* (2007) FRET study of membrane proteins: determination of the tilt and orientation of the N-terminal domain of M13 major coat protein. *Biophys. J.* 92, 1296–1305
21 Vos, W.L. *et al.* (2007) Structure of membrane-embedded M13 major coat protein is insensitive to hydrophobic stress. *Biophys. J.* 93, 3541–3547
22 Klimov, D.K. *et al.* (1998) Virtual atom representation of hydrogen bonds in minimal off-lattice models of alpha helices: effect on stability, cooperativity and kinetics. *Fold. Des.* 3, 481–496
23 Stopar, D. *et al.* (2006) Anchoring mechanisms of membrane-associated M13 major coat protein. *Chem. Phys. Lipids* 141, 83–93
24 Meijer, A.B. *et al.* (2001) Membrane-anchoring interactions of M13 major coat protein. *Biochemistry* 40, 8815–8820
25 Killian, J.A. (1998) Hydrophobic mismatch between proteins and lipids in membranes. *Biochim. Biophys. Acta* 1376, 401–415
26 White, S.H. and Wimley, W.C. (1999) Membrane protein folding and stability: physical principles. *Annu. Rev. Biophys. Biomol. Struct.* 28, 319–365
27 Ridder, A.N.J.A. *et al.* (2002) Importance of hydrophobic matching for spontaneous insertion of a single-spanning membrane protein. *Biochemistry* 41, 4946–4952
28 Chou, J.J. *et al.* (2001) Micelle-induced curvature in a water-insoluble HIV-1 Env peptide revealed by NMR dipolar coupling measurement in stretched polyacrylamide gel. *J. Am. Chem. Soc.* 124, 2450–2451
29 Becker, C.F.W. *et al.* (2005) Incorporation of spin-labelled amino acids into proteins. *Magn. Reson. Chem.* 43, S34–S39
30 Dougherty, D.A. (2000) Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* 4, 645–652
31 Hendrickson, T.L. *et al.* (2004) Incorporation of nonnatural amino acids into proteins. *Annu. Rev. Biochem.* 73, 147–176
32 Wang, L. *et al.* (2006) Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct.* 35, 225–249
33 Torres, J. *et al.* (2003) Membrane proteins: the 'Wild West' of structural biology. *Trends Biochem. Sci.* 28, 137–144
34 Kukol, A. (2005) Site-specific IR spectroscopy and molecular modelling combined towards solving transmembrane protein structure. *Spectroscopy* 19, 1–16
35 Strandberg, E. and Killian, J.A. (2003) Snorkeling of lysine side chains in transmembrane helices: how easy can it get? *FEBS Lett.* 544, 69–73
36 Egelman, E.H. (2007) Single-particle reconstruction from EM images of helical filaments. *Curr. Opin. Struct. Biol.* 17, 556–561
37 Sammalkorpi, M. and Lazaridis, T. (2007) Modeling a spin-labeled fusion peptide in a membrane: implications for the interpretation of EPR experiments. *Biophys. J.* 92, 10–22
38 Altenbach, C. *et al.* (2008) High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl. Acad. Sci. U. S. A.* 105, 7439–7444