# MOLECULAR MODELING OF SELECTED TRANSMEMBRANE PROTEINS

**Ph.D. Thesis**

**by**

**DENYS BASHTOVYY**

**Institute of Biophysics**
**Biological Research Center**
**Hungarian Academy of Sciences**

**Szeged, 2004**

**Contents**

# Introduction

## *Transmembrane Proteins*

Transmembrane (TM) proteins share a common property: part of their structure is embedded in a lipid bilayer. Therefore, being located at an interface, it is almost inevitable that they mediate communication between both sides of the membrane; receptors, pores and channels are all signal transducers. In their lipid-embedded domain, only two types of secondary structure have been observed – β-strands and α-helices. β-strands are found in outer membranes of Gram-negative bacteria, mitochondria and chloroplasts, forming rigid pores known as β-barrels. Single and bundled transmembrane α-helices have a broader range of functionalities and complexities. In some instances, extensive extramembrane domains complement the TM ones, and some membrane proteins consist of huge multisubunit complexes (e.g. cytochrome c oxidase, which contains up to 13 subunits in mammals).

Integral membrane proteins represent an important class of proteins which are employed in a wide range of cellular roles. The fact that these proteins are found in a lipid environment means that atomic resolution experimental structures for these proteins are few. To date only about 50 unique high-resolution integral membrane protein structures have been solved, whereas several dozens of thousands structures for globular proteins are known (Curran and Engelman, 2003).

Fortunately, despite the difficulty in experimental structure determination for transmembrane proteins (the classical structural methods, namely X-ray and NMR techniques, are subject to well-known limitations when applied to proteins in the membrane-bound state), the physicochemical constraints imposed by the lipid environment make the prediction of transmembrane protein structure somewhat more straightforward than for globular proteins. Algorithms for predicting trans-bilayer protein topology are currently very accurate (97% for a given helix). The best predictors are still based on neural-network (Krogh A et al., 2001) and hidden Markov model (Rost, 1995) methods. These algorithms take advantage of evolutionary data from multiple sequence alignments and consider sequence elements in a way that is context dependent.

Such extra features use more information than the simple hydrophobicity of TM residues (which alone can predict 80% of TM helices) to refine TM identification. For example, observation of the positive-inside rule for proteins allows the identification of otherwise ‘missing’ helices by checking for charge inversions at helix termini.

As the number of solved structures of TM proteins continues to grow, a dataset is developing that can serve to test the ability of sequence-based algorithms to predict TM-helix orientation and packing. Attempts have been made to correlate the surfaces of TM helices with their relative orientation (Pilpel Y et al., 1999). However, very subtle conformational complementarities appear to be responsible for most of the TM α-helix association in TM helical bundles. Thus, in contrast to soluble helical bundles, in most cases there is no clear helix amphipathicity (i.e. more hydrophobic residues facing the lipid). In addition, the sparse hydrophilic interactions in the TM domain are not a generalized driving force for helical association.

Residues that are structurally and functionally important can be detected using mutagenesis or sequence-based estimates of amino acid substitution (Stevens and Arkin, 2001), and general packing arrangements can be determined by observing the interface propensity (preferences of individual residues to be found at a given interface (Pilpel Y et al., 1999)) and hydrophobic moment (a measure of the direction of the amphipathic nature of an α-helix) (Eisenberg, 1984); these approaches are particularly effective when used in combination. However, the resolution of such analyses remains low. Despite being restricted to a lipid bilayer, it is not clear whether predicting the packing of TM helices is a conceptually easier task than predicting that of the hydrophobic cores of aqueous proteins, because the surfaces of TM helices are chemically homogeneous compared with those of aqueous bundles.

In the presented work we have focused on the structures of two distinct membrane proteins, the major coat protein of the bacteriophage M13 (M13 MCP) and the representative(s) of the cytochrome $b_{561}$ redox protein family. Our group dedicated substantial effort to experimental studies of these proteins in a membrane-bound form. However, for none of these proteins was an atomic structure of the

native, membrane-bound form available. Our task was, therefore, to build and analyze molecular models of these proteins on the basis of the relevant experimentally and *in silico* obtained information, in order to stimulate and design new experiments and also to verify available data.

### *The Major Coat Protein of the M13 Bacteriophage (M13 MCP)*

Filamentous bacteriophages (Ff), including fd, f1 and M13, are nonlytic, male-specific bacteriophages which infect *Escherichia coli* cells carrying an F-episome. The circular single-stranded genome encodes 10 proteins (for a review, see (Marvin, 1998)). Major coat protein pVIII is by far the most abundant protein of filamentous phage, present in about 2,700 copies per phage. It forms the cylinder architecture of virions with its C-terminus buried in the interior, whereas its N-terminus is exposed to the medium. Its number is correlated with the phage genome length. pVIII is synthesized as a precoat protein containing a 23-amino acid leader sequence which is cleaved to yield a 50-residue mature transmembrane protein (Sugimoto et al., 1997).

Three distinctive domains have been identified in pVIII: an acidic amphipathic N-terminal domain, which is exposed to the outer surface of the virion, a hydrophobic transmembrane domain and a positively charged C-terminus which is in close proximity to the viral DNA (Marvin et al., 1994;Almeida and Opella, 1997). The secondary structure of pVIII in phage particle is presumed to form an α-helical structure, with only the first 4–5 residues being mobile and unstructured (Kay et al., 1993;Marvin et al., 1994;Sugimoto et al., 1997). Upon reconstitution in the membrane the secondary structure of the major coat protein is thought to change dramatically.

Unfortunately a complete and reliable atomic structure of the coat protein in the phospholipid membranes cannot be obtained by solution NMR or X-ray crystallography. Structural rearrangements were studied, however, indirectly by applying different spectroscopic techniques (Hemminga et al., 1992;Wolkers et al., 1995;Wolkers et al., 1997;Blanch et al., 2002). Attempts were made to partially determine the structure of the major coat protein either in the phage particle or when

reconstituted in a model membrane environment (for recent review see, e.g. (Stopar et al., 2003)). Before further progress with structural studies on the M13 MCP the question was raised: are the high-resolution detergent-based nuclear magnetic resonance (NMR) structures (Almeida and Opella, 1997;Papavoine et al., 1998) compatible with the low resolution site-directed spin label electron paramagnetic resonance (EPR) data on the membrane-bound form (Stopar et al., 1997a)? We aimed to answer the question with molecular modeling.

***Cytochromes $b_{561}$***

The high-redox-potential *b*-type cytochrome (cytochrome $b_{561}$) of chromaffin granule membranes of the mammalian adrenal medulla can be fully reduced by ascorbate (Asc).The wavelength of its characteristic alpha-band absorbance maximum in the reduced-minus-oxidized absorbance spectra is close to 561 nm. The protein is capable of transporting electrons through the chromaffin granule membrane (Njus et al., 1983;Srivastava et al., 1984;Kent and Fleming, 1987;Kelley et al., 1990). In the past decades, evidence accumulated for the presence of a similar Asc-reducible cytochrome $b_{561}$ in plant plasma membranes (for a recent review, see (Asard et al., 2001)).This protein is also able to transfer electrons across the membrane (Asard et al., 1992) in a way that may be similar to that of the chromaffin granule membrane.

Genes coding for proteins with significant homology to the mammalian cytochrome $b_{561}$ have recently been identified in a large number of plant species (Asard et al., 2000). The mammalian and predicted plant cytochrome $b_{561}$ proteins are highly hydrophobic and transport electrons from the cytoplasmic side of the membrane in which they are embedded to the extracellular space or into intracellular vesicles. The physiological function of the plant plasma membrane cytochrome $b_{561}$ is yet to be elucidated. The first evidence for the existence of the cytochrome $b_{561}$ protein family in plants and animals was presented on the basis of a sequence analysis of 9 related sequences from different eukaryotic species (Asard et al., 2000).

It is generally believed that these redox proteins play an important role in a wide variety of physiological processes, including iron uptake, cell defense, nitrate reduction, and signal transduction. Recently a new member of this protein family has been located in the duodenal cells of the small intestine and demonstrated to play a role in the reduction of iron prior to its uptake (McKie et al., 2001). It was suggested that a cytochrome $b_{561}$ in these cells possesses ferric reductase activity.A similar function was proposed for cytochrome $b_{561}$ -like domains in larger proteins which play a potential role in neurodegenerative disorders (Ponting, 2001). This activity apparently contrasts with the activity of the chromaffin granule and plant plasma membrane cytochromes $b_{561}$, which are likely to function as monodehydroascorbate (MDA) reductases.

These proteins have been demonstrated to receive an electron from cytoplasmic Asc and transfer it across the membrane to MDA (Harnadek et al., 1992). Details of this process, in particular the transmembrane electron transfer mechanism, are not yet resolved, but almost certainly the two heme centers are involved (Tsubaki et al., 1997;Kobayashi et al., 1998;Trost et al., 2000). These proteins therefore represent an important and unique family of transmembrane electron transport proteins because of their putative and/or yet to be identified physiological functions in a variety of eukaryotic cells. In addition, the 1 eq reaction between Asc and the proteins and the long distance between the two hemes, which almost spans the membrane interior, make them a potentially very interesting model for redox reactions between metalloproteins and organic substrates and also for transmembrane electron transfer.

Since cytochrome $b_{561}$ proteins have not yet been crystallized, atomic-detail structural data about these proteins were lacking. On the other hand, essential structural features, including conserved heme-binding residues, transmembrane (TM) helices, and potential substrate binding sites have been identified (Okuyama et al., 1998;Asard et al., 2001;Takeuchi et al., 2001). Moreover, success in the purification of the plant plasma membrane cytochrome $b_{561}$ (Trost et al. 2000; Berczi et al. 2001, 2003) has been promising to render biophysical studies on plant proteins possible in the near future. Therefore, a working model of atomic detail representing the main

structural features of the cytochrome $b_{561}$ protein family would have been highly useful. The objectives of the present work were to identify new structural similarities in the cytochrome $b_{561}$ family and to build 3-dimensional (3-D) atomic models for representative plant and mammalian cytochrome $b_{561}$ proteins.

**Aims**

In the presented work various molecular modeling techniques (quantum chemical calculations, molecular mechanics optimizations, conformational searches, homology modeling; sequence database searches and alignments, transmembrane and lipid-facing propensities prediction etc) were applied to model and study selected transmembrane proteins, namely the Major Coat Protein of the M13 Bacteriophage M13 and 2 representatives (plant and mammalian) from the cytochrome $b_{561}$ family. The aims of these studies were as follows:

M13 MCP

- To test the bundle of distinct M13 MCP structures, refined in detergent micelles, against experimental constraints obtained from the protein embedded in a phospholipid bilayer in order to identify those structures that are most compatible with a lipid membrane environment;
- To test and refine earlier proposed indicator of the local packing density (the *f*-parameter), which is readily calculated from the coordinates of the optimized protein–lipid structural model;
- To investigate structural reasons for the increased outer hyperfine splitting ($2A_{max}$) values, observed in earlier EPR experiments from the spin-labeled residues 25 and 36;

Cytochrome $b_{561}$ family

- To perform comprehensive sequence analysis on the representatives of the family, predict transmembrane and lipid-facing propensities of the TM helices;
- To identify new structural similarities in the cytochrome $b_{561}$ family and to build 3-dimensional atomic models for representative plant and mammalian cytochrome $b_{561}$ proteins.

Detailed molecular models should aid the understanding of the available experimental data on these proteins and in the design of new experiments.

## Materials and Methods

### *M13 MCP structures*

Three-dimensional structures of the M13 MCP were taken from the Brookhaven Protein Data Bank in PDB format (Berman et al., 2000). The PDB IDs were 2CPB and 2CPS for structures determined in dodecylphosphocholine (DodPC) and sodium dodecyl sulfate (SDS) micelles, respectively, by various high-resolution NMR techniques (Papavoine et al., 1998).

### *Experimental data*

EPR outer hyperfine splittings (given in parenthesis after the each mutant name) and membrane topology data for the viable single cysteine mutants A25C (6.36 mT) , V31C (5.87 mT), T36C (5.92 mT), G38C (5.81 mT) and T46C (5.79 mT) of the M13 major coat protein reconstituted in 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) bilayers were taken from (Stopar et al., 1997a).

### *Molecular modeling of M13 MCP*

Quantum chemistry and molecular mechanics package Spartan v.5.0.4 (Wavefunction Inc., Irvine, CA) with MMFF94 force field (Halgren and Nachbar, 1996;Halgren, 1996a;Halgren, 1996b;Halgren, 1996c;Halgren, 1996d), and the interactive molecular mechanics package Sculpt v.2.1 (Surles et al., 1994) (Interactive Simulations Inc., San Diego, CA), were used for building and optimization of structures. Specifically, Spartan was used to generate the spin-labeled cysteine residue (validated by semi-empirical quantum chemical methods) and the phospholipid structure, and for reoptimization of the protein structure after single-residue replacement by spin-labeled cysteine. Additionally, Spartan was used for single-point energy calculations to obtain atomic charges. Adjustment of the phospholipid chain configuration and constrained molecular mechanics optimization of the protein–lipid assemblies were performed in Sculpt. MOLMOL (Koradi et al., 1996) was used for producing single-residue substitutions, construction of the lipid shell, and preparing the system for optimization by molecular mechanics. Insight II (Molecular Simulations Inc., San

Diego, CA) was used for visualization and presentation of structures. All modeling work was performed on a Silicon Graphics (Mountain View, CA) Origin 2000 server and O2 workstations.

***Molecular modeling of cytochromes $b_{561}$***

Related sequences were identified from 26 different tissues and organisms via PSI-BLAST database searches with default settings (Altschul S.F. et al., 1997) using Artb561-1, Artb561-4 and Hosb561-1 as queries (sequence names are used as defined in (Asard et al., 2001)). The sequences were aligned using MULTICLUSTAL (Yuan et al., 1999). Lipid facing propensities of the predicted TM regions have been analyzed by means of kPROT (Pilpel Y et al., 1999). 3-dimensional structures were built using Biopolymer and Homology modules in InsightII (Accelrys, 2000).

## Results

### *M13 MCP*

#### Starting structures

M13 MCP structures in SDS micelles, determined by NMR and optimization techniques (Papavoine et al., 1998) were subject to further experimental constraints in the membrane-bound state, which were obtained from viable cysteine mutants in spin-label EPR experiments (Stopar et al., 1997a). To do this, we built a spin-labeled cysteine, made single-residue replacements in the M13 MCP structures with cysteine-maleimide, optimized the modified structures, and then derived structural parameters from the models for comparison with the experimental EPR data. Some of the structures that are possible in small, highly curved micelles can be eliminated immediately as suitable candidates for the structure in planar membranes, because the charged N-terminal section would bend back into the hydrophobic interior of the membrane. These are the U-shaped structures that constitute not more than 7–8 members in each set (see Table 1).

In addition, solid-state NMR studies on aligned phospholipid bilayers have indicated that the N-terminal helix (specifically the section containing L14) of the closely related fd bacteriophage coat protein is oriented nearly perpendicular to the transmembrane helix (McDonnell et al., 1993). This further justifies elimination of the U-shaped structures.

#### Spin label structure

The 3-maleimido-2,2,5,5-tetramethylpyrrolidine-*N*-oxyl (5-MSL) spin label was built and minimized as the two stereoisomers arising from the chiral carbon in the proxyl ring. After connecting 5-MSL to the sulfhydryl group of the cysteine residue, four stereoisomers were obtained. The eight structures obtained by rotating the proxyl ring around the N–C bond by 180° for each stereoisomer were optimized using

MMFF94 (Halgren and Nachbar, 1996;Halgren, 1996a;Halgren, 1996b;Halgren, 1996c;Halgren, 1996d). One structure was chosen (Fig. 2) on the basis of lowest energy. Only two other of the alternative structures are expected to be comparably populated at room temperature. The remainder were predicted to lie higher in energy by 1.3–3.6 kcal/mol. In the chosen structure, the long axis of the maleimide spin label is oriented approximately perpendicular to the helix axis, in agreement with previous suggestions (Wolkers et al., 1997).

**Table 1.** ***Angle between transmembrane and N-terminal helices in NMR-derived structures of M13 major coat protein in DodPC (PDB: 2CPB) and SDS (PDB: 2CPS) micelles determined by Papavoine et al. (1998)[a]***

| Structure no. | DodPC | SDS |
|---|---|---|
| 1 | 146° | 100° |
| 2 | 117° | 129° |
| 3 | 99° | 139° |
| 4 | 69° | 30° |
| 5 | 123° | 111° |
| 6 | 70° | 90° |
| 7 | 117° | 122° |
| 8 | 59° | 17° |
| 9 | 64° | 63°[b] |
| 10 | 80° | 69°[b] |
| 11 | 35° | 108° |
| 12 | 52° | 146° |
| 13 | 150° | 113° |
| 14 | 121° | 111° |
| 15 | 132° | 61°[b] |
| 16 | 153° | 69° |
| 17 | 100° | 85° |
| 18 | 50° | 147° |
| 19 | 18° | 154° |
| 20 | 96° | 81° |
| 21 | 65°[b] | 56° |
| 22 | 59° | 51° |
| 23 | 150° | 93° |
| 24 | 92° | 125° |
| 25 | 44° | 91° |

[a] Structures with an angle between helix axes of <60° were discarded in the model building.

[b] These structures were also discarded because the non-helical part of the N-terminal forms a U-shape.

Spin-labeled M13 MCP without lipids

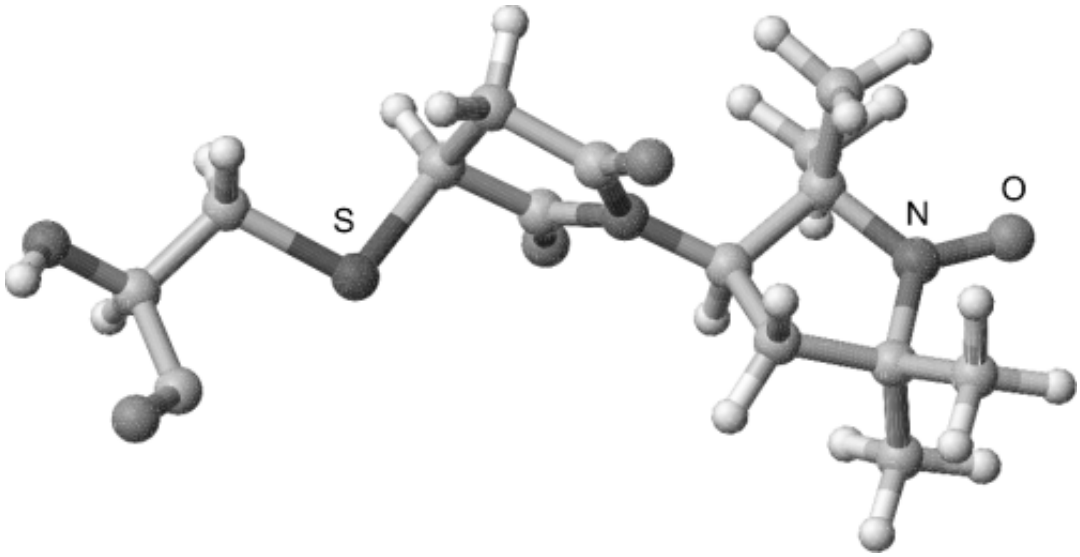


**Figure 1.** ***Optimized structure of spin-labeled cysteine. The 5-maleimidoproxyl spin label that is covalently attached at the sulfhydryl group (S) is indicated together with the nitroxide group (N–O) of the proxyl ring. The structure was built and optimized in Spartan with the MMFF94 force field, and is displayed using MOLMOL (Koradi et al., 1996).***

The spin-labeled cysteine (Fig. 1) was substituted in the original (from SDS) NMR structures and in the corresponding ones from DodPC, at positions corresponding to the single cysteine mutations (A25C, V31C, T36C, G38C, T46C, and A49C) used in the spin-label EPR studies (Stopar et al., 1997a). The residue replacements were performed in MOLMOL and are indicated, on a single peptide backbone, in Figure 2. Excluding U-shaped conformations, this generates a total of ~35 possible structures for each mutant. For mutation positions definitely within the hydrophobic region (residues 31, 36, and 38), the number of independent structures was reduced to minimally 26 by grouping together equivalent structures. These are representative, on the basis of root-mean-square (rms) differences lower than 0.25 Å, of the local structures ±2 residues from the mutated residue (it was verified that the structures assigned as equivalent are not differentially influenced by the presence of solvating phospholipids). Because structures at the ends of the hydrophobic region are not well aligned within a family, all 35 possibilities were treated explicitly for

mutated residues 25, 46, and 49, which are expected to be located close to the lipid headgroup regions. The protein-attached spin label structures were then reoptimized using the MMFF94 force field (Halgren and Nachbar, 1996;Halgren, 1996a;Halgren, 1996b;Halgren, 1996c;Halgren, 1996d). The resulting trial structures could then be compared with the EPR results on the different spin-labeled mutant proteins reconstituted in phospholipid membranes.

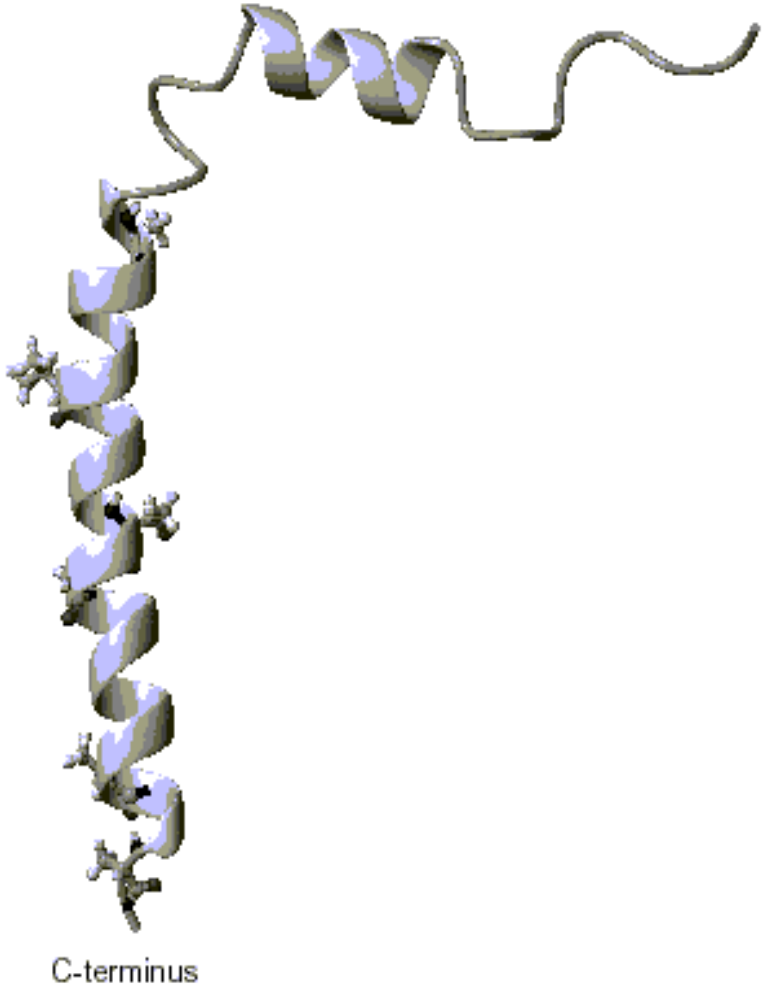


**Figure 2.** ***Ribbon representation of the M13 major coat protein structure no.20 of entry 2CPS from the Brookhaven Protein Data Bank (Papavoine et al., 1998). Side chains of the original residues that were changed singly to cysteine (top to bottom: A25C, V31C, T36C, G38C, T46C, and A49C), both by mutagenesis (Stopar et al. 1997) and by modeling in the present study, are shown in ball-and-stick representation. The figure was created using MOLMOL (Koradi et al., 1996).***

The relevant experimental data are not only measurements of the vertical position of the spin-labeled residue in the membrane, but importantly are also the profile of local mobility of the spin label with position in the sequence (Stopar et al., 1997a). The latter is parameterized by the outer hyperfine splitting, $2A_{max}$, in the spin-label EPR spectrum (see, e.g., (Marsh, 1981)). The experimental profile is characterized by a remarkably high value of $2A_{max}$ at C25, which decreases nonmonotonically on proceeding further along the sequence. The local rotational mobility of the spin label will depend on the atomic packing density in the region of attachment of the spin label. We have chosen to characterize this packing density by a parameter (*f*) that is defined by

$$f = \sqrt{\sum_i \frac{m_i}{d_i^2}} \qquad (1)$$

where $m_i$ is the mass and $d_i$ is the distance (in Å) of the *i*th atom measured from the reference atom. The reference position was the nitrogen in the proxyl ring of 5-MSL. All atoms of the protein, including hydrogens, were used in the summation. Only the atoms of the spin label were excluded from calculation of the *f*-parameter. The *f*-parameter in Equation 1 differs slightly from a previous one that was used successfully in a similar site-directed spin-labeling study on cytochrome *c* in solution (Turyna et al., 1998). (Note that the latter reference contains a printing error.) The extensions made here are to include mass weighting and to retain hydrogen atoms in the summation. Note that it is only relative values of $A_{max}$ (i.e., the shape of the profile) with which we seek to correlate the *f*-parameter. For this, a qualitative correlation should suffice. However, the final fit (see Fig. 7, below) does imply an approximately linear relation, most probably because the spectra analyzed are confined to the same motional regime of spin-label dynamics (see, e.g., Marsh and Horvath 1989).

The dependence of the *f*-parameter on the range over which the summation is made is shown in Figure 3. Data are given for three different mutants, and the effect of including or excluding hydrogen atoms is also shown. A 9-Å range was finally chosen for the summation, after testing various summations ranging from 4 to 50 Å, because this was found to give near-optimal discrimination between the different

mutants (see Fig. 3). The 9-Å radial region is large enough to include all atoms that can affect the spin label directly and is small enough to exclude atoms that have no direct influence on the spin label.

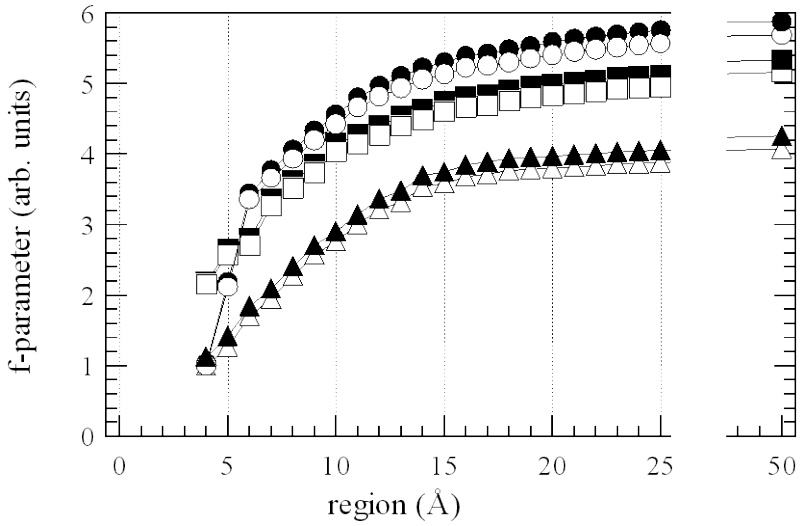


**Figure 3.** ***Dependence of the packing parameter (f, defined in Eq. 1) on the radial region over which the atom summation is made. Results of calculations are given for spin-labeled cysteine mutants: A25C (circles), V31C (squares), and A49C (triangles), either with (solid symbols) or without (open symbols) hydrogens included in the evaluation. The radius of the region was measured from the reference atom, which was the nitrogen atom of the proxyl ring of the spin label.***

The profiles of the calculated *f*-parameters are shown in Figure 4 for the two reduced sets of structures (i.e., from the families of structures in DodPC and SDS). Relative to the experimental EPR profile, the most significant feature of Figure 5 is that only structures in which the C25 residue is inside the interhelix hinge region between the N-terminal and transmembrane segments (cf. Fig. 2) give rise to a significantly higher *f*-parameter than do the other mutant positions. This substantiates previous suggestions from EPR and molecular modeling studies on the hinge region (Wolkers et al., 1997). The structures used in Figure 4 are still without the lipid shell, but it was found subsequently that the large *f*-parameter of C25, relative to other positions, could not be produced simply by adding lipids (data not shown). Therefore, this finding limits suitable candidates to those structures that are L-shaped

(see Table 1) and for which C25 is inside the hinge region. Surprisingly, no such

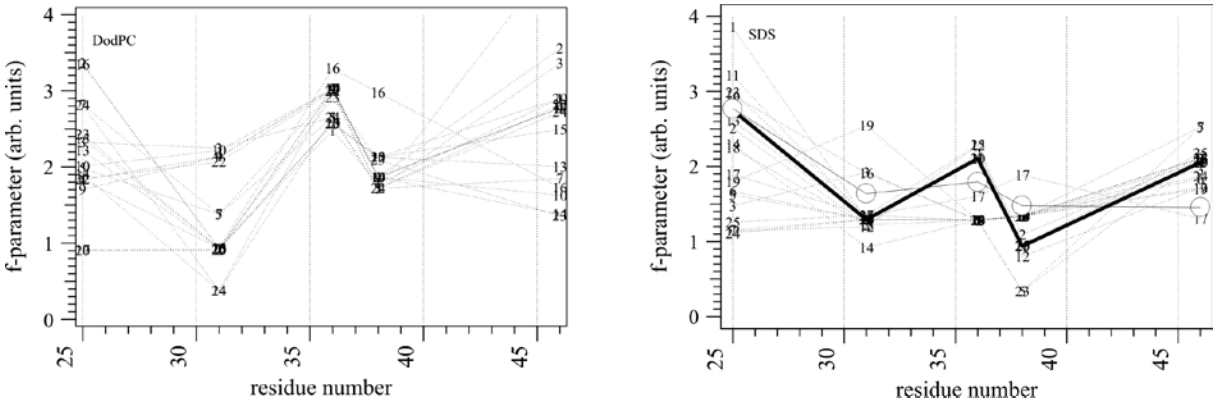


**Figure 4.** ***Sequence profiles for the f-parameter of M13 coat protein structures with single cysteine–maleimide replacements, in the absence of a lipid bilayer. Each x-coordinate represents a single mutant of the original families of structures determined in DodPC (left) or in SDS (right), in which the cysteine–maleimide side chain was locally optimized after amino acid replacement. Structure numbers indicated in the figure are those in the PDB files 2CPB and 2CPS for NMR-based structures in DodPC and SDS micelles, respectively, determined by (Papavoine et al., 1998). (Right) The heavy solid line indicates the selected structure no. 20; open circles are the experimental EPR outer hyperfine splittings scaled to best match this profile.***

structure was found in the DodPC family (i.e., PDB: 2CPB; Fig. 5, left). Out of the small subset (3–4) of very similar structures found to fulfil this condition in the SDS family (i.e., PDB: 2CPS), structure no. 20 was selected as representative for further model building because the *f*-parameter profile (see Fig. 4, right) most resembled the experimental mobility profile. In this subset of structures, the large *f*-parameter (and outer hyperfine splitting) for residue 25 is caused by the cysteine–maleimide being squeezed between W26 and I22 side chains. Additionally, putative hydrogen bonds between W26 and maleimide oxygens possibly contribute to immobilization of the spin label. Structure no. 20 from the SDS family was a favourable choice also because the C-terminal half of the hinge region (residues 17–24) of the protein is better defined in SDS micelles than in DodPC micelles (Papavoine et al., 1998).

## Spin-labeled protein with lipids

A single shell (the first solvation bilayer shell) of DOPC lipids was constructed around the selected M13 major coat protein structure. First, a single

DOPC molecule was built within Spartan. The glycerol region was created according to the *sc*/γ structure given by Pascher (1996). The electrostatic charges for 5-MSL and DOPC were calculated using density functional theory at the pBP86/DN** level. The acyl chains of the DOPC molecule were shortened using the tether tool in Sculpt (Surles et al., 1994) to match the apolar region of the protein. A lipid shell was then constructed in MOLMOL using this DOPC structure, which resulted in a bilayer thickness of ~33 Å between phosphorus atoms of opposing molecules. This is close to the length of the intramembranous part of the protein, which is 31.8 Å between the α carbons of residues 25 and 46 of structure no. 20 from PDB entry 2CPS. For comparison, steric bilayer thickness $D_B'$ (choline-choline distance) in a fully hydrated DOPC bilayer is ≈35.9 Å, and the hydrocarbon thickness $2 D_C$ is 27.1 Å (Nagle and Tristram-Nagle, 2000).

After a test minimization of the M13 MCP surrounded by 18 DOPC molecules in Sculpt, the number of lipid molecules was reduced to 12 (one shell of six lipids for each bilayer half). The diameter of this lipid shell was ~19 Å. Each of the 6 mutants of structure no. 20 was inserted in the center of the lipid shell in MOLMOL, with vertical positioning of the hydrophobic helix according to the distance measurements of Stopar et al. (Stopar et al., 1997a). This composite structure was minimized in Sculpt with only the protein backbone frozen. Additional forces (called springs in Sculpt) were applied to every atom of the lipid molecules, in a direction towards the axis of the protein hydrophobic helix. The default force field was used, Van der Waals interactions were modeled with a modified Lennard-Jones potential between atoms within 6 Å of each other, electrostatic interactions were treated with a Coulomb model, using a distance-dependent dielectric constant between atoms within 10 Å of each other (Surles et al., 1994). These structures were optimized until the fractional change in energy was <0.01. The resulting optimized model is shown for the spin-labeled A25C mutant in Figure 5.

Next, *f*-parameters were calculated for the minimized protein–lipid structures. In general, the lipid shell increased the *f*-parameter for residues 25, 31, 36, and 38 rather uniformly, whereas it had a discriminative effect at residues 46 and 49. For all

mutants, the spin label was directed pointing away from the helix in the optimized structures (see Table 2).

**Table 2.** ***Orientation of the spin-label N-O bond to the axis of the TM helix, for structure number 20 of the M13 coat protein (PDB: 2CPS) optimized with single cysteine-maleimide substitutions at different positions, in the absence and presence of the lipid shell***

| Spin-labeled residue no. | Without lipid | With lipid[a] |
|---|---|---|
| 25 | 67° | 51° |
| 31 | 104° | 111° |
| 36 | 107° | 95° |
| 38 | 100° | 97° |
| 46 | 77° | 40°[b] |
| 49 | 70° | 70° |

Exceptionally, this orientation produced an inconsistently high *f*-parameter for the maleimide connected to C46, which is situated close to the lipid headgroups. However, for this particular residue, an orientation pointing outside the lipid shell (along the protein helix) produced a negligible change in energy. Therefore, the label on C46 was fixed manually in the latter orientation.

Sequence profiles were then calculated for the *f*-parameter, not only with the protein in the position used for the structural optimization, but also with the protein displaced vertically in 1-Å increments, without further optimization. The resulting *f*-parameter profiles were then fitted to that of the EPR outer hyperfine splittings by using an adjustable linear scaling factor, plus offset, for the latter. The dependence of the fitting errors on the vertical displacement of the protein is given in Figure 6.

The minimum fitting error was not found for zero vertical displacement of the protein, but for a displacement by +2 Å. Therefore, optimization of the entire protein–lipid structure was repeated with different vertical displacements of the protein. A consistent minimum fitting error was obtained for the structure optimized with a vertical displacement of +1 Å. This represents a one-residue shift, or less, relative to the original EPR estimate.

[a] Final structure optimized with +1 Å vertical shift of the protein.
[b] Adjusted manually, producing an isoenergetic structure (see text)

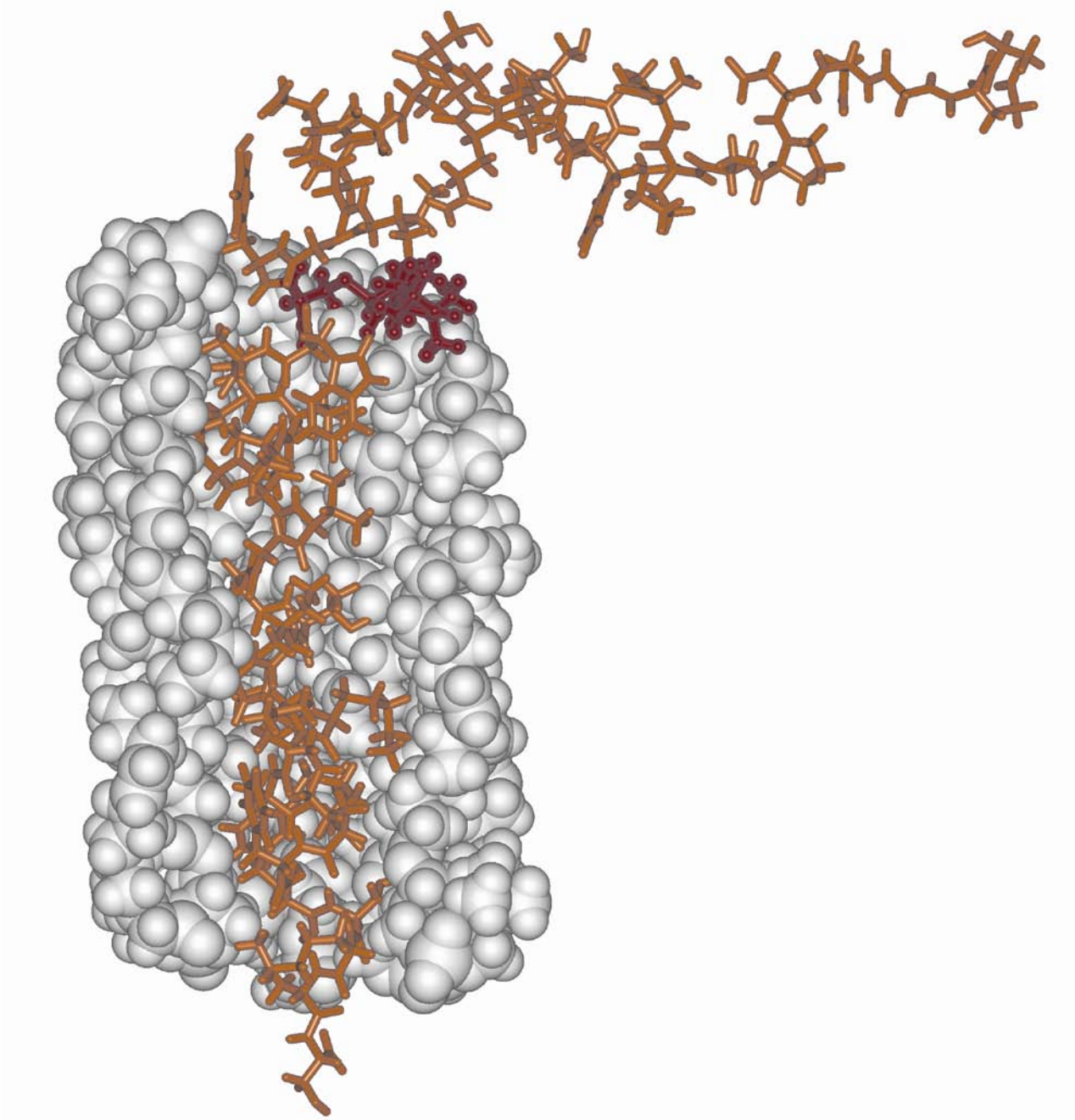

**Figure 5.** ***Structure of C25-spin labeled M13 MCP (stick representation) surrounded by a single bilayer solvation shell of DOPC phospholipids. For clarity, only part of the lipid shell is shown—in space-filling representation. The structure was chosen to satisfy the experimental EPR constraints and was obtained by geometry optimization of the lipids and all protein side chains. C25, with the 5-maleimidoproxyl label bound to it, is enhanced with dark ball-and-stick representation. The figure is created using Insight II***

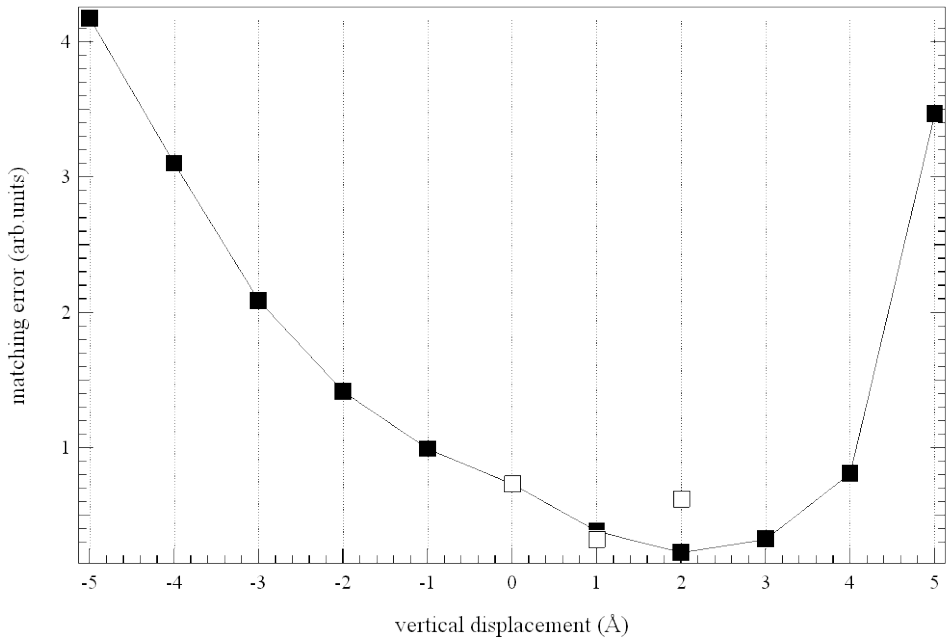


**Figure 6.** ***Matching error between the sequence profiles of the $^{14}N$ outer hyperfine splitting (Stopar et al., 1997a) and f-parameter, as a function of vertical displacement of the M13 major coat protein structure, relative to that shown in Fig. 6, in a DOPC bilayer. The minimum sum of squares of the differences was obtained by linear scaling of the outer splitting profile (with offset) to that of the f-parameter. Solid symbols indicate that the protein structure was not reoptimized at the new vertical position; open symbols denote reoptimization of the structure at the shifted position. Positive shift means that the N-terminal helix is moved closer to the membrane.***

The resulting sequence profile of the *f*-parameter is compared with that of the EPR outer hyperfine splittings in Figure 7. It is seen that the packing parameters deduced from the final optimized structure reproduce the major features of the experimentally determined mobility profile rather well. Including lipids in the model considerably improves the agreement with experiment (cf. Fig. 4); the rms matching error is reduced from 0.86 to 0.32, in the presence of lipids.

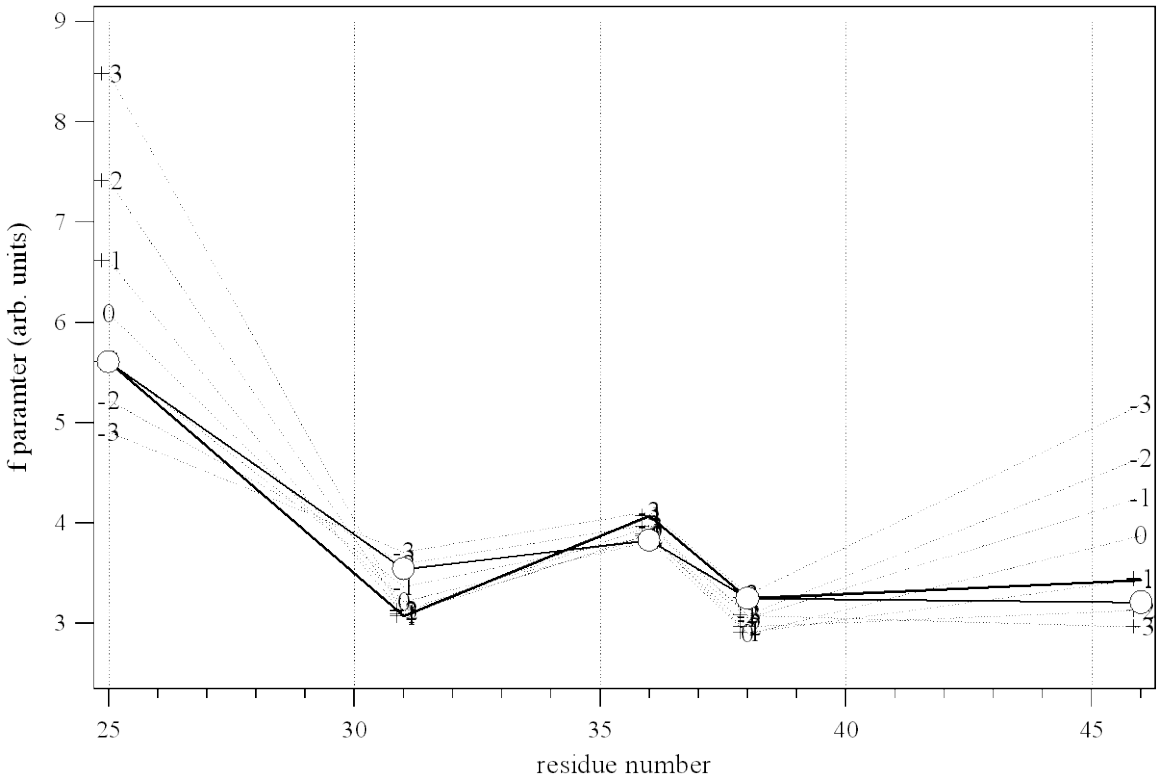


**Figure 7.** ***Sequence profiles of the f-parameter from cysteine–maleimide mutants of the M13 MCP at different vertical shifts (broken lines), for the structure shown in Fig. 5 that was optimized with zero shift. The heavy solid line gives the sequence profile for the final structure that is reoptimized with +1 Å vertical shift. Circles are the experimental outer hyperfine splitting profile that is scaled linearly (with offset) for best match to the f-parameter profile of the final (+1 Å-shifted) structure.***

## *Cytochromes $b_{561}$*

### Features of the cytochrome $b_{561}$ family at the primary structural level

Recent studies involving sequence comparisons have suggested a close relationship between cytochrome $b_{561}$ proteins of the plant and animal kingdom and demonstrated the conservation of a number of structural features (Asard et al., 2001;Ponting, 2001;McKie et al., 2001;Asada et al., 2002). In order to identify additional conserved properties and structural features of sequences related to the well-known plant and human cytochrome $b_{561}$ proteins, more members of the family had to be included in a sequence comparison. Related sequences were identified from

26 different tissues and organisms via PSI-BLAST searches with default settings (Altschul S.F. et al., 1997) using Artb561-1, Artb561-4 and Hosb561-1 as queries (sequence names are used as defined in (Asard et al., 2001)). Sequences were selected which (1) had previously been reported to belong to this family (Asard et al., 2001), (2) showed conserved functionally relevant key residues in similar locations, primarily the heme-ligating histidines and the SLHSW motif (a putative MDA binding site; (Tsubaki et al., 1997;Asard et al., 2001)) and (3) were at least 200 residues long. This latter constraint was needed to exclude incomplete gene fragments (Asard et al., 2001).

The sequences were aligned using MULTICLUSTAL (Yuan et al., 1999). Figure 8 represents the most complete and detailed sequence alignment of the cytochrome $b_{561}$ family to date. All sequences contain 6 regions that are rich in highly conserved amino acid residues or residue properties. These conserved regions are separated by regions with no or very little conservation and most of the gaps can be found in these nonconserved regions too. Clearly, the sequences display largest variability in these intermediate and the terminal nonconserved regions. Notable is that the nonconserved regions are too short to form membrane-spanning alpha-helices (TM helices), especially if the gaps are also taken into account. The conserved regions, on the other hand, overlap very well with the 6 TM helices (Fig. 8, boldface letters). These were predicted, independently from the alignment, for each sequence by TMHMM (version 2.0) (Krogh A et al., 2001).For one sequence (Caeb561-3) there was a seventh TM helix predicted at the N-terminal end of the protein (Fig. 8). However, this was unique to this sequence, since most of the other sequences contained very few residues in that region.

Based on the above observations, the conserved regions can be considered to be TM helices (named TMH1 through TMH6) and the nonconserved regions can be considered to be interconnecting loops and terminal domains. By any measure, the conservation of the TM helices 2–5 (i.e., TMH2 to -5) is much higher than that of the two terminal ones (TMH1 and -6).This is not surprising considering that all the known functionally important residues are located in the region defined by TMH2 to -5, i.e., the region 95–240 (numbering according to the “consensus” sequence shown

in the ruler of the alignment). The putative MDA binding site (the SLHSW motif;(Tsubaki et al., 1997;Asard et al., 2001)) is located sequentially at the beginning of TMH4 and spatially close to the lipid–water interface. Since MDA binds at the noncytoplasmic side of the plasma membrane in plants (Asard et al., 1992;Horemans et al., 1994), this orients the whole putative structure in the membrane. This automatically locates the putative Asc binding site, i.e., the sequence segment 114–122, on the cytoplasmic side of the membrane close to the membrane–water interface too (Okuyama et al., 1998). In this site, the ALLVYRVFR motif is fully conserved in 8 sequences in its full length and almost all sequences contain at least 4 residues from it (the presence of this motif was not a search criterion). In addition, residue properties (aliphatic, aromatic, and basic) at 4 positions in this motif are conserved to nearly 100%.The SLHSW motif contains one of the 4 histidine residues that are 100% conserved. These 4 histidines are located in all sequences on TMH2 (H99) and TMH4 (H177) and on TMH3 (H135) and TMH5 (H220) in a pair wise manner, ideal to anchor two hemes on the noncytoplasmic and cytoplasmic side of membrane, respectively, as proposed earlier (e.g., (Degli Esposti M. et al., 1989;Okuyama et al., 1998;Tsubaki et al., 2000)).

A number of highly conserved aromatic amino acid residues are identified in the TM helices between the heme-ligating histidine residues. TMH4 appears to be richest in conserved aromatic residues. Notable is the presence of 2 highly conserved basic residues located between TMH2 and TMH3 at positions 128, and 132. We propose that these basic residues participate at the Asc binding, i.e. sequential and not well conserved “ALLVYRVFR” motif should be extended with the following it 2 basic residues. The TM architecture in the alignment suggests cytoplasmic orientation for both the N and C termini. It should be noted, that TMHMM (Krogh A et al., 2001) predicted an opposite orientation for the sequences Artb561-1, Lyeb561-1, Zemb561, Orsb561-1, and Mecb561, but MEMSAT 2 (McGuffin LJ et al., 2000) predicted them to also have the N and C termini in the cytoplasmic side.

## Conservation of the hydrophobic moments of transmembrane helices of the cytochromes $b_{561}$

Considering the differences in conservation between the terminal and central TM helices, our efforts to build structures were restricted to the inner core of the cytochrome $b_{561}$ proteins consisting of 4 TM helices (TMH2 to -5). These TM helices define highly conserved membrane-spanning regions and the fully conserved histidine residues define their overall spatial relations. However, knowledge is still lacking on the angular orientation, i.e., the sides facing towards the interior of the core or towards the lipids, of the individual TM helices (heme ligation alone leaves still too much freedom).Therefore, as a further structural constraint, we tested the lipid-facing propensities of the TM helices and their conservation in the family. To do this, first the consensus regions of the 4 TM helices were defined in the multiple alignment (Fig. 8). Considering the position of the 4 by 26 individually predicted TM helices, we defined the consensus positions of TMH2, TMH3, TMH4, and TMH5 as 95–118(i), 133(i)–156, 175–200(i), and 218(i)–240, respectively (“i” indicates the cytoplasmic side of the helices). These regions were selected by considering the mathematical averages, but highly conserved residues known to prefer specific membrane locations (Killian JA and von Heijne G, 2000) were also taken into account. The TM helices TMH2 to -5 are indicated with horizontal bars under the sequences in Fig. 8. The 26 sequence segments of the 4 TM helix regions of the same length were subjected together for an analysis of their lipid-facing propensity in 4 separate submissions. The analysis was performed using the knowledge-based kPROT method, which scores residues with free-energy-like values (Pilpel Y et al., 1999). These kPROT values are related to the likelihood of a given residue to be oriented towards the lipids in TM alpha-helices. The algorithm treats TM helices as ideal alpha-helices. The vectorial sum of the kPROT values over the helical wheel of a single TM helix defines the side of the TM helix that is most likely oriented towards the lipids. This analysis was performed for all of the consensus TM helix

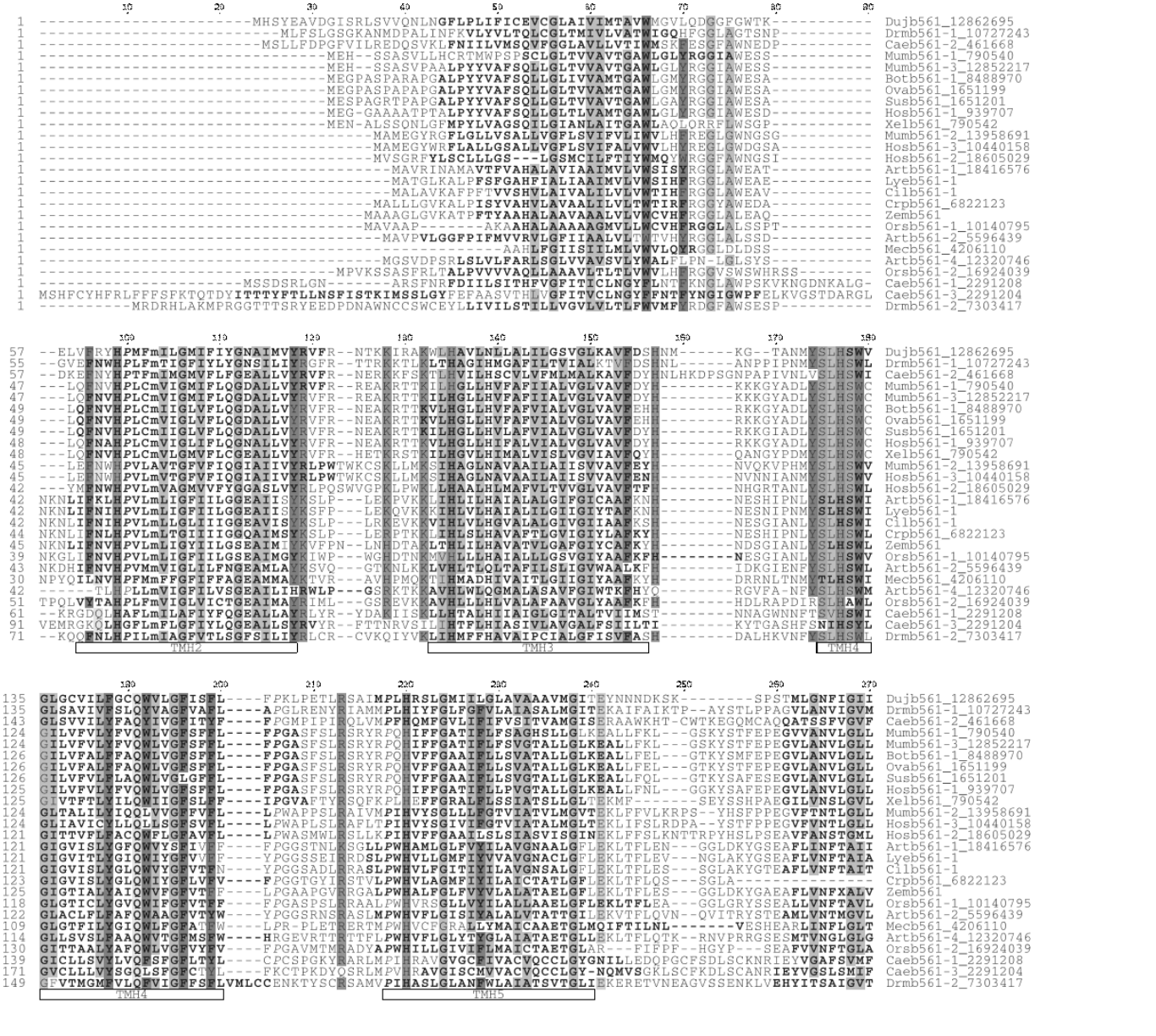
TMH2
TMH3
TMH4
TMH4
TMH5

**Figure 8.** ***MULTICLUSTAL alignment of 26 cytochrome $b_{561}$ sequences. The names and numbers before and after the underscore are names according to the proposed naming convention (Asard et al., 2001) and GenBank identifiers (Benson et al. 2002), respectively. Highlighting is performed with TeXshade (Beitz, 2000) above 85% threshold conservation of residue properties and according to the chemical mode. Concerning font styles and grey shades, on a relative scale between white and black of 0 to 100%, residue properties (Karlin 1985) are represented as follows: aromatic (F,W, Y), 50%; basic (H, K, R), 42%; hydroxyl (S, T), 34%; aliphatic (A, G, I, L, V), 26%; acidic (D, E), 18%; amide (N, Q), 10%; imino (P), font in italics (no shading); sulfur (C, M), lower case font (no shading). Amino acids in the TM segments, predicted by TMHMM (Krogh et al. 2001) independently for each sequence, are shown in boldface. The ruler above the alignment is numbered according to the consensus sequence. The horizontal bars indicate the consensus TM helices TMH2 to -5 that are most conserved***

regions, i.e., 4 by 26 sequence segments. The radial thin solid lines in Fig. 9 show these vectorial sums for the respective TM helices of all species. Also shown are the mean vectors, which were obtained as averages over the 26 vectors, defining the (conserved) side of each TM helix most likely facing the lipids (thick lines with arrows). The scattering around the mean vector varies for each helix, is generally small, and is remarkably better for TMH2 and TMH4. The larger scattering in TMH3 and TMH5 in part originates from imposing consensus helices on the TM helix positions, which are less well conserved than in TMH2 and TMH4.

The generally good conservation of the lipid-facing propensity in the cytochrome $b_{561}$ family not only supports our multiple alignment, but it also argues that other helix geometries are very unlikely. The imposed constraints allow two alternative helix topologies, which result in clockwise or counterclockwise top view (from the non-cytoplasmic side) arrangement for TMH2 to -5 in the membrane plane. These topologies allow heme ligation and face helices as favorably towards lipids as possible. Very striking is that the most variant residue positions (with a single exception at residue 117) and the mean direction of the most variable side of the helices (Fig. 9, italic typeface residues and dashed lines, respectively) are all located in the outer surface of the 4-helix core. In contrast, most conserved residues and residue properties (shown in boldface) are least frequent in these regions. Underlined numbers in Fig. 9 indicate sequence positions of highly conserved aromatic residues. Some of these, in particular those in TMH4, are located in favorable positions, both laterally and vertically, to participate in the electron transfer between the two hemes.

## Model structures for the cytochromes $b_{561}$

Model structures were built for the 4-helix core of representatives of a mammalian (*Homo sapiens*) and plant (*Arabidopsis thaliana*) cytochrome $b_{561}$ sequence, Hosb561-1 and Artb561-1 (Fig. 10). The primary constraints for the models were: (1) the predicted (and not consensus) TM helix sequence regions (Fig. 8), (2) the requirement for ligation of the two hemes by the fully conserved 4 histidine residues (Fig. 9), and (3) the ability to satisfy the lipid-facing propensities of the individual helices (Fig. 9) as far as possible. Residues with known preferential membrane location were also considered as additional constraints (Killian JA and von Heijne G, 2000).

It should be kept in mind that the accuracy in the orientation of the lipid-facing propensity of all the available prediction algorithms is rather limited because only few membrane protein structures are known to calibrate these algorithms (Pilpel Y et al., 1999). Therefore the third criterion was weaker than the first two. The two cysteine residues conserved at the positions 102 and 180 in 7 animal sequences were not bridged because they were apparently too far apart (Fig. 9) and a rotation of TMH2 and TMH4 to bring them closer together would destroy the optimal helix orientations. In addition, Kent and Fleming (Kent and Fleming, 1990) found that all cysteines were in the free sulfhydryl form in chromaffin granule cytochrome $b_{561}$. Both topologies still leave some “freedom” in orienting the helix axes relative to the membrane normal and along their long axis (Fig. 9). These uncertainties can be estimated to be below ca. 25° and 30°, respectively (see Fig. 9).

The locations of TMH1 and TMH6 are not firmly determined by the primary constraints imposed on the models (see above) and are different for the two TMH2 to -5 topologies (Fig. 9). The unique locations of TMH1 and TMH6 in the clockwise and counterclockwise topologies result from the requirement to avoid crossing of the helix-connecting loops and generating too large distances between sequentially adjacent helices. The topology with TMH1 and TMH6 located on the opposite sides of the core helices (Fig. 9, bottom) is somewhat more likely than the circular one

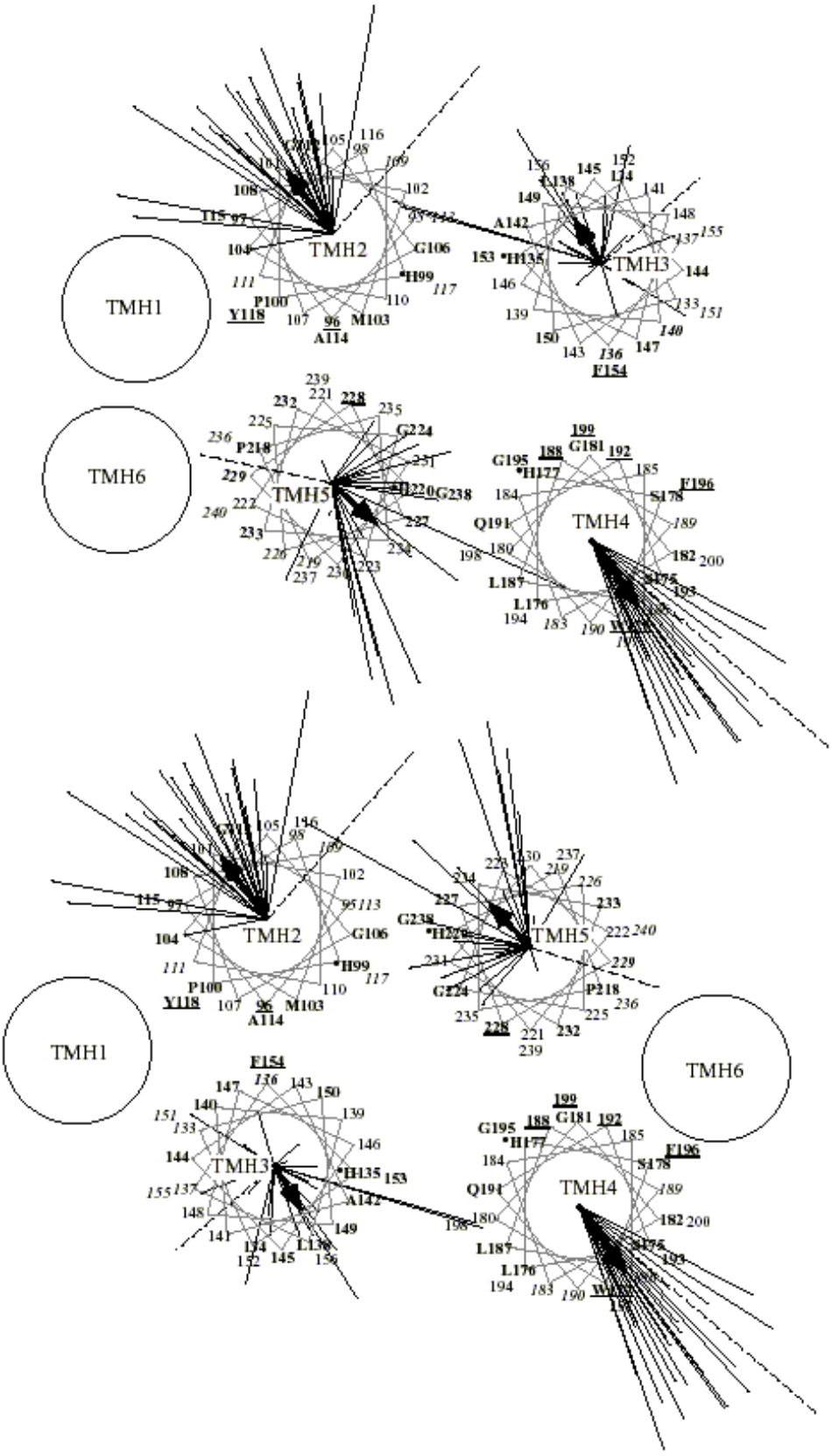
TMH1
TMH2
TMH3
TMH4
TMH5
TMH6
TMH1
TMH2
TMH3
TMH4
TMH5
TMH6

**Figure 9.** ***Lipid-facing propensity of the individual and consensus TM helices. kPROT vectors (Pilpel Y et al., 1999) for the individual proteins are shown with radial thin lines (not all being visible) in the helical wheels of the consensus transmembrane helices TMH2 to -5 defined by the sequence regions 95–118(i), 133(i)–156, 175–200(i), and 218(i)–240, respectively, according to the consensus sequence numbering (gaps were removed for the prediction). The two most likely helix topologies from which structures were built are shown as viewed from the noncytoplasmic side of the membrane (top view). Accordingly, for the N-terminal cytoplasmic (TMH3, TMH5) and noncytoplasmic (TMH2, TMH4) helices, sequence numbering goes counterclockwise and clockwise, respectively. The thick solid line with an arrow represents the mean vector that is the average over all the 26 corresponding TM helices. The dashed lines indicate the direction in which variation of the lipid-facing propensity of the residues over the different sequences is largest. All lines are proportional to the numerical values they represent. Consensus residue numbers are indicated along the wheel together with residue names in single-letter code at locations where residues are highly conserved (cf. Fig. 8). The most variant residue positions are indicated in italic typeface. Identifiers in bold typeface indicate highly conserved residues and residue properties, with and without residue names, respectively. The putative heme-ligating histidine residues are indicated with bullets. Underlined numbers indicate sequence positions of highly conserved aromatic residues***

(Fig. 9, top) because the latter one leaves a relatively large less “lipophilic” surface on TMH3 and TMH4 exposed to lipids.

Idealized alpha-helices were built for the predicted TMH2 to -5 helix sequence segments using Biopolymer in InsightII (Accelrys 2000). They were arranged manually according to the two topologies and the predicted lipid-facing propensities (Fig. 9) and to obtain an approximate distance of 1.2 nm between the $C_\alpha$ atoms of the pairs of the fully conserved heme-ligating histidine residues. This histidine-to-histidine distance was determined after an inspection of coordinates of 34 protein structures in the Brookhaven Protein Data Bank (Berman et al., 2000) which contained b-hemes ligated by the $N_\varepsilon$ atoms of histidine residue pairs. The helices could be easily arranged to satisfy all the criteria without tilting them significantly relative to the membrane normal. Interconnecting loops between the TM helices were identified after an extensive search using Homology (in InsightII). In this process preflex and postflex 5-residue-long sequence matches between the known and predicted helices were scored to identify the best matching loop template structures of the correct length. The coordinates of these loop templates were then assigned to the interconnecting loop sequences. Steric conflicts between amino acid side chains were removed by manual adjustment. In order to relax side chains into lower energy states, a conformational search of 50 cycles was performed for all the

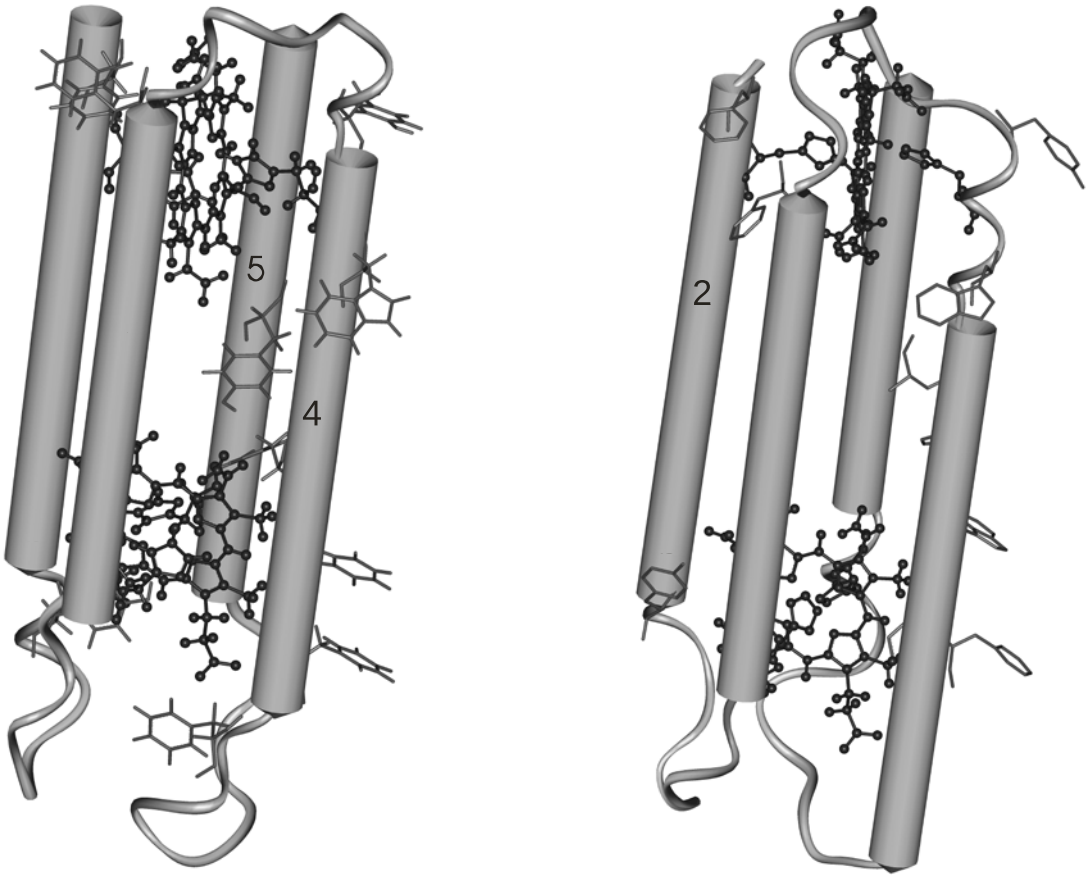


**Figure 10.** ***Predicted 3-D structure with counterclockwise helix topology (as defined in Fig. 9, bottom) for the Arabidopsis thaliana cytochrome $b_{561}$ sequence Artb561-1 (left) and Homo sapiens cytochrome $b_{561}$ sequence Hosb561-1 (right). Only the most conserved helices (TMH2 to -5) are shown, together with interconnecting loops. The ribbon cartoon was created with InsightII (Accelrys 2000) and represents secondary structural forms (i.e., ideal alpha-helix and random coil). Highly conserved aromatic amino acid residues are indicated in light grey thin bar presentation. The two pairs of fully conserved histidine residues (H99 – H176 and H135 – H220) and the hemes ligated by them are shown in dark grey ball-and-stick presentation. The TM helix numbers are indicated on the cylinders. The cytoplasmic side is below each structure. The putative MDA (N-terminus of TMH-4) and Asc (C-terminus of TMH-2) binding sites are indicated with the corresponding text. The helix axes are closely perpendicular to the membrane plane***

amino acid side chains of the models using Homology (in InsightII) with the default 0.8 nm cutoff for both Van der Waals and Coulomb interactions. The structures were further relaxed using Homology with 100 iterations of the steepest-descent algorithm followed by 1000 iterations of the conjugate-gradient algorithm. In this final step all the side chains of the TM regions and all the atoms in the loop regions were allowed

to move but, as additional constraints, the distances between the $N_\varepsilon$ atoms of the pairs of the heme-ligating histidine residues were fixed at about 0.4 nm. This distance was also determined from inspecting the relevant b-heme containing protein structures. The structure of the counterclockwise topology for the 4-helix core of the Artb561-1 protein is shown in Fig.10 together with the hemes that were inserted manually into the structures after optimization. Considering the high level of conservation of structural features in the cytochrome $b_{561}$ protein family, it is anticipated that the 3-D structure construction procedure described above would result in similar models for the other members of the family.

## Discussion

### *M13 MCP Final Model*

In the final optimized model, residues Y24 and F45 are located in the somewhat diffuse regions of the lipid phosphates on either side of the bilayer membrane (cf. A25 and T46 in the original experimental paper (Stopar et al., 1997b)). Our prediction of the shift in approximately 1 residue of the TM part of the M13 MCP, relative to the membrane normal towards the N-terminus, has been supported later experimentally (Marassi and Opella, 2003). This finding in turn justifies our manual orientation of the spin-label attached to the C46 – being positioned deeper in the aqueous phase the protein backbone will not preserve its α-helical conformation (responsible for the approximately perpendicular orientation of the long axis of the 5-MSL), but rather will achieve certain randomized configuration with comparably bulky side-chain of the spin-labelled C46 pointing away from the membrane surface.

Formation of H-bonds between the ε-amino groups of K43 and K44 and the carbonyl oxygens of the lipid fatty acid chains is suggested by the model. To some extent this resembles the snorkel effect (Monne et al., 1998). At the opposite side of the bilayer, the model places W26 in a position where it can function as a membrane-anchoring residue (Schiffer et al. 1992). In the mutated structure, the ε-amino group

of K40 interacts with oxygens of the maleimide ring for the spin label attached to C36. This interaction reduces the overall potential energy of the system by 20 kcal/mol (evaluated in Sculpt (Surles et al., 1994)). As we see it now, this may be the only reason for the increased outer hyperfine splitting at C36 reported in Stopar et al. (Stopar et al., 1996;Stopar et al., 1997a). Previously, it was speculated, that immobilisation of the spin-label at C36 could have been caused as well by the non-covalent dimerization of the TM helices of the adjacent (in the membrane) M13 MCP molecules. But recently Melnyk and colleagues (Melnyk et al., 2004) have described the exact dimer conformation for the TM parts of the M13 MCP (see Fig. 1 in (Melnyk et al., 2004), it could be clearly seen, that T36 is left far outside from the dimerization surface, formed by the sequential G(34)xxxG(38) motif). This last publication culminates the series of works performed in Charles M. Deber's lab, devoted to study of the dimerization of the TM helices of M13 MCP in artificial and natural membranes ((Deber et al., 1992;Deber et al., 1993;Li et al., 1993;Khan and Deber, 1995;Dawson et al., 2003)). It is should be noted, that comparing with other non-covalent TM dimers, those formed by the M13 MCP are denoted as "weakest" on relative scale of known non-covalent TM dimers ((Melnyk et al., 2004)), which is obviously related to the fact that M13 MCP exists in a few distinct conformations during the phage life cycle.

Putative H-bonds between the N–H group of the indole ring and the carbonyl oxygens of the fatty acid chains in adjacent lipid molecules possibly contribute to immobilization of the spin label on C25 in the mutated structure. The N-terminal helix of the final structure is oriented parallel to the membrane surface (see Fig. 5), consistent with solid-state NMR results on the closely related (D12N) fd coat protein in oriented membranes (McDonnell et al., 1993) and (Marassi and Opella, 2003). Refinement in 1997 of the 3-D structure of the MCP (PDB ID: 1FDM) from the fd phage in SDS micelles with specific loop in the hinge region (Almeida and Opella, 1997) and then publication of the structures of the M13 MCP (used in current work) (Papavoine et al., 1998), has caused a discussion, concerning the exact conformation of the hinge region. Recently, Marassi and Opella (Marassi and Opella, 2003) refined the 3-D structure of the fd MCP (PDB ID: 1MZT) in lipid bilayers, using an

alternative NMR approach. The hinge region now lacks any specific or unusual turn and reminds an unwound α-helix (i.e. finally, in general accordance with the structures from (Papavoine et al., 1998)).

An interesting feature of the amphipathic N-terminal surface helix (residues 8–16) of our M13 MCP is its azimuthal orientation relative to the membrane surface. This differs somewhat between the NMR structures determined in DodPC and SDS micelles (Papavoine et al., 1998). As already noted, the EPR data from phosphatidylcholine bilayer membranes, although not referring specifically to the N-terminal structure, are consistent only with the SDS-family. The N-terminal helix is oriented with the face containing the charged and polar residues K8 and Q15, and aromatic residue F11, directed towards the membrane. Correspondingly, the opposite face containing alanine residues 9, 10, and 16, and S13, is directed towards the aqueous phase (cf. (Papavoine et al., 1994)). Such an orientation may not be the optimal one, as one would expect the strongly hydrophobic L14 to be oriented toward the membrane rather than equatorially as it is in the model (as it appears in chronologically later work of (Marassi and Opella, 2003)). This orientation of the N-terminal helix is likely to be specific for the strongly negative surface potential of SDS micelles and may be relieved by a slight (not more than 30° clockwise, cf. Fig. 3c in (Marassi and Opella, 2003)) twist of the helix with respect to the membrane in zwitterionic lipids.

The configuration of the hinge region, proposed in (Marassi and Opella, 2003) differs from ours in relative positioning of the residue A25 – in former case this residue resides completely at the outer side of the hinge region, connecting amphipathic and TM helices of the M13 MCP (and positioning A25 inside of the loop was one of our strongest requirements during model building and selection). Nevertheless, it is necessary to have in mind, that the actual conformation of the hinge region from (Marassi and Opella, 2003) is not refined using any actual experimental constraints, but is a result of pure computational effort (Marassi and Opella, 2003).

The final model remains hypothetical both in the sense that the starting structures are for a micelle rather than a bilayer environment (it is worth to notice,

that while accepting, that the structures from the (Almeida and Opella, 1997) are most probably erroneous in the hinge region, Marassi and Opella (Marassi and Opella, 2003) strongly support an argument in favor of the use of lipid bilayer samples for structure determination of membrane proteins). More particularly that only one coarse constraint has been used to distinguish between the N-terminal starting structures.

### *Cytochromes $b_{561}$*

The present sequence alignment including plant and animal members of the cytochrome $b_{561}$ family (Fig. 8) supports the main conclusions on the conservation of functional elements from recent analyses on a smaller subset of the cytochrome $b_{561}$ family (Asard et al., 2001). Together with the structures presented in Figs. 9 and 10, this alignment sheds light on more structural details and raises a number of questions. Our observations provide evidence that the functionally relevant and structurally most conserved region in the cytochrome $b_{561}$ family is the TMH2 to -5 4-helix core with an amino acid composition that is very well conserved in the inner surface and somewhat less conserved in the outer surface of the core. The two terminal helices (TMH1 and TMH6) are less conserved (Fig. 8). They together with the interhelix loops and terminal regions are the main source of the variability in the family and may therefore define the specific subcellular location, physiological functions of the proteins they encode, and possibly their interactions with other proteins.

The 4-helix core surrounds and ligates two heme molecules by 4 fully conserved histidine residues, closely located to the membrane–water interface on the opposite sides of the membrane. Since the putative, well or highly conserved Asc and MDA binding sites, and other highly conserved residues with yet unknown function, are located in this region, this 4-helix core may represent conserved transmembrane electron transfer machinery. Recent findings demonstrating that this core structure occurs as a domain in other proteins in plants and animals support this idea (Ponting, 2001).

The membrane orientation of the 4-helix core presented in Fig.10 can be taken with quite some confidence, considering the experimental evidence for the location of MDA and Asc binding sites in the noncytoplasmic and cytoplasmic sides of the membranes, respectively, for the plant plasma membrane cytochrome $b_{561}$ and chromaffin granule cytochrome $b_{561}$ (Kelley et al., 1990;Asard et al., 1992;Okuyama et al., 1998). The orientation of this core is also in agreement with biochemical data on the location of some of the highly conserved histidine residues (Tsubaki et al., 2000). The position of TMH6 in the sequences is well conserved (Fig. 8), suggesting that the C terminus is located on the cytoplasmic side, i.e., on the same side as that of the Asc binding site. This is in agreement with experimental data (see, e.g., (Kent and Fleming, 1990)). However, the position of TMH1 in the primary sequences and residues in TMH1 are remarkably less conserved. Partly as a result, the two algorithms for the prediction of the membrane orientation of cytochrome $b_{561}$ yielded conflicting membrane sidedness for 5 of the 26 proteins. This result is particularly interesting in view of a recent study on the duodenal cytochrome $b_{561}$ (McKie et al., 2001). In that study, antibodies directed against C-terminal peptides inhibited cytochrome $b_{561}$-mediated ferric reductase activity in cell cultures expressing the cytochrome, suggesting that the C terminus was located on the extracellular surface. Further experiments are needed to clarify this apparent contradiction with the model supported by our work.

The high conservation of the motifs at 175–179 and 114–122, i.e., the putative MDA and Asc binding sites, respectively, is a further strong feature of the cytochrome $b_{561}$ family. This suggests a key functional role for these putative binding sites in transmembrane electron transfer common to this protein family. The role of aromatic residues in intramolecular electron transfer in proteins is an important topic in biophysics (see, e.g., Casimiro et al. 1993, Farver et al. 1997, Cheung et al. 1999). There are several highly conserved residues located favorably between the two pairs of heme ligating histidine residues (Figs. 9 and 10). Of these, the aromatic residues could indeed constitute the putative transmembrane electron transport pathway. In addition, there are a few additional (nonconserved) aromatic residues in many sequences that could also contribute to such a pathway. It will be an important

subject for future studies to test whether some or all of these conserved aromatic residues are essential for transmembrane electron transfer in cytochrome $b_{561}$ proteins. It is interesting that the location of the fully conserved histidine residues in TMH3 and TMH5 are rather close to the lipid-facing sides in these "mean" topologies (see bullets and solid vectors in Fig. 9). This raises the question whether structural details of heme ligation in the cytoplasmic side are possibly less conserved in the family, and/or heme ligation may be even less stable, than in the noncytoplasmic side. It should be noted that H139 and H157 are also well but not fully conserved. It has been argued that H139 does not participate in heme ligation because it is not present in some of the animal proteins (Asard et al., 2001;Asada et al., 2002). Chemical modification of histidine residues also supports this interpretation (Tsubaki et al., 2000). Histidine residues are believed to play an important direct role in proton movement coupled to the electron transfer reaction (Njus et al., 2001;Kipp et al., 2001). Whether H139 and H157 manifest some roles in heme ligation and/or in the electron transfer mechanism in different members of this protein family remains to be explored (note that some of the residue replacements appear simultaneously at these positions in some sequences; see Fig. 8).

The high structural similarity between the plant and animal cytochrome $b_{561}$ proteins, both at the sequence and protein structural level, suggests that the conserved machinery of transmembrane electron transfer mediated by these proteins serves diverse, yet to be explored physiological processes in eukaryotic cells. Of the di-heme proteins with known structure, the membranous subunit C of fumarate reductase from *Wolinella succinogenes*, available at 0.22 nm resolution ((Lancester CRD et al., 1999); PDB ID: 1QLA), is most relevant to the present study, although its sequence is very different from that of the cytochrome $b_{561}$ family (hence not shown in the alignment). Similarly to our models, two pairs of histidine residues on 4 TM helices coordinate the two hemes and the heme planes have similar orientations. The orientation of the 4-helix bundle relative to the direction of the electron flow is also the same. However, the TM helices are much longer than those in the cytochrome $b_{561}$ family and, consequently, are significantly tilted and kinked. Though there is quite some freedom in the helix orientations, a similar degree of TM

tilts and kinks in our models would result in a large hydrophobic mismatch with the membrane. The shorter heme-to-heme distance of about 1.5 nm in subunit C of fumarate reductase is in part due to the tilts and kinks in the TM helices in addition to the fact that the heme-ligating residues are located much closer to the center of the TM helices than in our models.

## Conclusion, Main Results

### *M13 MCP model*

- Relatively coarse-grained site-directed spin-label measurements have provided sufficient experimental constraints to select a single structural subclass from the family of high resolution NMR structures in micelles as being that most appropriate to the M13 MCP in lipid bilayer membranes.
- General configuration and topology of the selected structural subclass agrees with new, independent findings.
- A relatively simple indicator of the local packing density (the *f*-parameter), which is readily calculated from the coordinates of the optimized protein–lipid structural model, was found to be adequate for this purpose and has been further tested and developed.
- Extension of the approach to sparse experimental data on site-directed mutagenesis of other membrane proteins should be possible in the future.
- In our protein–lipid model, Y24 on one side of the membrane, and K43 and K44 on the other side, interact preferentially with the lipid head groups. The model indicates a hydrophobic mismatch of 3.5 Å or less (the protein is slightly shorter) between the unperturbed phospholipid bilayer and the intramembranous α-helix of the protein.
- Spin-labeled C25 is buried inside the hinge region, whereas C46 points towards the aqueous phase, in agreement with their strong and weak motional restriction, respectively.
- Shift in approximately 1 residue of the TM helix along the membrane normal towards the N-terminus is predicted: Y24 – F45 from previous experimental A25 – T46. The result has also been justified with independent, new data.
- Spin-label at C36 is restricted solely by involvement of its maleimide carbonyl oxygens in hydrogen bonding with K40.
- The model proves useful for the interpretation of future experimental data on membrane–M13 MCP systems. It became good starting point for full-scale

molecular dynamics simulations and for the design of further site-specific spectroscopic experiments.

### *Cytochrome $b_{561}$ Model*

- The most detailed and extensive sequence alignment and analysis to date for the representatives of the cytochrome $b_{561}$ family was performed.
- Transmembrane regions and lipophilic properties of the sequences have been obtained.
- 2 possible topological models of 2-D TMH arrangement were proposed and discussed.
- 3-D atomic models of the 4-TMH core of the mammalian and plant sequences were built for both 2-D topologies, representing transmembrane electron transfer machinery.
- The present 3-D structures provide useful working models for designing combined point mutation and biophysical experiments targeting heme ligation and putative electron transport pathways.
- The present models will be further refined as new structural data emerge in the future.

**Acknowledgements**

I am indebted to my mentors, Dr. László Dux and Dr. Tibor Páli for their continuous scientific guidance, encouragement and support throughout my Ph.D studies. I express my gratitude to Dr. Pál Ormos, director of the Institute of Biophysics at Biological Research Center, for allowing me to work at this research facility. I am thankful for the help I received from Dr. Alajos Bérczi and Dr. Balázs Szalontai. I am grateful to Dr. Derek Marsh from the Max-Planck Institute for Biophysical Chemistry, for the opportunity to visit his research group. I would like to thank Dr. Peter Knox for his advices and support. I am grateful to Dr. Erzsébet Szatmári and the rest of my family for their trust in me. Finally, I would like to thank all the members of the Institute of Biophysics, Usenet and Debian communities for the friendly and supportive environment.

**Bibliography**


1. Accelrys. InsightII Modeling Environment, Release 2000 (June 2000) and Biopolymer and Homology (March 2000). Accelrys Inc., San Diego, Calif. Accelrys Inc., San Diego, CA . 2000.

2. Almeida,F.C.L. and S.J.Opella. 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.

3. Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., and Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389-3402.

4. Asada,A., T.Kusakawa, H.Orii, K.Agata, K.Watanabe, and M.Tsubaki. 2002. Planarian cytochrome b(561): conservation of a six transmembrane structure and localization along the central and peripheral nervous system. *J. Biochem. (Tokyo)* 131:175-182.

5. Asard,H., N.Horemans, and R.J.Caubergs. 1992. Transmembrane electron transport in ascorbate-loaded plasma membrane vesicles from higher plants involves a b-type cytochrome. *FEBS Lett.* 306:143-146.

6. Asard,H., J.Kapila, W.Verelst, and A.Berczi. 2001. Higher-plant plasma membrane cytochrome b561: a protein in search of a function. *Protoplasma* 217:77-93.

7. Asard,H., J.Terol-Alcayde, V.Preger, J.Favero, W.Verelst, F.Sparla, M.Perez-Alonso, and P.Trost. 2000. *Arabidopsis thaliana* sequence analysis confirms the presence of cyt b561 in plants: evidence for a novel protein family. *Plant. Physiol. Biochem* 38:905-912.

8. Beitz,E. 2000. TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics.* 16:135-139.

9. Berman,H.M., J.Westbrook, Z.Feng, G.Gilliland, T.N.Bhat, H.Weissig, I.N.Shindyalov, and P.E.Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235-242.

10. Blanch,E.W., L.Hecht, C.D.Syme, V.Volpetti, G.P.Lomonossoff, K.Nielsen, and L.D.Barron. 2002. Molecular structures of viruses from Raman optical activity. *J. Gen. Virol.* 83:2593-2600.

11. Curran,A.R. and D.M.Engelman. 2003. Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr. Opin. Struct. Biol.* 13:412-417.

12. Dawson,J.P., R.A.Melnyk, C.M.Deber, and D.M.Engelman. 2003. Sequence context strongly modulates association of polar residues in transmembrane helices. *J. Mol. Biol.* 331:255-262.

13. Deber,C.M., A.R.Khan, Z.Li, C.Joensson, M.Glibowicka, and J.Wang. 1993. Val-->Ala mutations selectively alter helix-helix packing in the transmembrane segment of phage M13 coat protein. *Proc. Natl. Acad. Sci. U. S. A* 90:11648-11652.

14. Deber,C.M., Z.Li, C.Joensson, M.Glibowicka, and G.Y.Xu. 1992. Transmembrane region of wild-type and mutant M13 coat proteins. Conformational role of beta-branched residues. *J. Biol. Chem.* 267:5296-5300.

15. Degli Esposti M., Y.Kamensky, A.M.Arutjunjan, and A.A.Konstantinov. 1989. A model for the molecular organization of cytochrome beta-561 in chromaffin granule membranes. *FEBS Lett.* 254:74-78.

16. Eisenberg,D., E.Schwarz, Komaromy,M., and R.Wall. 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* 179:125-142.

17. Halgren,T.A. 1996a. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization and Performance of MMFF94. *J. Comp. Chem* 17:490-519.

18. Halgren,T.A. 1996b. Merck Molecular Force Field. II. MMFF94 van der Waals and Electrostatic Parameters for Intermolecular Interactions. *J. Comp. Chem* 17:520-552.

19. Halgren,T.A. 1996c. Merck Molecular Force Field. III. Molecular Geometrics and Vibrational Frequencies for MMFF94. *J. Comp. Chem* 17:553-586.

20. Halgren,T.A. 1996d. Merck Molecular Force Field. V. Extension of MMFF94 using Experimental Data, Additional Computational Data and Empirical Rules. *J. Comp. Chem* 17:616-641.

21. Halgren,T.A. and R.B.Nachbar. 1996. Merck Molecular Force Field. IV. Conformational Energies and Geometries. *J. Comp. Chem* 17:587-615.

22. Harnadek,G.J., E.A.Ries, D.G.Tse, J.S.Fitz, and D.Njus. 1992. Electron transfer in chromaffin-vesicle ghosts containing peroxidase. *Biochim. Biophys. Acta* 1135:280-286.

23. Hemminga,M.A., J.C.Sanders, and R.B.Spruijt. 1992. Spectroscopy of lipid-protein interactions: structural aspects of two different forms of the coat protein of bacteriophage M13 incorporated in model membranes. *Prog. Lipid Res.* 31:301-333.

24. Horemans,N., H.Asard, and R.J.Caubergs. 1994. The Role of Ascorbate Free Radical as an Electron Acceptor to Cytochrome b-Mediated Trans-Plasma Membrane Electron Transport in Higher Plants. *Plant Physiol* 104:1455-1458.

25. Kay,B.K., N.B.Adey, Y.-S.He, J.P.Manfredi, A.H.Mataragon, and D.M.Fowlkes. 1993. An M13 phage library displaying 38-amino-acid peptides as a source of novel sequences with affinity to selected targets. *Gene* 128:59-65.

26. Kelley,P.M., V.Jalukar, and D.Njus. 1990. Rate of electron transfer between cytochrome b561 and extravesicular ascorbic acid. *J. Biol. Chem.* 265:19409-19413.

27. Kent,U.M. and P.J.Fleming. 1987. Purified cytochrome b561 catalyzes transmembrane electron transfer for dopamine beta-hydroxylase and peptidyl glycine alpha-amidating monooxygenase activities in reconstituted systems. *J. Biol. Chem.* 262:8174-8178.

28. Kent,U.M. and P.J.Fleming. 1990. Cytochrome b561 is fatty acylated and oriented in the chromaffin granule membrane with its carboxyl terminus cytoplasmically exposed. *J. Biol. Chem.* 265:16422-16427.

29. Khan,A.R. and C.M.Deber. 1995. An engineered disulfide bridge in the transmembrane region of phage M13 coat protein stabilizes the alpha-helical dimer. *Biochem. Biophys. Res. Commun.* 206:230-237.

30. Killian JA and von Heijne G. 2000. How proteins adapt to a membrane water interface. *Trends Biochem Sci* 25:429-434.

31. Kipp,B.H., P.M.Kelley, and D.Njus. 2001. Evidence for an essential histidine residue in the ascorbate-binding site of cytochrome b561. *Biochemistry* 40:3931-3937.

32. Kobayashi,K., M.Tsubaki, and S.Tagawa. 1998. Distinct roles of two heme centers for transmembrane electron transfer in cytochrome b561 from bovine adrenal chromaffin vesicles as revealed by pulse radiolysis. *J. Biol. Chem.* 273:16038-16042.

33. Koradi,R., M.Billeter, and K.Wüthrich. 1996. MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14:51-55.

34. Krogh A, Larsson B, von Heijne G, and Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305:567-580.

35. Lancester CRD, Kroger A, Auer M, and Michel H. 1999. Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 A resolution. *Nature* 402:377-385.

36. Li,Z., M.Glibowicka, C.Joensson, and C.M.Deber. 1993. Conformational states of mutant M13 coat proteins are regulated by transmembrane residues. *J. Biol. Chem.* 268:4584-4587.

37. Marassi,F.M. and S.J.Opella. 2003. Simultaneous assignment and structure determination of a membrane protein from NMR orientational restraints. *Protein Sci* 12:403-411.

38. Marsh,D. 1981. Electron spin resonance: Spin labels. *In* Membrane spectroscopy. Molecular biology, biochemistry and biophysics. E.Grell, editor. Springer-Verlag, New York. 51-142.

39. Marvin,D.A. 1998. Filamentous phage structure, infection and assembly. *Curr. Opin. Struct. Biol.* 8:150-158.

40. Marvin,D.A., R.D.Hale, C.Nave, and M.H.Citterich. 1994. Molecular model and structural comparisons of native and mutant class I filamentous bacteriophage Ff (fd, f1, M13), If1 and Ike. *J. Mol. Biol.* 111:487-507.

41. McDonnell,P.A., K.Shon, Y.Kim, and S.J.Opella. 1993. fd coat protein structure in membrane environments. *J. Mol. Biol.* 233:447-463.

42. McGuffin LJ, Bryson K, and Jones DT. 2000. The PSIPRED protein structure prediction server. *Bioinformatics*. 16:404-405.

43. McKie,A.T., D.Barrow, G.O.Latunde-Dada, A.Rolfs, G.Sager, E.Mudaly, M.Mudaly, C.Richardson, D.Barlow, A.Bomford, T.J.Peters, K.B.Raja, S.Shirali, M.A.Hediger, F.Farzaneh, and R.J.Simpson. 2001. An iron-regulated ferric reductase associated with the absorption of dietary iron. *Science* 291:1755-1759.

44. Melnyk,R.A., S.Kim, A.R.Curran, D.M.Engelman, J.U.Bowie, and C.M.Deber. 2004. The affinity of GxxxG motifs in transmembrane helix-helix interactions is modulated by long-range communication. *J. Biol. Chem.* 279:16591-16597.

45. Monne,M., I.Nilsson, M.Johansson, N.Elmhed, and G.von Heijne. 1998. Positively and negatively charged residues have different effects on the position in the membrane of a model transmembrane helix. *J. Mol. Biol.* 284:1177-1183.

46. Nagle,J.F. and S.Tristram-Nagle. 2000. Structure of lipid bilayers. *Biochim. Biophys. Acta* 1469:159-195.

47. Njus,D., J.Knoth, C.Cook, and P.M.Kelly. 1983. Electron transfer across the chromaffin granule membrane. *J. Biol. Chem.* 258:27-30.

48. Njus,D., M.Wigle, P.M.Kelley, B.H.Kipp, and H.B.Schlegel. 2001. Mechanism of ascorbic acid oxidation by cytochrome b(561). *Biochemistry* 40:11905-11911.

49. Okuyama,E., R.Yamamoto, Y.Ichikawa, and M.Tsubaki. 1998. Structural basis for the electron transfer across the chromaffin vesicle membranes catalyzed by cytochrome b561: analyses of cDNA nucleotide sequences and visible absorption spectra. *Biochim. Biophys. Acta* 1383:269-278.

50. Papavoine,C.H., B.E.Christiaans, R.H.Folmer, R.N.Konings, and C.W.Hilbers. 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.

51. Papavoine,C.H., R.N.Konings, C.W.Hilbers, and F.J.van de Ven. 1994. Location of M13 coat protein in sodium dodecyl sulfate micelles as determined by NMR. *Biochemistry* 33:12990-12997.

52. Pilpel Y, Ben-Tal N, and Lancet D. 1999. kPROT: a knowledge-based scale for the propensity of residue orientation in transmembrane segments. Application to membrane protein structure prediction. *J. Mol. Biol.* 294:921-935.

53. Ponting,C.P. 2001. Domain homologues of dopamine beta-hydroxylase and ferric reductase: roles for iron metabolism in neurodegenerative disorders? *Hum. Mol. Genet.* 10:1853-1858.

54. Rost,B. 1995. Transmembrane helices predicted at 95-percent accuracy. *Protein Sci* 4:521-533.

55. Srivastava,M., L.T.Duong, and P.J.Fleming. 1984. Cytochrome b561 catalyzes transmembrane electron transfer. *J. Biol. Chem.* 259:8072-8075.

56. Stevens,T.J. and I.T.Arkin. 2001. Substitution rates in a-helical transmembrane proteins. *Protein Sci* 10:2507-2517.

57. Stopar,D., K.A.Jansen, T.Pali, D.Marsh, and M.A.Hemminga. 1997a. Membrane location of spin-labeled M13 major coat protein mutants determined by paramagnetic relaxation agents. *Biochemistry* 36:8261-8268.

58. Stopar,D., R.B.Spruijt, C.J.Wolfs, and M.A.Hemminga. 1996. Local dynamics of the M13 major coat protein in different membrane-mimicking systems. *Biochemistry* 35:15467-15473.

59. Stopar,D., R.B.Spruijt, C.J.Wolfs, and M.A.Hemminga. 1997b. In situ aggregational state of M13 bacteriophage major coat protein in sodium cholate and lipid bilayers. *Biochemistry* 36:12268-12275.

60. Stopar,D., R.B.Spruijt, C.J.Wolfs, and M.A.Hemminga. 2003. Protein-lipid interactions of bacteriophage M13 major coat protein. *Biochim. Biophys. Acta* 1611:5-15.

61. Sugimoto,K., H.Sugisaki, T.Okamoto, and M.Takanami. 1997. Studies on bacteriophage fd DNA. The sequence of messenger RNA for the major coat protein gene. *J. Mol. Biol.* 111:487-507.

62. Surles,M.C., J.S.Richardson, D.C.Richardson, and F.P.Brooks, Jr. 1994. Sculpting proteins interactively: Continual energy minimization embedded in a graphical modeling system. *Protein Sci.* 3:198-210.

63. Takeuchi,F., K.Kobayashi, S.Tagawa, and M.Tsubaki. 2001. Ascorbate inhibits the carbethoxylation of two histidyl and one tyrosyl residues indispensable for the transmembrane electron transfer reaction of cytochrome b561. *Biochemistry* 40:4067-4076.

64. Trost,P., A.Berczi, F.Sparla, G.Sponza, B.Marzadori, H.Asard, and P.Pupillo. 2000. Purification of cytochrome b-561 from bean hypocotyls plasma membrane. Evidence for the presence of two heme centers. *Biochim. Biophys. Acta* 1468:1-5.

65. Tsubaki,M., K.Kobayashi, T.Ichise, F.Takeuchi, and S.Tagawa. 2000. Diethyl pyrocarbonate modification abolishes fast electron accepting ability of cytochrome b561 from ascorbate but does not influence electron donation to monodehydroascorbate radical: identification of the modification sites by mass spectrometric analysis. *Biochemistry* 39:3276-3284.

66. Tsubaki,M., M.Nakayama, E.Okuyama, Y.Ichikawa, and H.Hori. 1997. Existence of two heme B centers in cytochrome b561 from bovine adrenal

chromaffin vesicles as revealed by a new purification procedure and EPR spectroscopy. *J. Biol. Chem.* 272:23206-23210.

67. Turyna,B., A.Osyczka, A.Kostrzewa, W.Blicharski, J.J.Enghild, and W.Froncisz. 1998. Preparation and electron paramagnetic resonance characterization of spin labeled monoderivatives of horse cytochrome c. *Biochim. Biophys. Acta* 1386:50-58.

68. Wolkers,W.F., P.I.Haris, A.M.Pistorius, D.Chapman, and M.A.Hemminga. 1995. FT-IR spectroscopy of the major coat protein of M13 and Pf1 in the phage and reconstituted into phospholipid systems. *Biochemistry* 34:7825-7833.

69. Wolkers,W.F., R.B.Spruijt, A.Kaan, R.N.Konings, and M.A.Hemminga. 1997. Conventional and saturation-transfer EPR of spin-labeled mutant bacteriophage M13 coat protein in phospholipid bilayers. *Biochim. Biophys. Acta* 1327:5-16.

70. Yuan,J., A.Amend, J.Borkowski, R.DeMarco, W.Bailey, Y.Liu, G.Xie, and R.Blevins. 1999. MULTICLUSTAL: a systematic method for surveying Clustal W alignment parameters. *Bioinformatics.* 15:862-863.