Protoplasma (2003) 221: 31–40
DOI 10.1007/s00709-002-0065-0

PROTOPLASMA
Printed in Austria

Structure prediction for the di-heme cytochrome b_{561} protein family

Denys Bashtovyy[1], **Alajos Bérczi**[1], **Han Asard**[2], and **Tibor Páli**[1,*]

[1] Institute of Biophysics, Biological Research Centre, Szeged
[2] Beadle Center for Genetic Research, University of Nebraska-Lincoln, Lincoln, Nebraska

Received May 12, 2002; accepted September 20, 2002; published online May 21, 2003

Summary. Atomic models possessing the common structural features identified for the cytochrome b_{561} (cyt b_{561}) protein family are presented. A detailed and extensive sequence analysis was performed in order to identify and characterize protein sequences in this family of transmembrane electron transport proteins. According to transmembrane helix predictions, all sequences contain 6 transmembrane helices of which 2–6 are located closely in the same regions of the 26 sequences in the alignment. A mammalian (*Homo sapiens*) and a plant (*Arabidopsis thaliana*) sequence were selected to build 3-dimensional structures at atomic detail using molecular modeling tools. The main structural constraints included the 2 pairs of heme-ligating His residues that are fully conserved in the family and the lipid-facing sides of the helices, which were also very well conserved. The current paper proposes 3-dimensional structures which to our knowledge are the first ones for any protein in the cyt b_{561} family. The highly conserved His residues anchoring the two hemes on the cytoplasmic side and noncytoplasmic side of the membrane are in all proteins located in the transmembrane helices 2, 4 and 3, 5, respectively. Several highly conserved amino acids with aromatic side chain are identified between the two heme ligation sites. These residues may constitute a putative transmembrane electron transport pathway. The present study demonstrates that the structural features in the cyt b_{561} family are well conserved at both the sequence and the protein level. The central 4-helix core represents a transmembrane electron transfer architecture that is highly conserved in eukaryotic species.

Keywords: Cytochrome b_{561}; Structure prediction; Heme protein; Protein family; Ascorbic acid; Electron transport.

Abbreviations: Asc ascorbate; cyt cytochrome; MDA monodehydroascorbate; TM transmembrane; 3-D 3-dimensional.

Introduction

The high-redox-potential *b*-type cytochrome (cyt 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, Kelley and Njus 1986, Kent and Fleming 1987). In the past decades, evidence accumulated for the presence of a similar Asc-reducible cyt 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 cyt b_{561} have recently been identified in a large number of plant species (Asard et al. 2000). The mammalian and predicted plant cyt 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 cyt b_{561} is yet to be elucidated. The first evidence for the existence of the cyt 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 cyt b_{561} in these cells pos-

* Correspondence and reprints: Institute of Biophysics, Biological Research Centre, P.O. Box 521, 6701 Szeged, Hungary.
E-mail: tpali@nucleus.szbk.u-szeged.hu

sesses ferric reductase activity. A similar function was proposed for cyt 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 (Kelley and Njus 1986, 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 cyt b_{561} proteins have not yet been crystallized, atomic-detail structural data about these proteins are 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 cyt b_{561} (Trost et al. 2000; Bérczi et al. 2001, 2003) may 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 cyt b_{561} protein family may be highly useful. The objectives of the present work were to identify new structural similarities in the cyt b_{561} family and to build 3-dimensional (3-D) atomic models for representative plant and mammalian cyt b_{561} proteins. In the absence of direct structural data, this was done in three steps: (1) exploring the structural features of the family via sequence alignment and predicting the TM helices, (2) establishing the intra- and extracellular topology and the lipid-facing propensities of the TM helices, and (3) building 3-D atomic models that satisfy all known and newly explored structural constraints. Although we used most recent modeling tools and we built structures, our strategy turned out to be essentially similar to that followed by Hägerhäll and Hederstedt (1996). The structural features of the predicted models are discussed in relation to the structural data and the putative electron transport mechanism.

Results

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

Recent studies involving sequence comparisons have suggested a close relationship between cyt b_{561} proteins of the plant and animal kingdom and demonstrated the conservation of a number of structural features (Asard et al. 2000, 2001; McKie et al. 2001; Ponting 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 cyt 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 database searches with default settings (Altschul 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 1 represents the most complete and detailed sequence alignment of the cyt 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

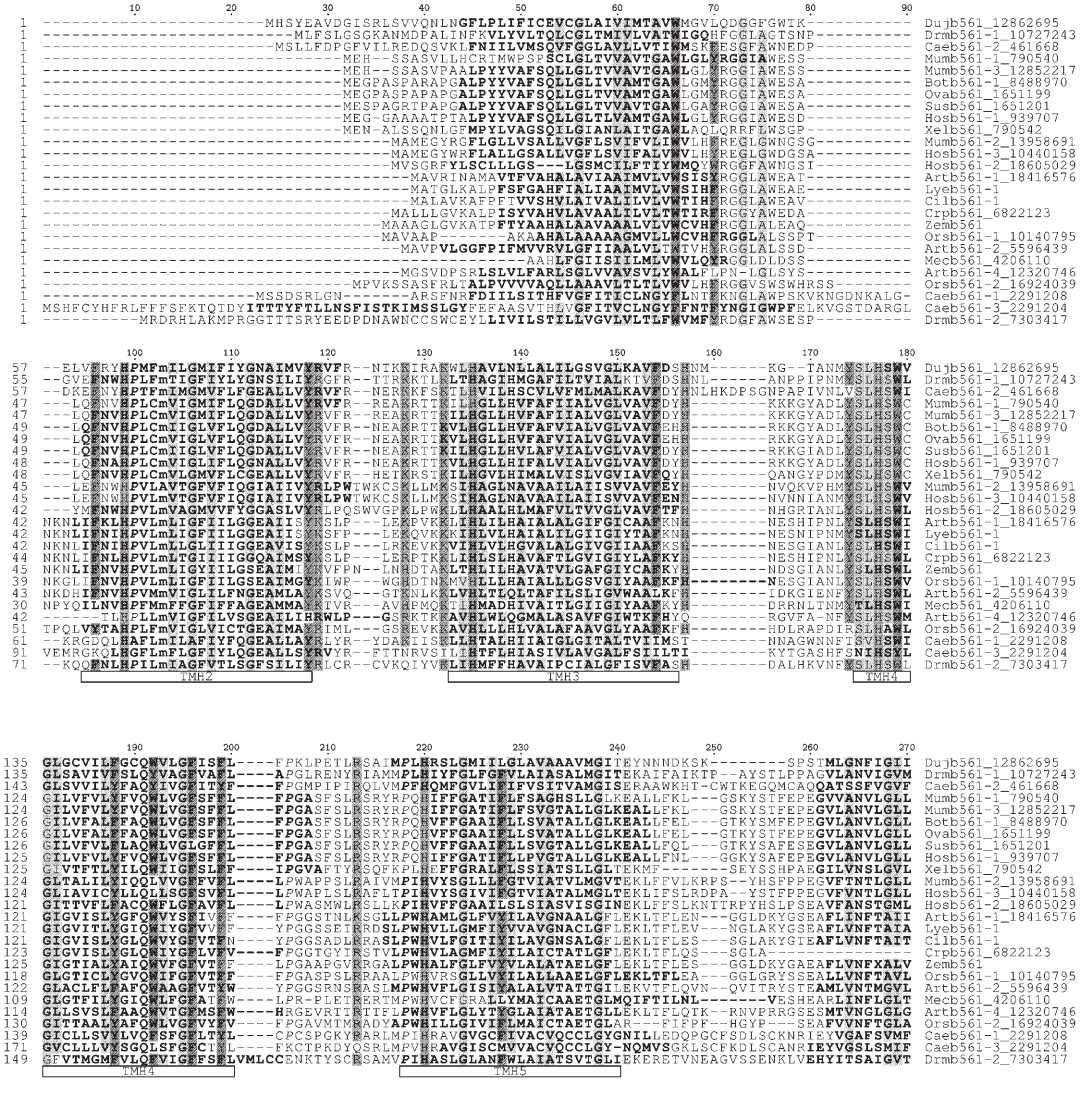

Fig. 1. MULTICLUSTAL alignment of 26 cyt b_{561} sequences. The names and numbers before and after the underscore are names according to the proposed naming convention (Asard et al. 2001) and GenBank identifiers (Benson et al. 2002), respectively. Highlighting is performed with TeXshade (Beitz 2000) above 85% threshold conservation of residue properties and according to the chemical mode. Concerning font styles and gray 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

with the 6 TM helices (Fig. 1, boldface letters). These were predicted, independently from the alignment, for each sequence by TMHMM (version 2.0) (Krogh et al. 2001, Möller et al. 2001). For one sequence (Caeb561-3) there was a seventh TM helix predicted at the N-terminal end of the protein (Fig. 1). 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 His residues that are 100% conserved. These 4 histidines are located in all sequences on TMH2 (His99) and TMH4 (His177) and on TMH3 (His135) and TMH5 (His220) in a pairwise manner, ideal to anchor two hemes on the noncytoplasmic and cytoplasmic side of membrane, respectively, as proposed earlier (e.g., 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 His residues. TMH4 appears to be richest in conserved aromatic residues. Notable is the presence of highly conserved basic residues located between TMH2 and TMH3 at positions 119, 128, and 132. The TM architecture in the alignment suggests cytoplasmic orientation for both the N and C termini. It should be noted, that TMHMM (Krogh et al. 2001) predicted an opposite orientation for the sequences Artb561-1, Lyeb561-1, Zemb561, Orsb561-1, and Mecb561, but MEMSAT 2 (McGuffin 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

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 cyt b_{561} proteins consisting of 4 TM helices (TMH2 to -5). These TM helices define highly conserved membrane-spanning regions and the fully conserved His 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. 1). 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 and von Heijne 2000) were also taken into account. The TM helices TMH2 to -5 are indicated with horizontal bars under the sequences in Fig. 1. 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 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 regions, i.e., 4 by 26 sequence segments. The radial thin solid lines on Fig. 2 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 origi-

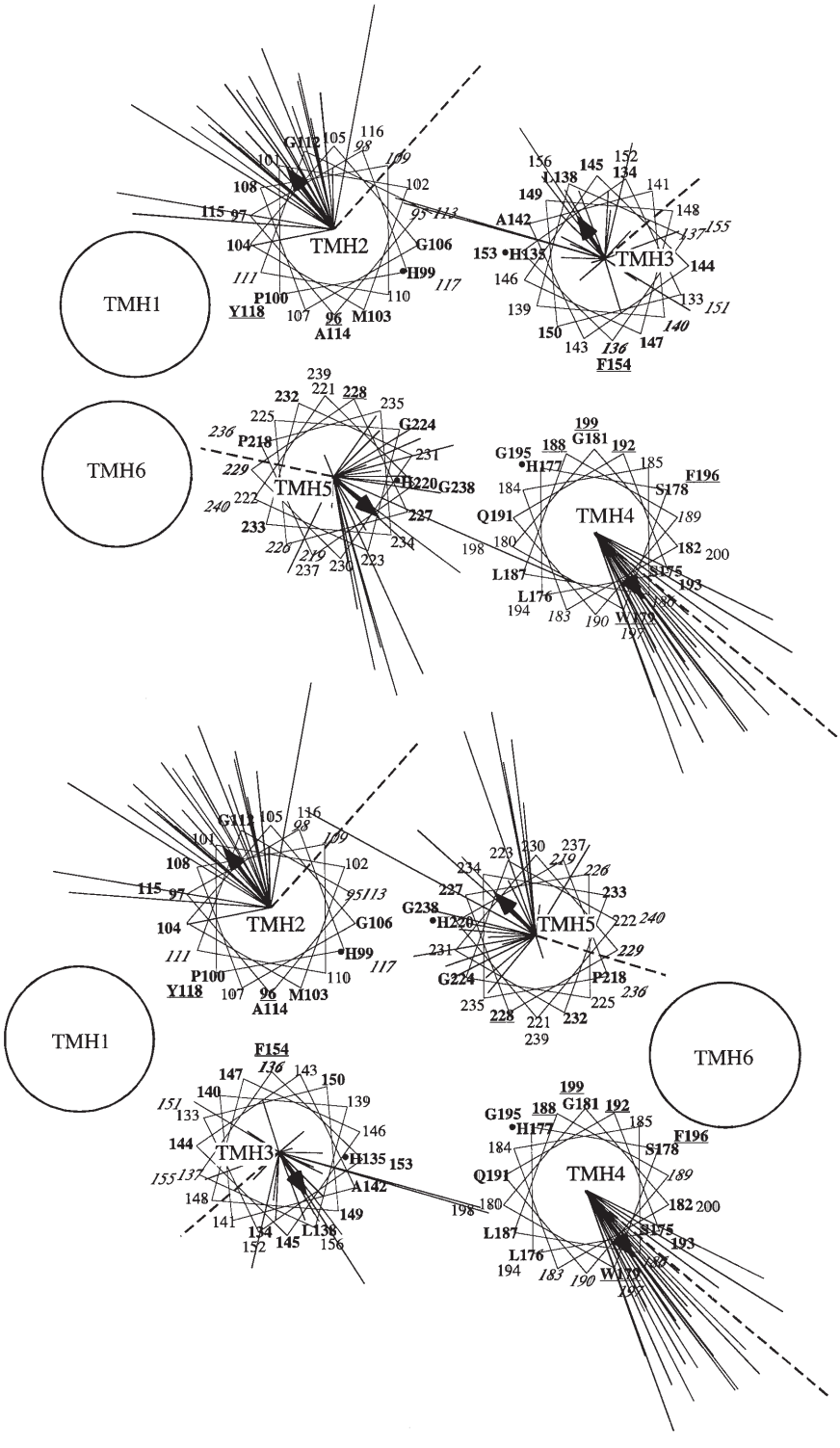

Fig. 2. Lipid-facing propensity of the individual and consensus TM helices. kPROT vectors (Pilpel et al. 1999) for the individual proteins are shown with radial thin lines (not all being visible) in the helical wheels of the consensus transmembrane helices TMH2 to -5 defined by the sequence regions 95–118(i), 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. 1). 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 His residues are indicated with bullets. Underlined numbers indicate sequence positions of highly conserved aromatic residues

nates 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 cyt 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 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. 2, 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. 2 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

Model structures were built for the 4-helix core of representatives of a mammalian (*Homo sapiens*) and plant (*Arabidopsis thaliana*) cyt b_{561} sequence, Hosb561-1 (Fig. 3) and Artb561-1 (Fig. 3). The primary constraints for the models were: (1) the predicted (and not consensus) TM helix sequence regions (Fig. 1), (2) the requirement for ligation of the two hemes by the fully conserved 4 His residues (Fig. 2), and (3) the ability to satisfy the lipid-facing propensities of the individual helices (Fig. 2) as far as possible. Residues with known preferential membrane location were also considered as additional constraints (Killian and von Heijne 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 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. 2) and a rotation of TMH2 and TMH4 to bring them closer together would destroy the optimal helix orientations. In addition, Kent and Fleming (1990) found that all cysteines were in the free sulfhydryl form in chromaffin granule cyt b_{561}. Both topologies still leave some "freedom" in orienting the helix axes relative to the membrane normal and along their long axis (Fig. 2). These uncertainties can be estimated to be below ca. 25° and 30°, respectively (see Fig. 2). 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. 2). 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. 2, bottom) is somewhat more likely than the circular one (Fig. 2, 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. 2) and to obtain an approximate distance of 1.2 nm between the C^{α} atoms of the pairs of the fully conserved heme-ligating His residues. This His-to-His distance was determined after an inspection of coordinates of 34 protein structures in the Brookhaven Protein Data Bank which contained b-hemes ligated by the N^{ε} atoms of His 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 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

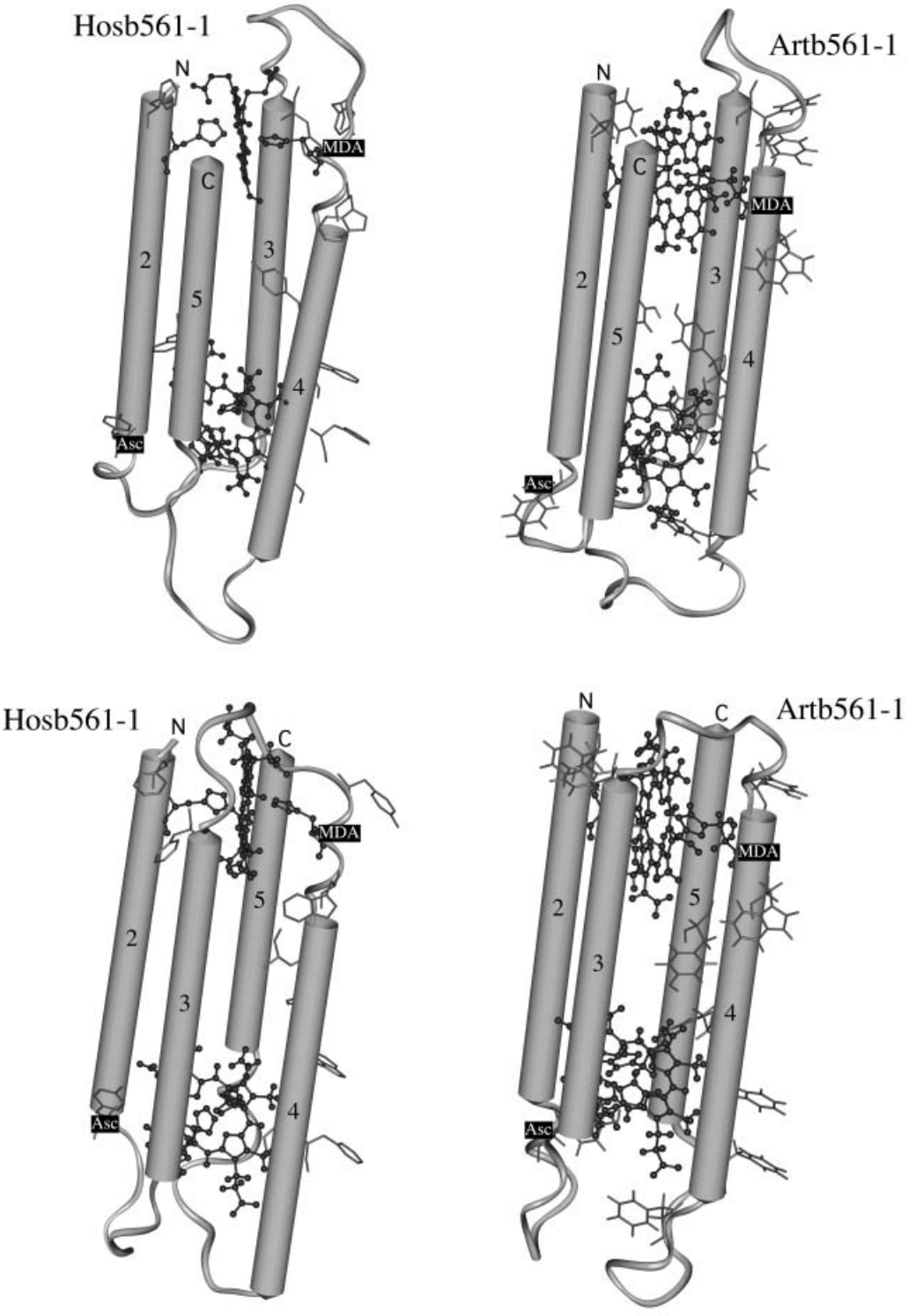

Fig. 3. Predicted 3-D structures for the *Homo sapiens* and *Arabidopsis thaliana* cyt b_{561} sequences, Hosb561-1 and Artb561-1, respectively. Two structures are shown for both sequences, namely, with clockwise and counterclockwise helix topology (top and bottom, respectively, as defined in Fig. 2). 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 gray thin bar presentation. The two pairs of fully conserved His residues (His139 and His157) and the hemes ligated by them are shown in dark gray ball-and-stick presentation. The N- and C-termini of helices 2 and 5 are indicated with N and C, respectively. The TM helix numbers are indicated on the cylinders. The cytoplasmic side is below each structure. The putative MDA and Asc binding sites are indicated with the corresponding text boxes. The helix axes are closely perpendicular to the membrane plane

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^{ε} atoms of the pairs of the heme-ligating His residues were fixed at about 0.4 nm. This distance was also determined from inspecting the relevant b-heme-containing protein structures. The 2 by 2 structures of the 4-helix cores for the two proteins are shown in Fig. 3 together with the hemes that were inserted manually into the structures after optimization. Considering the high level of conservation of structural features in the cyt 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

The present sequence alignment including plant and animal members of the cyt b_{561} family (Fig. 1) supports the main conclusions on the conservation of functional elements from recent analyses on a smaller subset of the cyt b_{561} family (Asard et al. 2000, 2001). Together with the structures presented in Figs. 2 and 3, 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 cyt 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. 1). 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 His 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 a 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. 3 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 cyt b_{561} and chromaffin granule cyt b_{561} (Kelley and Njus 1986, 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 His residues (Tsubaki et al. 2000).

The position of TMH6 in the sequences is well conserved (Fig. 1), 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 cyt 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 cyt b_{561} (McKie et al. 2001). In that study, antibodies directed against C-terminal peptides inhibited cyt 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 cyt 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 His residues (Figs. 2 and 3). 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 cyt b_{561} proteins. It is interesting that the location of the fully conserved His residues in TMH3 and TMH5 are rather close to the lipid-facing sides in these "mean" topologies (see bullets and solid vectors in Fig. 2). 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 His139 and His157 are also well but not fully conserved. It has been argued that His139 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 His residues also supports this interpretation (Tsubaki et al. 2000). His 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 His139 and His157 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. 1). The high structural similarity between the plant and animal cyt 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 et al. 1999; PDB entry 1QLA), is most relevant to the present study, although its sequence is very different from that of the cyt b_{561} family (hence not shown in the alignment). Similarly to our models, two pairs of His 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 cyt 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. 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 latter one is particularly interesting because of the large distance between the heme centers, 2.2–2.9 nm in the present models, supported by an early estimate of 2.1–3.2 nm for a cyt b_{561} (Esposti et al. 1989). The present models will be further refined as new structural data emerge in the future.

Acknowledgments

This work was supported by grants from the Flemish (BIL 99/48) and Hungarian (OMFB B-4/99) governments, the Volkswagen-Stiftung, and the Hungarian National Science Fund (OTKA T-029458 and T-034488). Work by H.A. was a contribution of the University of Nebraska, Agricultural Research Division, Lincoln, Nebraska. Journal Series nr. 13700.

References

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

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389–3402

Asada A, Kusakawa T, Orii H, Agata K, Watanabe K, Tsubaki M (2002) Planarian cytochrome b561: conservation of a six transmembrane structure and localization along the central and peripheral nervous system. J Biochem 131: 175–182

Asard H, Horemans N, Caubergs RJ (1992) Transmembrane electron transport in ascorbate-loaded plasma membrane vesicles from higher plants involves a b-type cytochrome. FEBS Lett 306: 143–146

– Terol-Alcayde J, Preger V, Del Favero J, Verelst W, Sparla F, Perez-Alonso M, Trost P (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

– Kapila J, Verelst W, Bérczi A (2001) Higher-plant plasma membrane cytochrome b_{561}: a protein in search of a function. Protoplasma 217: 77–93

Beitz, E (2000) TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. Bioinformatics 16: 135–139

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL (2002) GenBank. Nucleic Acids Res 30: 17–20

Bérczi A, Lüthje S, Asard H (2001) *b*-Type cytochromes in plasma membranes of *Phaseolus vulgaris* hypocotyls, *Arabidopsis thaliana* leaves, and *Zea mays* roots. Protoplasma 217: 50–55

– Caubergs RJ, Asard H (2003) Partial purification and characterization of an ascorbate-reducible *b*-type cytochrome from the plasma membrane of *Arabidopsis thaliana* leaves. Protoplasma 221: 47–56

Casimiro DR, Richards JH, Winkler JR, Gray HB (1993) Electron transfer in ruthenium-modified cytochromes c: sigma-tunneling

pathways through aromatic residues. J Phys Chem 97: 13073–13077

Cheung MS, Daizadeh I, Stuchebrukov AA, Heelis PF (1999) Pathways of electron transfer in *Escherichia coli* DNA photolase: Trp306 to FADH. Biophys J 76: 1241–1249

Esposti MD, Kamensky YA, Arutjunjan AM, Konstantinov AA (1989) A model for the molecular organization of cytochrome β-561 in chromaffin granule membranes. FEBS Lett 254: 74–78

Farver O, Skov LK, Young S, Bonander N, Karlsson BG, Vanngard T, Pecht I (1997) Aromatic residues may enhance intramolecular electron transfer in azurin. J Am Chem Soc 119: 5453–5454

Hägerhäll C, Hederstedt L (1996) A structural model for the membrane-integral domain succinate:quinone oxidoreductases. FEBS Lett 389: 25–31

Harnadek GJ, Ries EA, Tse DG, Fitz JS, Njus D (1992) Electron transfer in chromaffin-vesicle ghosts containing peroxidase. Biochim Biophys Acta 1135: 280–286

Horemans N, Asard H, Caubergs R (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

Jones DT, Taylor WR, Thornton JM (1994) A model recognition approach to the prediction of all-helical membrane protein structure and topology. Biochemistry 33: 3038–3049

Karlin S, Ghandour G (1985) Multiple-alphabet amino acid sequence comparisons of the immunoglobulin kappa-chain constant domain. Proc Natl Acad Sci USA 82: 8597–8601

Kelley PM, Njus D (1986) Cytochrome b561 spectral changes associated with electron transfer in chromaffin-vesicle ghosts. J Biol Chem 261: 6429–6432

Kent UM, Fleming PJ (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

– – (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

Killian JA, von Heijne G (2000) How proteins adapt to a membrane-water interface. Trends Biochem Sci 25: 429–434

Kipp BH, Kelley PM, Njus D (2001) Evidence for an essential histidine residue in the ascorbate-binding site of cytochrome b561. Biochemistry 40: 3931–3937

Kobayashi K, Tsubaki M, Tagawa S (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

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

Lancester CRD, Kröger A, Auer M, Michel H (1999) Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution. Nature 402: 377–385

McGuffin LJ, Bryson K, Jones DT (2000) The PSIPRED protein structure prediction server. Bioinformatics 16: 404–405

McKie AT, Barrow D, Latunde-Dada GO, Rolfs A, Sager G, Mudaly E, Mudaly M, Richardson C, Barlow D, Bomford A, Peters TJ, Raja KB, Shirali S, Hediger MA, Farzaneh F, Simpson RJ (2001) An iron-regulated ferric reductase associated with the absorption of dietary iron. Science 291: 1755–1759

Möller S, Croning MDR, Apweiler R (2001) Evaluation of methods for the prediction of membrane spanning regions. Bioinformatics 17: 646–653

Njus D, Knoth J, Cook C, Kelley PM (1983) Electron transfer across the chromaffin granule membrane. J Biol Chem 258: 27–30

– Wigle M, Kelley PM, Kipp BH, Schlegel HB (2001) Mechanism of ascorbic acid oxidation by cytochrome b561. Biochemistry 40: 11905–11911

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

Pilpel Y, Ben-Tal N, 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

Ponting CP (2001) Domain homologues of dopamine beta-hydroxylase and ferric reductase: roles for iron metabolism in neurodegenerative disorders? Human Mol Genet 10: 1853–1858

Srivastava M (1986) *Xenopus* cytochrome b561: molecular confirmation of a general five transmembrane structure and developmental regulation at the gastrula stage. DNA Cell Biol 15: 1075–1080

– Duong LT, Fleming PJ (1984) Cytochrome b561 catalyzes transmembrane electron transfer. J Biol Chem 259: 8072–8075

Takeuchi F, Kobayashi K, Tagawa S, Tsubaki M (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

Trost P, Bérczi A, Sparla F, Sponza G, Marzadori B, Asard H, Pupillo P (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

Tsubaki M, Nakayama M, Okuyama E, Ichikawa Y, Hori H (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

– Kobayashi K, Ichise T, Takeuchi F, Tagawa S (2000) Diethylpyrocarbonate 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

Yuan J, Amend A, Borkowski J, DelMarco R, Bailey W, Liu Y, Xie G, Blevins R (1999) MULTICLUSTAL: a systematic method for surveying Clustal W alignment parameters. Bioinformatics 15: 862–863