Original article

Homology modeling and ligand interaction of Cytochrome b protein

Rohith Kumar Anugolu, Shravan Kumar Gunda, Rajiv Varman Kakarla

Bioinformatics Division, Osmania University, Hyderabad-500007, Andhra Pradesh, India.

Abstract

Cytochrome b is a component of respiratory chain complex III, also known as the bc1 complex or ubiquinol-cytochrome c reductase. These complexes are involved in electron transport and the generation of ATP and thus play a vital role in the cell. The cytochrome bc1 complex is a membrane-bound enzyme that catalyses the transfer of electrons from ubiquinol to cytochrome c, coupling this process to the translocation of protons across the inner mitochondrial membrane. The function of proteins is generally determined by its three-dimensional (3D) structure. Thus, it would be useful to know the 3D structure of the thousands of protein sequences that are emerging from many genome projects. Structural studies on biomolecules have changed our perception of the biological world in the last twenty years. A number of efforts have been made on structure prediction. One such technique that has found a wide appreciation is Homology modeling which aims at predicting the 3D structure of biomolecules, relying heavily on resources such as pattern/function and sequence. However, the three-dimensional structure of cytochrome b subunit protein (Accession number C4PKA1) from Homo sapiens remains unknown. In the present study, effort was made to generate the three-dimensional (3D) structure of the cytochrome b based on available template (1BGY) structural homologues from Protein Data Bank and the model validated with standard parameters (Procheck). With the predicted model, the ligand was subjected to docking study using FlexX docking tool. Flexible docking was carried out with the HEM - Protoporphyrin X [Heme] as ligand; which was found to bind at His267, Ile268 and Val343 residues on given generated protein. We therefore concluded that the above mentioned residues were the key residue sites for ligand binding. The predicted model showed better results than the template structure with 0% disallowed regions. This study will be used in broad screening of the protein in the respiratory process and can be further implemented in future drug designing.

Key words: Homology modeling, cytochrome b, Modeller9v7, FlexX, PROCHECK

© 2011 Deccan College of Medical Sciences. All rights reserved.
Cytocromes were initially described in 1884 by MacMunn as respiratory pigments (myohematin or histohematin). In the 1920s, Keilin rediscovered these respiratory pigments and named them the cytochromes, or "cellular pigments", and classified these heme proteins, on the basis of the position of their lowest energy absorption band in the reduced state, as cytochromes a (605 nm), b (~565 nm), and c (550 nm). The UV-visible spectroscopic signatures of hemes are still used to identify heme type from the reduced bis-pyridine-ligated state, i.e., the pyridine hemochrome method.

In the mitochondrion of eukaryotes and in aerobic prokaryotes, cytochrome b is a component of respiratory chain complex III, also known as the bc1 complex or ubiquinol-cytochrome c reductase. In plant chloroplasts and cyanobacteria, there is an analogous protein, cytochrome b6, a component of the plastocyanine-plastocyanin reductase, also known as the b6f complex. These complexes are involved in electron transport and the generation of ATP and thus play a vital role in the cell. The cytochrome bc1 complex is a membrane-bound enzyme that catalyses the transfer of electrons from ubiquinol to cytochrome c, coupling this process to the translocation of protons across the inner mitochondrial membrane. In higher eukaryotes, the enzyme complex is composed of 11 polypeptide subunits. One subunit, cytochrome b, is encoded by the mitochondrial genome (mtDNA) whilst the others are encoded by the nucleus and synthesized on cytoplasmic ribosomes prior to being imported into mitochondria and assembled into a functional complex. As a hydrophobic, integral membrane protein consisting of eight transmembrane helices, Cytochrome b is fundamental for the assembly and function of complex III, and together with cytochrome c1 and the iron-sulfur protein (ISP) it forms the catalytic core of the enzyme.

The heme group is a highly-conjugated ring system (which allows its electrons to be very mobile) surrounding a metal ion, which readily interact between the oxidation states. For many cytochromes, the metal ion present is that of iron, which interconverts between Fe\(^{2+}\) (reduced) and Fe\(^{3+}\) (oxidised) states (electron-transfer processes) or between Fe\(^{2+}\) (reduced) and Fe\(^{3+}\) (formal, oxidised) states (oxidative processes). Cytochromes are, thus, capable of performing oxidation and reduction. Because the cytochromes (as well as other complexes) are held within membranes in an organized way, the redox reactions are carried out in the proper sequence for maximum efficiency.

Materials and Methods

Sequence Alignment

The FASTA sequence of cytochrome b (Accession number C4PKA1) from Homo sapiens was retrieved from the UniProt database that has 380 amino acids. Comparative modeling usually starts by searching the PDB of known protein structures using the target sequence as the query. This search is generally done by comparing the target sequence with the sequence of each of the structure in the database. The target sequence was searched for similar sequence using the BLAST (Basic Local Alignment Search Tool) against Protein Database (PDB). The BLAST results yielded X-ray structure of 1BGY (cytochrome BC1 complex from bovine) with 79% similarity to our target protein.

Comparative Modeling

The theoretical structure (Fig 1) of cytochrome b (Accession number C4PKA1) from Homo sapiens is generated using MODELLER9v7 by comparative modeling of protein structure prediction. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints. The program was designed to use as many different types of information about the target sequence as possible.

Validation of Cytochrome b Model

PROCHECK

A versatile protein structure analysis program was used in validation of protein structure and models by verifying the parameters like Ramachandran plot quality, peptide bond planarity, Bad non-bonded interactions, main chain hydrogen bond energy, C-alpha chirality and over-all G factor and the side chain parameters like standard deviations of chi1 gauche minus, trans and plus, pooled standard deviations of chi1 with respect to refined structures.

Fig 1. Predicted 3D structure of Human Cytochrome b represented in yellow color (C4PKA1.B99990018) superimposed with the template (1BGY) represented in red color.
**RMSD**

Root Mean Squared Deviation (RMSD) is commonly used to represent the distance between two objects. In a structural sense, this value indicates the degree to which two three-dimensional structures are similar. The lower the value, the more similar the structures are. The RMSD value between the template 1BGY and our model structure was calculated using spdbv program.

**Docking studies**

Docking studies were carried out using the FlexX program interfaced with SYBYL 6.7. The 3D coordinates of the active sites were taken from the cytochrome b model reported as complex with the corresponding HEM-Protoporphyrin X [Heme] ligand (Fig 2). The ligand was preprocessed before docking calculations by giving charges according to the Gasteiger-Hückel method followed by energy minimization with 10,000 iterations of conjugate gradient algorithm using Tripos force field. HEM-Protoporphyrin X [Heme] ligand was docked with the protein with the active site defined within 6.5Å radius around the co-crystallized ligand.

![Image showing docking of HEM-Protoporphyrin X with the predicted cytochrome b protein](image)

**Results**

**Comparative Modeling of Cytochrome b Model**

Tertiary structure of a protein is built by packing of its secondary structure elements to form discrete domains or autonomous folding units. Comparative modeling predicts the 3-D structure of Cytochrome b model a given protein sequence (target) based primarily on its alignment to 1BGY as a template (a determined structure experimentally). The hypothetical protein models created were stored as PDB output file. This latter was visualized by Pymol program.

**Validation of Protein Structures of Cytochrome b Model**

The hypothetical protein models generated were analyzed. Accuracy of the protein model generated was judged by validity report generated by PROCHECK. Parameter comparisons of these proteins were made with well-refined structures that have similar resolution. The main chain parameters plotted are Ramachandran plot quality, peptide bond planarity, Bad non-bonded interactions, main chain hydrogen bond energy, C-alpha chirality and over-all G factor. In the Ramachandran plot analysis, the residues were classified according to their regions in the quadrangle. The Ramachandran map for Cytochrome b is represented in (Fig 3) and the plot statistics (Table 1).

![Ramachandran plot showing the human cytochrome b model. The most favored regions are colored red, additional allowed, generously allowed and disallowed regions are indicated as yellow, light yellow and white fields, respectively.](image)

**Docking**

Based on rigid-body docking by using FlexX, the protein and ligand were analyzed for shape complementary, hydrophobic effects resulting from a decreasing in the solvent accessible surface, and electrostatic interactions. The key amino acid residues within the docking complex model involved in the interaction between human...
cytochrome b and HEM- Protoporphyrin X [Heme] ligand are His267, Ile268 and Val343 (Fig 2). These residues were determined based on intermolecular hydrogen bond lengths of amino acid residues interacted between human cytochrome b and heme ligand. The distance between protein and the ligand appear to be less than 2.6Å. This suggests that hydrogen bonds can be plausibly formed. The docking result indicated that the complex could be stabilized by hydrogen bonding. The Docking interactions as observed with the heme ligand, shows H-bonds with three residues. The residue His267 show only one hydrogen bond with a distance of 1.77Å, and the residue Ile268 showed only one interaction with the protein, showing the interaction distance of 2.22Å, where as residue Val343 shows three hydrogen bond interactions, with an interaction distance of 1.89Å, 2.13Å and 2.13Å. The table 1 explains the docking interactions of the two proteins.

**Table 1.** The number of Hydrogen bonds formed and the catalytic site residues involved in protein-ligand complex.

<table>
<thead>
<tr>
<th>Compound</th>
<th>No. of H-bonds</th>
<th>Interacting residues</th>
<th>Distances (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEM-Protoporphyrin X [Heme]</td>
<td>1</td>
<td>HIS257</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>ILE268</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>VAL343</td>
<td>1.89, 2.13, 2.13</td>
</tr>
</tbody>
</table>

Discussion

Homology protein modeling uses experimentally determined protein structures (templates) to predict the 3-D of another protein that has a similar amino acid sequence (the target). This approach to modeling is possible since a small change in the protein sequence usually results in a small change in its 3D structure. Homology modeling remains the only modeling method that can provide models with a root mean square error lower than 2Å. The FASTA sequence Cytochrome b (Accession number C4PKA1) from Homo sapiens was obtained from UniProt. The primary requirement for reliable homology modeling is a detectable similarity between the sequence of interest (target sequence) and a know structure (template) by protein BLAST query. Based on sequence similarity analysis between target and other proteins known structures showed that our Cytochrome b from Homo sapiens (1BGY) has 79% of amino acid sequence identity with Cytochrome b (Accession number C4PKA1) from Bovine. Practically, at this level of sequence identity, it is good enough to use crystallographic structures of 1BGY as a template in order to obtain high quality alignment for structure prediction by homology modeling.

Homology modeling is currently restricted to protein sequences (targets) that share 30% or more sequence identity to an experimentally solved protein structure template. Under this sequence identity the reliability of the sequence alignment between target and template decreases fast, resulting into significant modeling errors, low accuracy models should still be treated with attention. Medium accuracy models, obtained with a template-target sequence identity of 30–50%, tend to have nearly 85% of their C-a atom within 3.5 Å of the correct position. These models often fit a variety of applications, including the testing of ligand binding states by designing site directed mutants with altered binding capacity, and computational screening of databases listing small molecules for potential lead compounds or inhibitors. Top accuracy models, based on sequence identities more than 50%, usually have structures comparable to 3 Å resolution X-ray structures and can be used for more reliable calculations as (ligand docking, drug design), however sequence identities more than 90% can be used to facilitate a meaningful biophysical description of the active site.

In this study we also checked for φ and ψ torsion angles using the Ramachandran plots. A comparison of the results shows that one of the models generated by Modeller program is more acceptable. The molecular visualization Pymol program was used to manipulate the models based on residue interactions, energy minimization and steric hindrance. The best model predicted by Modeller (Cytochrome b) was used for further analysis by PROCHECK. The Ramachandran plot contributed final values of Cytochrome b i.e. 94.6% of residues in most favored regions and the allowed regions in additional residues of 5.4%. Non-proline residues, non-glycine residue regions were 100.0% and most disallowed regions were 0.0% in the plot. In general, a score close to 100% implies good stereochemical quality of the models.

The overall general similarities and subtle difference among the 3D structure of template 1BGY and predicted model Cytochrome b can be seen from the backbone superposition. As evident from superposition, general folding topology of the structure is similar; however, some structural differences appear between the predicted model and template. These differences are mainly due to insertion and deletions in different loop regions. The RMSD (Root Mean Square Deviation) between predicted model and template is 0.44 Å. The low RMSD between the target and template reflects the presence of strong homology (The lower the value, the more similar the structures are). The z-score indicates overall model quality and measures the deviation of the total energy of
the structure with respect to an energy distribution derived from random conformations. In order to facilitate interpretation of the z-score of the specified protein, its particular value is displayed in a plot that contains the z-scores of all experimentally determined protein chains in current PDB.

Conclusion

Based on the Template structure it is clearly observed that the theoretical structure generated is structurally similar to the template structure which is highly sufficient for the development of specific ligand for Cytochrome b from Homo sapiens. Our model of Cytochrome b is only a predictive, and needs to be confirmed experimentally. This modeled structure can be used to predict the molecular function and the active sites which received less attention in previous reports.

Acknowledgments: None

Conflict of Interest: None

References

14. SYBYL 6.7, Tripos Associates, 1699 South Hanley Road, St. Louis, USA, MO 63144.