Chapter 4: Molecular cloning and characterization of three cytochrome P450 genes from Coleus forskohlii

4.1 Introduction
Cytochrome P450 monooxygenases (CYP) are heme-thiolate proteins, which catalyzes majority of the oxidative reactions, including hydroxylation, dehydration, epoxidation, carbon–carbon bond cleavage and dealkylation. CYPs constitute one of the largest gene family of heme-thiolate proteins widely distributed in a different kingdom of life. About 1% of the total genes in a typical plant genome, encodes for CYPs. Arabidopsis thaliana has 272 CYP genes (246 predicted full-length genes and 26 pseudogene fragments) making it one of the largest gene families in the plant. All plant CYPs have been clustered into ten clans including four multiple-family clans (CYP71, CYP72, CYP85, CYP86) and six single-family clans (CYP51, CYP74, CYP97, CYP710, CYP711, CYP727) (Nelson and Werck-Reichhart 2011). Plant CYPs are known to play an important role in the biosynthesis of diverse phytochemical such as flavonoids, terpenoids, alkaloids, essential oils etc. (Paul R 2005). Cytochrome P450 nomenclature is based on the amino acid sequence identity. A CYP protein with >40% identity is grouped into the same family, and that with >55% identity is grouped into the same subfamily (Mizutani and Ohta 2010). CYPs diversification in plants has been closely associated with the phytochemical diversity in plants (Nelson et al. 2008; Schuler and Werck-Reichhart 2003). In order to face the challenges posed by the environment, plants produce a wide variety of secondary metabolites. (Neilson et al. 2013). CYPs play an important role in the biosynthesis of these secondary metabolites (Bak et al. 2011). Around 7000 CYPs, belonging to 95 families, have been identified in plants (Nelson 2009). CYP710, CYP 51 and CYP97 are the families conserved throughout the plant kingdom and are involved in the biosynthesis of xanthophylls and sterols (Bak et al. 1997; Morikawa et al. 2006; Kim and DellaPenna 2006). CYP71, CYP76, CYP88, CYP701, CYP714, CYP720 and CYP725 families are generally involved in terpenoid metabolism, while members belonging to CYP81 and CYP93 function in flavonoid metabolism (Bak et al. 2011). Medicinal plants are known to produce pharmacologically important secondary metabolites, such as taxol from Taxus baccata L. (Taxaceae), artemisinin from Artemisia annua L. (Asteraceae) and Indole alkaloid (vincristine and vinblastine)
from *Catharanthus roseus* (L.) G. Don (Apocynaceae) (Afendi et al. 2012). CYPs catalyzing key steps in the biosynthesis of these compounds have been identified. Approaches such as differential display, homology based cloning and random sequencing of cDNA library were used in the isolation of CYPs involved in the biosynthesis of taxol (Schoendorf et al. 2001). Some of the examples of CYPs involved in the biosynthesis of pharmacologically important secondary metabolites in medicinal plants are: (1). CYP725A4 catalyzes the conversion of taxa-4(5),11(12)-diene to 5(12)-oxa-3(11)-cyclohexane in *T. baccata* (Jennewein et al. 2004; Rontein et al. 2008). (2). In artemisinin biosynthesis, CYP71AV1 converts oxidized amorphadiene to artemisinic alcohol (Shen et al. 2012; Paddon and Keasling 2014; Teoh et al. 2006). (3). Indole alkaloid (vincristine and vinblastine) biosynthesis in *C. roseus* requires the action of Geraniol 10- hydroxylase (CYP76B6) and a light inducible tabersonine 16-hydroxylase (CYP71D12) (He, Zhao, and Hu 2012; Collu et al. 2001). Compared to the model plant *A. thaliana*, there is relatively little information available regarding CYPs in medicinal plants (Zerbe et al. 2013; Guo et al. 2014).

*C. forskohlii* is an important medicinal plant, producing a pharmacologically important secondary metabolite known as forskolin. Forskolin is a potent activator of adenylyl cyclase. It imparts the antihypertensive property to the plant. Apart from forskolin, other minor diterpenoids such as deacetylforskolin, 1,9-dideoxyforskolin, 1,9-dideoxy-7-deacetylforskolin, 9-deoxyforskolin also accumulate in the roots. (Grayer et al. 2003; Alasbahi and Melzig 2010). Chemical structure of forskolin shows the presence of several hydroxyl moieties, suggesting the involvement of CYPs in the biosynthesis of forskolin and other related diterpenoids. *C. forskohlii* is also known to produce rosmarinic acid, genkwanin (7-O-methylapigenin) and guaiacol glycerin (Alasbahi and Melzig 2010). The biosynthesis of genkwanin likely involves the action of CYP involved in flavonoid metabolism (Jeon, Kim, and Ahn 2009). However, no CYP genes have been identified from this plant. In this study, we have carried out the molecular cloning and characterization of three differentially expressed CYPs from this medicinal herb.
4.2 Results

4.2.1 Differential display study of CYPs from *Coleus forskohlii*

Leaf, stem and root tissues of *C. forskohlii* were used to carry out differential display study of CYPs from *C. forskohlii* (Figure 4.1). Differentially expressed PCR bands were cloned and sequenced. Three CYP gene sequences were identified using BLASTX analysis. These CYPs were named on the basis of sequence identity with other members of various CYP families as; CfCYP93B, CfCYP706C, and CfCYP76. CfCYP93B and CfCYP706C were found to be 100% identical at the nucleotide level to EST identified as CfP450C1and CfP450C6 in chapter 3. The amplicon size of CfCYP93B, CfCYP706C, and CfCYP76 transcript fragment were 360 bp, 363 bp and 334 bp respectively.

4.2.2 Cloning and sequence analysis of CYPs isolated from *C. forskohlii*

4.2.2.1 CfCYP93B

CfCYP93B ESTs isolated from differential display study of CYPs was 360 bp in length. This sequence information was used to design 5’RACE primer for CfCYP93B. RACE-PCR was carried out to obtain the 5’ end of CfCYP93B, giving an amplicon size of 1470 bp. Full-length cDNA of CfCYP93B is 1776 bp in size, with ORF of 1530 bp and having 73 bp 5’UTR and 173 bp 3’UTR. Full-length sequence of CfCYP93B was submitted to NCBI genbank database and accession ID (KF606861) was obtained. CfCYP93B encodes a protein of 509 amino acids, having a molecular weight of 57.542 kDa and theoretical pI 8.49. On BLASTX analysis, CfCYP93B was found to share the maximum identity of 85% with flavone synthase II (*Perilla frutescens* var. *crispa*). The secondary structure of CfCYP93B was predicted using SOPMA (Geourjon and Deléage 1995). The result showed that the CfCYP93B contain 47.94% of α-helices, 4.72 of β-turns, 13.36% of extended strands, and 33.99% of random coils.

CYPs are well conserved in structure but not in sequence. The conserved structure comprises of eleven α helices (labeled A-K) and four β-pleated sheets (labeled 1-4) that surround the buried catalytic site and contribute to overall fold (Stout 2004; Graham and Peterson 1999). Using the CYP module of cytochrome P450 engineering database we analyzed the sequence of CfCYP93B (Fischer et al. 2007). All the eleven α helices (labeled A-K) and four β-pleated sheets were found to be conserved in
CfCYP93B. The (D/E)T pair and EXXR motif were present in I helix and K helix of CfCYP93B respectively. Different protein homologs of CfCYP93B that were identified using BLASTP were downloaded from NCBI database and multiple sequence alignment was carried out using these protein homologs. Cysteine heme-iron ligand signature (FGXGRRXCXG) was found to be conserved within these protein homologs (Figure 4.2). Phylogenetic tree analysis was also carried out using representative members of CYP families of plants. On phylogenetic tree analysis, CfCYP93B was found to cluster with members of the CYP93 family (Figure 4.3).

4.2.2.2 CfCYP706C

The 3’ EST sequence information of CfCYP706C was used to design the 5’ RACE primer. RACE-PCR was carried out to obtain the 5’ end of CfCYP706C. The amplicon size obtained for the 5’end of CfCYP706C was 1552 bp. Full-length cDNA of CfCYP706C is 1701 bp in size, with ORF of 1521 bp and having 35 bp 5’ UTR and 145 bp 3’UTR. Full-length sequence of CfCYP706C was submitted to NCBI genbank database and accession ID (KC307774) was obtained. The protein encoded by CfCYP706C consists of 506 amino acids with molecular weight of 56.04 kDa and theoretical pI 8.14. On BLASTX analysis CfCYP706C was found to have a maximum identity of 79% with CYP706C35 (Salvia miltiorrhiza f. alba). Secondary structure prediction of CfCYP706C was carried out using SOPMA (Geourjon and Deléage 1995). Predicted secondary structure of CfCYP706C, comprises of 49.80 % of α-helices, 6.32 % of β-turns, 11.26 % of extended strands, and 32.61 % of random coils. CfCYP706C protein was further analyzed for its conserved structure using CYP module of cytochrome P450 engineering database (Fischer et al. 2007). All the 11 alpha helices and 4 beta sheets were found to be conserved in CfCYP706C. Different homologs of CfCYP706C were identified and downloaded using BLASTP tool. Multiple sequence alignment was carried out using these sequences. Cysteine heme-iron ligand signature (FGXGRRXCXG) was found to be conserved within these homologs (Figure 4.4). Phylogenetic tree analysis was also carried out using representative members of CYP families of plants. On phylogenetic tree analysis, CfCYP706 was found to cluster with the members of the CYP706 family (Figure 4.3).
Figure 4.1: Differential display profile of P450 in *C. forskohlii*. Differential display study was carried out in different tissues (leaf, stem and root) of *C. forskohlii*. 
Figure 4.2: Multiple sequence alignment of CfCYP93B with other homologs. Conserved residues are shaded with dark red color. Amino acid sequence with the Cytochrome P450 cysteine heme-iron ligand signature (FGXGRRXCGXG) is underlined. Coleus forskohlii (Accession No. KF606861); Ocimum basilicum (Accession No. AGF30365.1); Perilla frutescens (Accession No. BAB59004.1); Salvia miltiorrhiza (Accession No. AJD25217.1); Picrophiza kurrooa (Accession No. AEG73887.1).
Figure 4.3: Clustering of CfCYP93B, CfCYP706C and CfCYP76 from C. forskohlii with members from all major plant CYP families.

Figure 4.3: Clustering of CfCYP93B, CfCYP706C and CfCYP76 from C. forskohlii with members from all major plant CYP families. MEGA5 software was used for the analysis. Tree topology support was assessed by bootstrap analysis (1000 replicates).
### Figure 4.4: Multiple sequence alignment of CfCYP706 with other homologs.

Conserved residues are shaded with dark red color. Amino acid sequence with the Cytochrome P450 cysteine heme-iron ligand signature (FGXGRXXCXXG) is underlined. *Coleus forskohlii* (Accession No. KC307774); *Salvia miltiorrhiza* (Accession No. AJD25237.1); *Solanum lycopersicum* (Accession No. XP_004245721.1); *Vitis vinifera* (Accession No. XP_002263860.2); *Glycine max* (Accession No. XP_003549879.1).

<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>C. forskohlii</em></td>
<td></td>
</tr>
<tr>
<td><em>S. miltiorrhiza</em></td>
<td></td>
</tr>
<tr>
<td><em>S. lycopersicum</em></td>
<td></td>
</tr>
<tr>
<td><em>V. vinifera</em></td>
<td></td>
</tr>
<tr>
<td><em>G. max</em></td>
<td></td>
</tr>
<tr>
<td><strong>Consensus</strong></td>
<td></td>
</tr>
</tbody>
</table>

The table continues with the alignment of sequences for each species, highlighting conserved residues with shading.
Figure 4.5: Multiple sequence alignment of CfCYP76 with other homologs.

Conserved residues are shaded with red dark color. Amino acid sequence with the Cytochrome P450 cysteine heme-iron ligand signature (FGXGRXCGXG) is underlined.

C. forskohlii (Accession No. KF643241); Sesamum indicum (Accession No. XP_011069673.1); Catharanthus roseus (Accession No. AKH60842.1); Rosmarinus officinalis (Accession No. AJQ30187.1); Salvia miltiorrhiza (Accession No. AJD25181.1).
4.2.2.3 CfCYP76

Sequence information of 3’ EST of CfCYP76 was used to design the 5’ RACE primer. RACE PCR was carried out to obtain the 5’ end of CfCYP76. The amplicon size for 5’ RACE product was 1480 bp. Full-length cDNA of CfCYP76 is 1622 bp in size, with ORF of 1479 bp, 41 bp 5’UTR, and 102 bp 3’UTR. Full-length sequence of CfCYP76 was submitted to NCBI genbank database and accession ID (KF643241) was obtained. CfCYP76, encodes a protein of 492 amino acid, having a molecular weight of 55.568 kDa and theoretical pI 8.76. On BLASTX analysis, CfCYP76 was found to have a maximum identity of 53% with ferruginol synthase (Sesamum indicum var. malabaricum).

The secondary structure of CfCYP76 was predicted using SOPMA (Geourjon and Deléage 1995). The result showed that CfCYP76 contain 48.58% of α-helices, 7.93% of β-turns, 12.80% of extended strands, and 30.69% of random coils. All the 11 alpha helices and 4 beta sheets were found to be conserved in CfCYP76. This analysis was carried out using CYP module of cytochrome P450 engineering database (Fischer et al. 2007). Different homologs of CfCYP76 were identified and downloaded using BLASTP tool. Multiple sequence alignment was carried out using these sequences. Cysteine heme-iron ligand signature (FGXGRRXCXG) was found to be conserved in these homologs (Figure 4.5). Phylogenetic tree analysis was also carried out using representative members of CYP families of plants. On phylogenetic tree analysis, CfCYP76 was found to cluster with a member of the CYP76 family (Figure 4.3).

4.2.3 Expression study

4.2.3.1 Tissue expression study for CfCYP93B, CfCYP706B, and CfCYP76

Semi quantitative RT-PCR was carried out for CfCYP93B, CfCYP706B and CfCYP76 in young leaves, mature leaves, stem, roots and root tips of C. forskohlii. Actin was used as an internal control for the expression study. CfCYP93B showed higher expression in young leaves, mature leaves and roots as compared to other tissues. Whereas, CfCYP706C expression was dominant in young leaves and mature leaves. CfCYP76 expression was found to be dominant in roots and root tips (Figure 4.6).
Figure 4.6: Semi quantitative expression study in different tissues.

Figure 4.6: Semi quantitative expression study in different tissues. Semi quantitative expression study of CfCYP93B, CfCYP706C, CfCYP76 in different tissues (young leaf, mature leaf, stem, root and root tip) of C. forskohlii.
4.2.3.2 qPCR study of CfCYP93B and CfCYP706C under stress treatment

Time course expression profiles of CfCYP93B and CfCYP706C were studied in response to mannitol stress using quantitative real time RT-PCR analysis. Actin was used as housekeeping control for normalization and relative quantification was carried out by taking the expression of the gene of interest at 0 h (just before treatment), as a baseline for calculating fold change. CfCYP93B was upregulated in response to mannitol treatment. We observed that CfCYP93B mRNA accumulation reaches a maximum level after 5 h of treatment and then gradually reduces to almost the same levels as in untreated plants after 10 h of treatment. CfCYP706C was down regulated in response to mannitol stress. CfCYP706C expression reaches to its minima after 10 h of treatment and then its expression level gradually increase to 0.16 fold after 24 h of mannitol treatment (Figure 4.7).

4.2.4. Quantification of total flavonoid content

Total flavonoid content in leaves, stems, and roots of Coleus forskohlii was determined using spectrophotometry. Total flavonoid content was highest in leaves followed by roots and stems. Quercitin was used as a standard for quantification of total flavonoid content (Table 4.1).

Table 4.1: Total flavonoid content in different tissues of C. forskohlii.

<table>
<thead>
<tr>
<th>Plant species</th>
<th>Leaves</th>
<th>Stems</th>
<th>Roots</th>
</tr>
</thead>
<tbody>
<tr>
<td>C. forskohlii</td>
<td>19.09228442</td>
<td>11.21368</td>
<td>16.09394</td>
</tr>
</tbody>
</table>

4.2.5 Homology modeling and docking studies

4.2.5.1 Protein modeling and docking studies of CfCYP93B and CfCYP706C

Homology modeling was carried out for CfCYP93B and CfCYP706C using CYP17A1 (PDB ID: 3RUKA) as a template. Protein model was developed using schrödinger suite 2013 (Maestro, version 9.4, Schrödinger, LLC, New York, NY, 2013). Heme group copied from the template was incorporated in the protein models. The stereo-chemical analysis of the predicted models of CfCYP93B and CfCYP706C proteins was analyzed by PROCHECK (http://services.mbi.ucla.edu/SAVES/). Ramachandran plot analysis of CfCYP93B showed 83.7% residues in the most
Figure 4.7: Relative expression profile of CYPs in response to mannitol treatment at different time points.

Figure 4.7: Relative expression profile of CYPs in response to mannitol treatment at different time points (0 h, 2 h, 5 h, 10 h and 24 h). (A) CfCYP706C (B) CfCYP93B. Actin was used as housekeeping control and expression of gene-of-interest at 0 h was used as baseline for calculating fold change.
favorable region, 13.9% residues in the additional allowed region, 1.0% in the generously allowed region and 1.4% in the disallowed region (Figure 4.8). Whereas CfCYP706C model, showed 84.3% residues in the most favorable region, 12.1% residues in the additional allowed region, 2.5% residues in the generously allowed region and 1.0% in the disallowed region (Figure 4.9). The results of the PROCHECK analysis indicate that a relatively low percentage of residues have phi/psi angles in the disallowed regions suggesting the acceptability of Ramachandran plots for CYPs proteins. The percentage of residues in the allowed region were found to be 98.6 and 99% for CfCYP93B and CfCYP706C respectively, while residues in disallowed regions were found to be 1.4% and 1% for CfCYP93B and CfCYP706C respectively.

As per the previous report, CYP93B is known as flavones synthase and it acts on various flavonones such as naringenin and related molecules. CYP706C is not a well-characterized gene and its shows homology with flavonoids 3’ hydroxylase (F3H). Flavonones are known substrates for F3H. Docking experiments were performed using flavonones (naringenin, butin, isosakuranetin and eriodictyol) as ligand data set and CfCYP93B and CfCYP706C as protein models (Figure 4.10 to 4.17), to identify the likely substrate. Based on the docking score, CfCYP706C showed a similar binding affinity for all the ligand dataset i.e., naringenin, isosakuranetin, eriodictyol and butin (Table 4.2), whereas CfCYP93B showed a higher binding affinity for eriodictyol, naringenin, and butin. Amino acids surrounding four ligand dataset were found to be almost conserved in CfCYP706 and CfCYP93B.

**Table 4.2:** Docking score of ligands, docked on modeled structure of CfCYP93B and CfCYP706

<table>
<thead>
<tr>
<th>S. No</th>
<th>Ligand</th>
<th>Docking Score</th>
<th>CfCYP93B</th>
<th>CfCYP706</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Naringenin</td>
<td>-3.164433</td>
<td>-8.35417</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Isosakuranetin</td>
<td>-2.570103</td>
<td>-8.655886</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Eriodictyol</td>
<td>-4.273103</td>
<td>-8.817555</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Butin</td>
<td>-3.649534</td>
<td>-8.779972</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.8: Protein modelling and topological assessment of CfCYP93B. (A) Three-dimensional structure of CfCYP93B. Protein model for CfCYP93B was made using homology modeling module of schrödinger suite 2013 and heme group was incorporated. (B) Ramachandran plot of CfCYP706. The plot calculations of CfCYP706 was computed using PROCHECK. Most favoured regions are coloured red (A, B, L), additional allowed (a, b, l, p), generously allowed (~a,~b,~l,~p), and disallowed regions are indicated as yellow, light yellow and white regions, respectively.
Figure 4.9: Protein modelling and topological assessment of CfCYP706. (A) Three dimensional structure of CfCYP706. Protein model for CfCYP706 was made using homology modeling module of schrödinger suite 2013 and heme group was incorporated. (B) Ramachandran plot of CfCYP706. The plot calculations of CfCYP706 was computed using PROCHECK. Most favoured regions are coloured red (A, B, L), additional allowed (a, b, l, p), generously allowed (~a, ~b, ~l, ~p), and disallowed regions are indicated as yellow, light yellow and white regions, respectively.
Figure 4.10: Molecular docking study of CfCYP93B with naringenin. (A) Molecular docking of CfCYP93B with naringenin. (B) 2D representation of interaction between CfCYP93B and naringenin.
Figure 4.11: Molecular docking and ligand interaction of CfCYP93B. (A) Molecular docking of CfCYP93B with butin. (B) 2D representation of interaction between CfCYP93B and butin. Arrow indicate the interaction between the butin and CfCYP93B.
Figure 4.12: Molecular docking study of CfCYP93B with Isosakuranetin. (A) Molecular docking of CfCYP93B with Isosakuranetin. (B) 2D representation of interaction between CfCYP93B and Isosakuranetin.
Figure 4.13: Molecular docking study of CfCYP93B with eriodictyol. (A) Molecular docking of CfCYP93B with Eriodictyol. (B) 2D representation of interaction between CfCYP93B and Eriodictyol. Arrow indicate the interaction between eriodictyol and CfCYP93B.
Figure 4.14: Molecular docking study of CfCYP706C with naringenin. (A) Molecular docking of CfCYP706C with naringenin. (B) 2D representation of interaction between CfCYP706C and naringenin. Arrow indicate the interaction between ligand and surrounding amino acids.
Figure 4.15: Molecular docking study of CfCYP706C with butin. (A) Molecular docking of CfCYP706C with butin. (B) 2D representation of interaction between CfCYP706C and butin.
Figure 4.16: Molecular docking study of CfCYP706C with Isosakuranetin. (A) Molecular docking of CfCYP706C with Isosakuranetin. (B) 2D representation of interaction between CfCYP706C and Isosakuranetin. Arrow indicates the interaction between ligand and surrounding amino acids.
Figure 4.17: Molecular docking study of CfCYP706C with eriodictyol. (A) Molecular docking of CfCYP706C with Eriodictyol. (B) 2D representation of interaction between CfCYP706C and Eriodictyol. Arrow indicate the interaction between ligand and surrounding amino acids.
### 4.2.5.2 Three dimensional structure prediction and docking studies of CfCYP76

Protein model for CfCYP76 was predicted using CYP17A1 (PDB ID: 3RUKA) as a template. Protein model was developed using schrödinger suite 2013 (Maestro, version 9.4, Schrödinger, LLC, New York, NY, 2013). Heme group copied from the template was incorporated in the protein models. The stereo-chemical analysis of CfCYP76 was checked by PROCHECK server (http://services.mbi.ucla.edu/SAVES/). Ramachandran plot analysis of CfCYP76 showed 82.7% residues in the most favorable region, 13.6% residues in the additional allowed region, 2.0% in the generously allowed region and 1.7% in the disallowed region (Figure 4.18). The results of the PROCHECK analysis indicate that a relatively low percentage of residues have phi/psi angles in the disallowed regions suggesting the acceptability of Ramachandran plots for CYPs proteins. The percentage of residues in the allowed region was found to be 98.3%, while the percentage of residues in disallowed regions was found to be 1.7%. A total of 44 terpenoids from *C. forskohlii* were downloaded from a dictionary of natural product database. Docking of these ligands was carried out using virtual screening workflow module of Schrödinger Suite 2013 (Maestro, version 9.4, Schrödinger, LLC, New York, NY, 2013). The best two docked ligands are shown in table 4.3. The binding affinity of both the ligand molecule was found to be almost same. Amino acids that were found to be interacting with coleolactone and 3-hydroxyisoforskolin are shown in figure 4.19.

#### Table 4.3: Docking score of ligands, docked on modeled structures of CfCYP76

<table>
<thead>
<tr>
<th>S. No</th>
<th>Ligand</th>
<th>Docking Score</th>
<th>CfCYP76</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Coleolactone</td>
<td>-6.166</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3- hydroxyisoforskolin</td>
<td>-6.055</td>
<td></td>
</tr>
</tbody>
</table>

### 4.3 Discussions

Secondary metabolites play wide range of functions in plants and have been extensively used in pharmaceuticals (e.g. artemisinin and taxol), herbal medicines (e.g. glycyrrhizin and ginsenosides), fragrances (e.g. citronellol and geraniol), nutraceuticals (e.g. astaxanthin and lycopene), flavors (e.g., limonene and linalool), etc. (Vining 1990; Verpoorte, van der Heijden, and Memelink 2000; Hussain et al. 2012). CYPs play an important role in the biosynthesis of secondary metabolites ( 
Figure 4.18: Protein modelling and topological assessment of CfCYP76. (A) Three-dimensional structure of CfCYP76. Protein model for CfCYP76 was made using homology modeling module of Schrödinger suite 2013 and heme group was incorporated. (B) Ramachandran plot of CfCYP76. The plot calculations of CfCYP706 was computed using PROCHECK. Most favoured regions are coloured red (A, B, L), additional allowed (a, b, l, p), generously allowed (~a,~b,~l,~p), and disallowed regions are indicated as yellow, light yellow and white regions, respectively.
Figure 4.19: Molecular docking study of CfCYP76. (A) Molecular docking of CfCYP93B with coleolactone. (B) Molecular docking of CfCYP93B with 3-hydroxyisoforskolin. (C) 2D representation of interaction between CfCYP76 and coleolactone. (D) 2D representation of interaction between CfCYP76 and 3-hydroxyisoforskolin. Arrow indicate the interaction between ligands and surrounding amino acids.
Molecular cloning........... Coleus forskohlii

Bak et al. 2011). Few small motifs are conserved among different families (Schopfer and Ebel 1998), making the cloning of CYPs, using homology based approach a difficult task. In this study, we have used the differential display technique to isolate the differentially expressed CYPs from *C. forskohlii*. Differential display study has earlier been used in plants (*Glycine max* (L.) Merr., *Gerbera* (hybrid), *Astragalus chrysochlorus* Boiss. & Kotschy) to isolate the differentially expressed CYP transcripts (Turgut-kara and Ari 2011; Schoendorf et al. 2001; Martens and Forkmann 1999; Schopfer and Ebel 1998). Differential display study has played an important role in the identification of CYPs involved in the biosynthesis of taxol in *Taxus cuspidata* Siebold & Zucc. (Schoendorf et al. 2001). In this study, we have cloned three CYP genes (CfCYP93B, CfCYP706C, and CfCYP76) using this approach.

Phylogenetic tree analysis, reveals that CfCYP93B and CfCYP706C are found to be clustered with members of CYP93 and CYP706 families of CYPs respectively. CYP93B family is generally involved in the biosynthesis of flavonoids (Ayabe and Akashi 2006; Akashi, Aoki, and Ayabe 1999). CYP93C2 isolated from *Glycyrrhiza echinata* L. (Fabaceae) encodes a 2-hydroxyisoflavanone synthase, which catalyzes the hydroxylation associated with 1, 2-aryl migration of flavanones (Akashi, Aoki, and Ayabe 1999). So far none of the CYP706C family members have been characterized in detail. On BLASTP analysis, CfCYP706C was also found to have high sequence identity with flavonoid 3’ monooxygenases from other plants. Till date, there has been no report of functional characterization of CYP706C member in plants.

One of the member of another subfamily (CYP706B) from *Gossypium hirsutum* L. in known to be involved in the biosynthesis of terpenoids (Luo et al. 2001; Bak et al. 2011). There are seven CYP706 genes in the *Arabidopsis* genome but their function so far has not been reported ( Bak et al. 2011). The expression level of CfCYP93B was up regulated in young leaves, mature leaves, and roots. Whereas, CfCYP706C expression was dominant in young leaves and mature leaves. CYP93B11 that encodes a putative flavone synthase exhibits maximum expression in leaves and siliques in *Medicago trunculata* Gaertn. (Fabaceae). The expression study of CfCYP93B and CfCYP706C was studied in response to mannitol stress. CfCYP93B expression was up regulated in response to mannitol, whereas CfCYP706C expression was down regulated. Flavone synthase genes (GmFNSII-1 and GmFNSII-2) from *Glycine max* (L.) Merrill was also reported to be induced by mannitol stress (Yan et al. 2014). Only one flavonoid named as genkwanin is reported from *C. forskohlii* (Alasbahi and
Melzig 2010). Total flavonoid content was found to be prominent in leaves, which is correlating with the tissue expression profile of CfCYP93B and CfCYP706C. This result shows that CfCYP93B and CfCYP706C might play a role in the biosynthesis of flavonoids.

Functional characterization of CYPs is time-consuming, labor intensive, and requires biochemical studies, as well as the study of site-specific mutants. Elucidation of plant CYPs crystal structure is a daunting task. Only one solved the crystal structure of unusual CYP is available from plant i.e. CYP74A. Molecular modeling and docking study is becoming an important tool to guide the biochemical and functional characterization of the large number of CYPs, as well as facilitating CYP enzyme engineering and metabolic pathway engineering of crop plants. To circumvent this bottleneck in the study of CYPs, homology modeling has proved to be a strong tool for the identification of potentially important residues of SRS (Substrate recognition sites) (Rupasinghe and Schuler 2006). In Arabidopsis thaliana, homology modeling based structure analysis has been carried out for CYPs (CYP73A5, CYP84A1, CYP98A3, CYP75B1) involved in the biosynthesis of flavonoids (Rupasinghe and Schuler 2006). In our study, protein models of CfCYP93B and CfCYP706C were predicted using homology modeling. These protein models were further assessed by Ramachandran plot for stereo-chemical property. Docking studies of these modeled proteins were carried out with four different flavonones (naringenin, butin, isosakuranetin and eriodictyol) to study the binding pocket and probable substrate candidates for CfCYP93B and CfCYP706C.

CfCYP76 belongs to the CYP76 family having a maximum identity of 53% with ferruginol synthase (Sesamum indicum) on BLASTX analysis. This information tells us that, CfCYP76 could be the new subfamily in CYP76 as per P450 nomenclature system given by Dr. Nelson (Nelson 2009). Members of the CYP76 family are generally involved in the biosynthesis of terpenoids (Collu et al. 2001; Swaminathan et al. 2009). In Salvia miltiorrhiza Bunge (Lamiaceae) and Rosmarinus officinalis L. (Lamiaceae), CYP76AH1 and CYP76H4 respectively, are involved in hydroxylation of norditerpene abietatriene (Zi and Reuben 2013). CfCYP76 expression was found to be confined to the roots and root tips of C. forskohlii. Since labdane diterpenoids are restricted to roots of C. forskohlii, it might be possible that CfCYP76 may have a role in the biosynthesis of labdane diterpenoids in C. forskohlii. Homology modeling and docking study of CfCYP76 showed that 3-hydroxyforskolin or coleolactone could be
the probable substrate for CfCYP76. 3-hydroxyforskolin is a labdane diterpenoid isolated from of *C. forskohlii* (Alasbahi and Melzig 2010), suggesting that CfCYP76 could be involved in the biosynthesis of labdane diterpenoids, which needs to be further validated through experimental analysis.