

Structure prediction and functional characterization of secondary metabolite proteins of *Ocimum*

Sudeep Roy*, Nidhi Maheshwari, Rashi Chauhan, Naresh Kumar Sen, Ashok Sharma

Biotechnology Division, Central Institute of Medicinal and Aromatic Plants, (Council of Scientific and Industrial Research), Kukrail Picnik Spot Road, P.O. CIMAP, Lucknow - 226015, India; Sudeep Roy - Email: roysudeep28@gmail.com; Phone: +919450614880; *Corresponding author

Received June 18, 2011; Accepted June 27, 2011; Published July 06, 2011

Abstract:

Various species of *Ocimum* have acquired special attention due to their medicinal properties. Different parts of the plant (root, stem, flower, leaves) are used in the treatment of a wide range of disorders from centuries. Experimental structures (X-ray and NMR) of proteins from different *Ocimum* species, are not yet available in the Protein Databank (PDB). These proteins play a key role in various metabolic pathways in *Ocimum*. 3D structures of the proteins are essential to determine most of their functions. Homology modeling approach was employed in order to derive structures for these proteins. A program meant for comparative modeling-Modeller 9v7 was utilized for the purpose. The modeled proteins were further validated by Procheck and Verify-3d and Errat servers. Amino acid composition and polarity of these proteins was determined by CLC-Protein Workbench tool. Expasy's Prot-param server and Cys_rec tool were used for physico-chemical and functional characterization of these proteins. Studies of secondary structure of these proteins were carried out by computational program, Profunc. Swiss-pdb viewer was used to visualize and analyze these homology derived structures. The structures are finally submitted in Protein Model Database, PMDB so that they become accessible to other users for further studies.

Keywords: *Ocimum*, Homology modeling, CLC protein work bench, Secondary structure prediction, Swiss-PDB Viewer

Background:

Botanically, basil belongs to the genus *Ocimum* of the family Lamiaceae. More than 160 species of *Ocimum* are reported from different parts of the world. Different parts (roots, stem, leaves, seeds and flowers) of *Ocimum* have been used for treatment of variety of diseases such as bronchitis, malaria, diarrhea, dysentery, skin diseases, arthritis etc. *Ocimum* sp. contains monoterpene derivatives such as camphor, limonene, thymol, citral, geraniol and linalool. A detailed analysis of protein sequences from *Ocimum*, their probable structures and mode of action are yet to be accomplished. Plants synthesize chemicals in their leaves in order to protect themselves from herbivores. One such class of defense compounds that has been used extensively by humans are members of phenylpropanoid class namely eugenol, chavicol and their derivatives. It has been reported that in basil glands, two closely related (90% identical) enzymes chavicol o-methyltransferase (CVOMT) and eugenol o-methyltransferase (EOMT) catalyze the formation of methylchavicol and methyleugenol from chavicol and eugenol respectively [1]. The enzymes are involved in aroma production in basil. From an evolutionary perspective plant and microbial PALs (phenylalanine ammonia lyase) are part of superfamily of enzymes from plants, fungi and bacteria, and are likely derived from a precursor of the widespread histidine ammonia lyase (HAL) family in the histidine degradation pathway [2]. PAL catalyses the non-oxidative deamination of phenylalanine to trans-cinnamate and directs the carbon flow from the shikimate pathway to the various branches of the general phenylpropanoid metabolism. Lipooxygenase (Fatty-acid metabolism) is one of the most widely studied enzymes found in more than 60 species of plant and animal kingdom. The enzyme catalyses the biooxygenation of polyunsaturated fatty acids (PUFA) containing a cis, cis-1, 4-pentadiene unit to form conjugated hydroperoxydienoic acids. Lipooxygenase

has considerable application in food related products such as in bread making. The enzyme plays a significant role in formation of secondary metabolites in sweet basil.

Enzymatic browning of fruits and vegetables is caused mainly by the conversion of native phenolic compounds to quinones which are then polymerized to brown, red or black pigments imparting colour to various plant parts. The enzymes responsible for catalyzing this sequence of reactions are termed as polyphenol oxidases, but are also known as tyrosinases, catecholases, cresolases and phenolases [3]. Because of deleterious effect of enzymatic browning on fruits and vegetables much work is devoted as to retard or at least delay the browning process. Polyphenol oxidase being the causative agent responsible for browning is exploited for the purpose. The enzyme is involved in iso-quinoline alkaloid biosynthesis and in biosynthesis of other secondary metabolites. In order to understand biochemical function and interaction properties of the protein at molecular level, three dimensional structure of protein is foremost requirement. However, the number of available protein sequences exceeds far behind the available three dimensional protein structures. In order to compensate this, homology modeling approach came into being. These methods are believed to be cost-effective and time-effective when compared to X-rays crystallography and NMR techniques. Computational methods make use of hidden information inside amino acid sequences in order to predict protein structure and function. In the present study, *In silico* analysis and homology modeling studies on uncharacterized proteins in different species of *Ocimum* like *O. basilicum*, *O. tenuiflorum*, *O. citriodorum*, *O. seloi*, *O. gratissimum* and *O. americanum* whose structures are not yet available in PDB have been accomplished.

Materials and Methodology:

The amino acid sequences of secondary metabolite proteins of *Ocimum* whose structures are not yet available in RCSB Protein Databank (PDB) were retrieved from SWISSPROT, a public resource of curated protein sequences [4] and subjected to NCBI BLAST [5]. Based on high score, lower e-value and maximum sequence identity, the best template was selected which was then used as reference structure to build a 3D model. Template and target proteins considered for the study have been shown in **Table 1** (see **Supplementary material**).

Model building and evaluation:

The three dimensional structures of proteins were modeled using Modeler 9v8 [6]. Quality of generated models was evaluated with PROCHECK [7] by Ramachandran plot analysis [8]. Stereochemical quality and accuracy of the selected models was further improved by subjecting it to energy minimization with the GROMOS 96 43B1 parameters set, implementation of Swiss-PDB Viewer [9]. Validation of generated models was further performed by VERIFY 3D [10] and ERRAT [11] programs. ProSA [12] was used for the analysis of Z-scores and energy plots. The three dimensional structures of modeled proteins were analyzed using Deep View Swiss PDB viewer. Root Mean Square Deviation (RMSD) values were calculated between the set of targets and template protein to see how much modeled protein deviates from the template protein structure.

Computation of amino acid composition:

Amino acid composition (**Table 2** see **supplementary material**) of *Ocimum* proteins under study was calculated using CLC protein workbench tool (www.clcbio.com/protein). The tool also provides estimation of percentage of hydrophobic and hydrophilic residues present in the protein (**Table 3** see **Supplementary material**).

Physicochemical characterization:

For physicochemical characterization, theoretical pI (isoelectric point), molecular weight, -R and +R (total number of positive and negative residues), EI (extinction coefficient) [13], II (instability index [14] [15], AI (aliphatic index) and GRAVY (grand average hydropathy) [16] were computed using the ExPasy's ProtParam server [17] for set of proteins (<http://us.expasy.org/tools/protparam.html>). The results are shown in **Table 4** (see **Supplementary material**).

Functional characterization:

CYS_REC (<http://sunl.softberry.com/berry.phtml?topic>) was used to locate "SS bond" between the pair of cysteine residues, if present. The tool yields position of cysteins, total number of cysteins present and pattern, if present, of pairs in the protein sequence as output. All the *Ocimum* proteins under study showed absence of disulphide bonds. The results are presented in **Table 5** (see **Supplementary material**).

Secondary structure prediction:

Profunc [18] was employed for calculating the secondary structural features of *Ocimum* protein sequences. The results are presented in **Table 6** (see **Supplementary material**).

Submission of the modeled proteins in protein model database (PMDB):

The models generated for various *Ocimum* proteins were successfully submitted in Protein model database, PMDB [19] without any stereochemical errors. The submitted models can be accessed via their PMIDs (**Table 7** see **Supplementary material**).

Results and Discussion:

As experimental structures of some of the important secondary metabolite proteins of *Ocimum* are not available, homology modeling approach was used in order to derive their structures.

Model building, refinement and evaluation:

PROCHECK analysis:

Ramachandran plot for Chavicol O-methyltransferase (D3KYA1) has been illustrated in **Figure 1**. Altogether more than 90% of the residues were found to be in favoured and allowed regions, which validate the quality of homology models. The overall G-factor for D3KYA1 was -0.19. As the value is greater than the acceptable value -0.50, this suggests that the modeled structure is acceptable. The modeled structures were also validated by other structure verification servers such as Verify 3D and Errat. Verify 3D assigned a 3D-1D score of >0.2 for all the modeled proteins. This implies that the models are

compatible with its sequence. ERRAT showed overall quality factor of 49.62 for D3KYA1. The plot generated by Verify-3D and Errat for Chavicol omethyltransferase has been illustrated in **Figure 2A & 2B**.

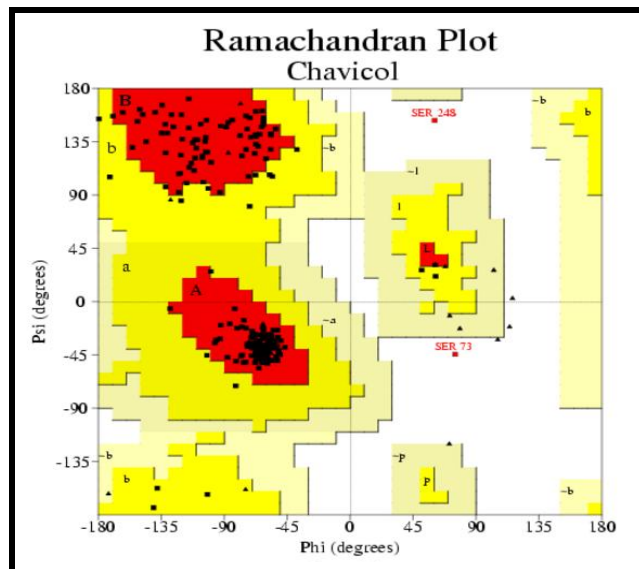


Figure 1: Ramachandran plot of chavicol obtained by PROCHECK. 91.2% residues in favourable regions; 8.0% residues in additional residue regions; 0.0% residues in generously regions; 0.9% residues in disallowed regions; Over all G-factor: 0.00.

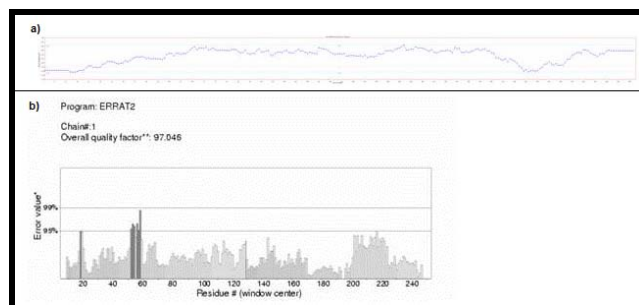


Figure 2: (A) Verify-3D plot, (B) Errat plot

PROSA analysis:

The z-score for all the modeled proteins was found to be within the range of scores typically found for native proteins of similar size showing good quality of the model. Energy Plot for chavicol o-methyltransferase (D3KYA1) with chain length (257 AA) and z-score (-7.26) is presented in **Figure 3A & B**.

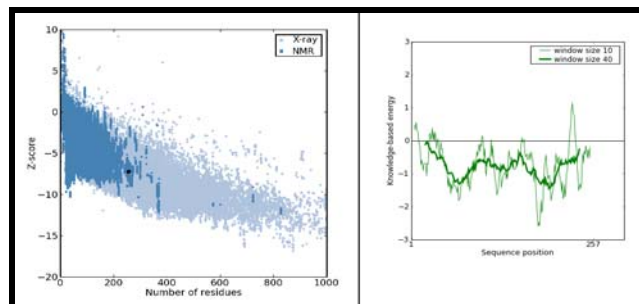


Figure 3: (A) Prosa-web z-score plot, (B) Prosa-web plot of residue scores

Swiss-PDB viewer analysis of predicted model:

Visualization and analysis of the model using Swiss-PDB reveals that there are no steric hindrances between the residues and thus modeled structures are stable. Structure-structure superimposition was done in order to calculate Root

Mean Square Deviation (RMSD) between the target and template sequence. RMSD values for D3KYA1 were found to be 0.94. This implies good quality of the modeled structures. **Figure 4** represents modeled structure of Chavicol o-methyltransferase.

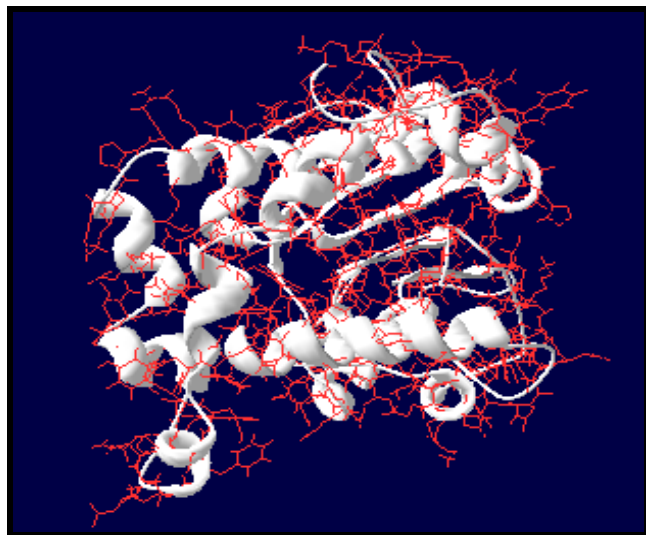


Figure 4: Modeled structure of Chavicol o-methyltransferase as viewed by Swiss-PDB viewer

Physiochemical characterization:

The physiochemical parameters viz., theoretical isoelectric point (Ip), molecular weight, total number of positive and negative residues, extinction coefficient, half-life, instability index, aliphatic index and grand average hydropathy (GRAVY) were computed using the ExPASy's ProtParam tool (**Table 4**). The computed pI value for A8D7D8, B2ZA17, B6VQV5, B6VQV6, D3KYA1 (pI<7) indicated their acidic nature, whereas pI for A8D6D7, B2ZA12, B2ZA16 (pI>7) revealed their basic behaviour. The computed isoelectric point (pI) will be useful for developing buffer system for purification by isoelectric focusing method. Extinction coefficient values for *Ocimum* proteins at 280 nm ranged from 1490 to 50795 M⁻¹cm⁻¹ for B6VQV6 and D3KYA1 indicating the presence of higher concentration of Tyr and Trp. Cys was very low in concentration in all the eight *Ocimum* proteins studied. This indicates that these proteins cannot be analyzed using UV spectral methods. On the basis of instability index ExPASy's ProtParam classified the B2ZA17 (Eugenol o-methyltransferase), A8D7D8 (Lipoxygenase), B2ZA12 (Eugenol o-methyltransferase) and B2ZA16 (Eugenol o-methyltransferase) proteins as unstable (Instability index>40) and other *Ocimum* proteins as stable (Instability index<40). The aliphatic index (AI) which is defined as the relative volume of a protein occupied by aliphatic side chain is regarded as the positive factor for the increase of thermal stability of globular proteins. The very high aliphatic index of all *Ocimum* proteins infers that these proteins may be stable for a wide range of temperature. The very low GRAVY index of proteins B6VQV6 and D3KYA1 infers that these proteins could result in a better interaction with water.

Functional characterization:

The result of primary analysis suggests that all the *Ocimum* proteins under study were hydrophobic in nature due to the presence of high non-polar residues content (**Table 2 & 3**). As percentage of Cysteine(C) is very low in all the *Ocimum* proteins under study (**Table 2**), none of these proteins have disulphide bond linkages, as indicated by CYS_REC result (**Table 4**). The extensive hydrogen bonding may provide stability to these proteins in absence of disulphide bonds. Proteins B2ZA12, A8D6D7 and B6VQV5 have high percentage of methionine(M), alanine(A), leucine(L) and lysine(K). As these amino acids have high helix-forming propensities, alpha helix are dominant in these proteins. This is also evident from analysis of PROFUNC result (**Table**

6). Rest of the *Ocimum* proteins had mixed secondary structures i.e. alpha-helices, beta-strands and coils. All the proteins showed high percentage of glycine and proline (**Table 2**). As these amino acids are common in turns, other secondary structures such as Beta turns and Gamma turns are dominant in these proteins (**Table 6**).

Submission of modeled proteins in PMDB:

The modeled structures of proteins from various species of *Ocimum* were successfully deposited in Protein Model Database (PMDB). The PMDB ID for the submitted structures has been presented in **Table 7** (see **Supplementary material**). These 3D structures may be further used in characterizing the protein experimentally.

Conclusion:

In this study proteins from various species of *Ocimum* were modeled using homology modeling approach. Different parameters such as isoelectric point, molecular weight, total number of positive and negative residues, extinction coefficient, instability index, aliphatic index and grand average hydropathy (GRAVY) were computed for these proteins in order to determine their physiochemical characteristics. All the proteins were found to be deficient in amino acid cysteine, and therefore lack presence of disulphide linkages as also inferred from analysis of cys_rec result. In the absence of disulphide bond, extensive hydrogen bonding is believed to be responsible for stability of these proteins. Polarity studies using CLC protein work bench tool confirmed all the studied proteins to be hydrophobic in nature. This may be due to the presence of a large number of non-polar residues. Secondary structure studies showed that all the studied proteins contain high proportion of other secondary structures i.e. Beta-turns and Gamma-turns. This is attributed to the presence of higher concentration of proline and glycine residues. The modeled structures can be accessed through protein model database PMDB via their PMID's. Homology derived models are extensively used in wide range of applications such as virtual screening, site-directed mutagenesis experiments or in rationalizing the effects of sequence variation. These structures will serve as cornerstone for functional analysis of experimentally derived crystal structures.

Acknowledgements:

Financial support of Department of Biotechnology (DBT) Govt of India; New Delhi under BTISnet program is gratefully acknowledged. Sudeep Roy is thankful to Council of Scientific and Industrial Research for Senior Research Fellowship.

References:

- [1] Koeduka T *et al. Plant Physiol.* 2009 **149**: 384 [PMID: 18987218]
- [2] Rother D *et al. Eur J Biochem.* 2002 **269**: 3065 [PMID: 12071972]
- [3] Mayer AM & Harel E. *Phytochemistry.* 1979 **18**: 193
- [4] Bairoch A & Apweiler R. *Nucleic Acids Res.* 2000 **28**: 45 [PMID: 10592178]
- [5] Altschul SF *et al. J Mol Biol.* 1990 **215**: 403 [PMID: 2231712]
- [6] Sali A & Blundell TL. *J Mol Biol.* 1993 **234**: 779 [PMID: 8254673]
- [7] Laskowski RA *et al. J Appl Cryst.* 1993 **26**: 283
- [8] Ramachandran GN *et al. J Mol Biol.* 1963 **7**: 95 [PMID: 13990617]
- [9] Kaplan W & Littlejohn TG. *Brief Bioinform.* 2001 **2**: 195 [PMID: 11465736]
- [10] Eisenberg D *et al. Methods Enzymol.* 1997 **277**: 396 [PMID: 9379925]
- [11] Colovos VC & Yeates TO. *Protein Sci.* 1993 **2**: 1511 [PMID: 8401235]
- [12] Sippl MJ. *Proteins.* 1993 **17**: 355 [PMID: 8108378]
- [13] Gill SC & von Hippel PH. *Anal Biochem.* 1989 **182**: 319 [PMID: 2610349]
- [14] Guruprasad K *et al. Protein Eng.* 1990 **4**: 155 [PMID: 2075190]
- [15] Ikai A. *J Biochem.* 1980 **88**: 1895 [PMID: 7462208]
- [16] Kyte J & Doolittle RF. *J Mol Biol.* 1982 **157**: 105 [PMID: 7108955]
- [17] Wilkins MR *et al. Methods Mol Biol.* 1999 **112**: 531 [PMID: 10027275]
- [18] Laskowski RA *et al. Nucleic Acids Res.* 2005 **33**: W89 [PMID: 15980588]
- [19] Castrignanò T *et al. Nucleic Acids Res.* 2006 **34**: D306 [PMID: 16381873]

Edited by P Kanguane

Citation: Roy *et al.* Bioinformation 6(8): 315-319 (2011)
provided the original author and source are credited.

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes,

Supplementary material:

Table 1: Target protein and template protein considered for the study

Target protein	Uniprot ID	Length	Template protein	PDB ID	e-value
Polyphenol oxidase	A8D6D7	139 AA.	Catechol oxidase from <i>Ipomoea batatas</i> (Sweet Potatoes)	1BT1	4e-42
Lipoxygenase	A8D7D8	156 AA.	Crystal structure of delta 413-417:gs I805w Lox	3FG3	0.004
Eugenol o-methyltransferase	B2ZA12	200 AA.	Crystal structure of isoflavone o-methyltransferase homolog	2QYO	1e-35
Eugenol o-methyltransferase	B2ZA16	200 AA.	Crystal structure of Bchu involved in bacteriochlorophyll C biosynthesis	1X19	0.23
Eugenol o-methyltransferase	B2ZA17	154 AA.	Crystal structure analysis of isoflavone o-methyltransferase	1FP2	3e-23
Phenylalanine ammonia-lyase	B6VQV5	145 AA.	Phenylalanine ammonia-lyase (Pal) from <i>Petroselinum crispum</i>	1W27	2e-62
Chavicol o-methyltransferase	B6VQV6	257 AA	Crystal structure of isoflavanone 4'-o-methyltransferase complexed with (+)-6a-hydroxymaackiain	1ZGA	1e-56
Chavicol o-methyltransferase	D3KYA1	356 AA.	Crystal structure of a putative o-methyltransferase	2R3S	0.005

Table 2: Amino acid composition (%) of *Ocimum* proteins

Proteins	A8D6D7	A8D7D8	B2ZA12	B2ZA16	B2ZA17	B6VQV5	B6VQV6	D3KYA1
Amino acids								
Alanine(A)	5	7.1	8.5	8.5	9.7	11.7	7.8	7.6
Cysteine(C)	0	1.3	2.0	2.0	1.3	0.7	2.3	2.0
Aspartic acid(D)	6.5	5.8	3.5	3.5	3.9	2.1	7.8	6.5
Glutamic acid(E)	2.2	6.4	5.0	5.0	6.5	6.9	5.4	5.3
Phenylalanine(F)	6.5	1.3	4.0	4.0	5.2	4.1	4.3	4.5
Glycine(G)	7.2	6.4	5.5	5.5	5.8	9.7	7.4	5.6
Histidine(H)	2.9	5.8	7.0	7.0	7.1	2.8	3.5	3.9
Isoleucine(I)	3.6	7.7	5.0	5.0	3.2	7.6	5.8	6.2
Lysine(K)	2.9	7.1	5.5	5.5	4.5	4.1	6.2	5.9
Leucine(L)	5.8	11.5	10.5	10	9.1	15.2	8.6	10.1
Methionine(M)	3.6	1.9	5.5	5.5	5.2	1.4	4.3	4.5
Asparagine(N)	7.2	5.1	4.5	4.5	4.5	4.8	3.9	4.8
Proline(P)	10.8	5.1	5.0	5.0	4.5	6.2	3.9	3.9
Glutamine(Q)	4.3	3.8	3.0	3.0	1.9	2.8	1.9	3.9
Arginine(R)	5.8	4.5	4.0	4.0	3.9	1.4	2.7	2.5
Serine(S)	7.2	5.1	7.0	7.0	5.8	5.5	4.7	6.2
Threonine(T)	7.2	7.1	5.0	5.0	5.8	6.2	5.8	5.1
Valine(V)	4.3	3.2	6.0	6.5	7.1	6.2	8.9	7.3
Tryptophan(W)	2.2	1.3	2.0	2.0	2.6	0	2.3	2.0
Tyrosine(Y)	5.0	2.6	1.5	1.5	1.9	0.7	2.3	2.2

Table 3: Hydrophilic and hydrophobic residues content

Accession Number	Hydrophobic residues (%)	Hydrophilic residues (%)	Net hydrophobic residues content
A8D6D7	48.9	30.9	High
A8D7D8	45.5	25	High
B2ZA12	52	23	High
B2ZA16	52	23	High
B2ZA17	52.6	21.4	Very high
B6VQV5	62.1	20.7	Very high
B6VQV6	53.3	21	Very high
D3KYA1	51.7	24.2	Very High

Table 4: Physicochemical characters as predicted by ExPASy's prot-param program

Uniprot ID	Protein	Sequence length	M.wt.	pI	-R	+R	Extinction Coefficient	Instability Index	Aliphatic Index	GRAVY
A8D6D7	Polyphenol oxidase	139	15815.6	7.45	12	12	26930	34.50	54.03	-0.660
A8D7D8	Lipoxygenase	156	17511.9	6.82	19	18	17085-16960	57.08	91.35	-0.494
B2ZA12	Eugenol o-methyltransferase	200	22492.1	8.38	17	19	26720-26470	48.32	86.35	-0.114
B2ZA16	Eugenol o-methyltransferase	200	22478.1	8.38	17	19	26720-26470	46.93	85.85	-0.112
B2ZA17	Eugenol o-methyltransferase	154	17386.9	6.39	16	13	26595-26470	48.62	78.57	-0.161
B6VQV5	Phenylalanine ammonia-lyase	145	15121.5	5.28	13	8	1490-1490	30.71	118.48	0.423
B6VQV6	Chavicol o-methyltransferase	257	28569.8	5.22	34	23	42315-41940	30.88	89.88	-0.003

D3KYA1	Chavicol o-methyltransferase	356	39950	5	42	30	50795-50420	38.34	92.30	-0.038
--------	------------------------------	-----	-------	---	----	----	-------------	-------	-------	--------

Table 5: Presence of disulphide (ss) bond as predicted by Cys_Rec

Accession number	CYS_REC	Score
A8D6D7	No cystein found	-
A8D7D8	Cys_114	-6.8
	Cys_124	-18.4
B2ZA12	Cys_5	-7.9
	Cys_12	-17.4
	Cys_129	-31.4
	Cys_151	-31.4
B2ZA16	Cys_5	-7.9
	Cys_12	-17.4
	Cys_129	-28.6
	Cys_151	-34.8
B2ZA17	Cys_83	-25.8
	Cys_105	-33.0
B6VQV5	Cys_17	-24.7
B6VQV6	Cys_6	-29.3
	Cys_62	-24.1
	Cys_84	-26.7
	Cys_130	-36.6
	Cys_186	-36.7
	Cys_225	-27.0
D3KYA1	Cys_37	-25.0
	Cys_97	-16.8
	Cys_153	-26.7
	Cys_175	-25.6
	Cys_221	-42.8
	Cys_277	-29.2
	Cys_316	-35.5

Table 6: Calculated secondary structure elements by Profunc

Secondary Structure	B2ZA12	B2ZA17	B6VQV6	D3KYA1	B2ZA16	A8D7D8	A8D6D7	B6VQV5
Strand	5.0%	2.6%	11.7%	13.8%	6.5	12.2%	0.0%	5.5%
Alpha helix	58.5%	48.1%	46.7%	45.2%	43.5%	18.6%	38.1%	53.8%
3-10 helix	1.5%	1.9%	2.3%	1.4%	1.0%	5.1%	2.2%	2.8%
Other	35.0%	47.4%	39.3%	39.6%	49.0%	64.1%	59.7%	37.9%
Total residues	200	154	257	356	200	156	139	145
Beta sheets	1	1	2	2	1	1	-	1
Beta alpha beta motifs	-	-	4	3	2	-	-	-
Strands	3	2	7	9	3	4	-	3
Helices	15	11	15	19	11	6	7	7
Helix-helix interactions	25	13	20	23	8	1	6	8
Beta turns	8	14	14	25	16	16	14	8
Gamma turns	3	3	1	3	3	5	3	-
Beta hairpins	1	1	-	2	-	3	-	2
Beta bulges	-	-	-	2	-	1	-	1

Table 7: Protein submitted in PMDB

Uniprot ID	Protein name	PMID
B6VQV6	chavicol o-methyltransferase	PM0077211
D3KYA1	chavicol o-methyltransferase	PM0077212
A8D6D7	Polyphenol oxidase	PM0077059
A8D7D8	Lipoxygenase	PM0077191
B2ZA16	Eugenol o-methyltransferase	PM0077192
B2ZA17	Eugenol o-methyltransferase	PM0077201
B6VQV5	Phenylalanine ammonia-lyase	PM0077206