CHAPTER 7
ISOLATION AND CHARACTERIZATION OF ETHYLENE RELATED GENES IN POLLINATED *DENDROBIUM* POMPADOUR

7.1 INTRODUCTION

In many flowers pollination causes an initial, dramatic increase in ethylene production in the stigma and style and a subsequent increase in ethylene production by other floral organs (Larsen *et al.*, 1995; O'Neill *et al.*, 1993; Zhang and O'Neill, 1993). Ethylene is synthesized by plants through the conversion of S-adenosyl-L-Met to ACC, which is then oxidized to ethylene (Adams and Yang, 1979). The former is catalyzed by the enzyme ACCS, while the latter is catalysed by the enzyme ACCO. The importance of ACCO in ethylene regulation was demonstrated in transgenic tomatoes with antisense copies of ACCO where ethylene production was greatly reduced (Hamilton *et al.*, 1990). Studies on *Phalaenopsis* showed that the production of ACCO increased following pollination alongside the increase in ethylene (Nadeau *et al.*, 1993). Thousands of genes have been cloned from different plants and plant tissues (for review, see Zarembinski and Theologis, 1994; Klee and Clark, 2010). The role of ACCS during pollination in orchids has also been investigated. Bui and O'Neill (1998) reported increase in ACCS production especially in the ovary. They hypothesized that a portion of the ACC synthesized in the ovary is available for translocation to other floral organs, especially the perianth. The control of these enzymes may also be key in controlling ethylene synthesis in climacteric plants. Thus, many biochemical studies have been carried out to characterize this protein.

In concert with these biochemical studies, this chapter looks into the molecular characterization of this important enzyme. To complete the process of ethylene regulation, ethylene is then perceived by a group of ethylene receptors. They function as negative regulators as the binding of ethylene results in the inactivation of receptor function. The first and most widely studied ethylene receptor is the Ethylene Receptor 1, (*ETR1*), reported in *Arabidopsis* (Bleeker *et al* , 1988 and Chang *et al.,* 1993), tobacco (Knoester *et al.,* 1998), rice (Yau and Yip, 1997) and *Phalaenopsis* (Do *et al.,* 1999). Subsequently, *ETR2, ERS1, ERS2* and *EIN4* were also identified as the ethylene receptors involved in ethylene perception (Hua *et al.,* 1995, 1998; Sakai *et al.,* 1998; Harada *et al.,* 2011). Ethylene receptor genes with homology to receptors found in *Arabidopsis* have also been identified in flowers such as carnation, rose, delphinium and geranium. All ethylene receptors across species are predicted to contain three N-terminal transmembrane domains, a GAF domain and a kinase domain.

The aim of this chapter is to characterize these important genes using bioinformatics tools. With the progress of bioinformatics, many programmes and tools have been developed to allow swift data generation, processing, analysis and storage. Various bioinformatics centers and programmes offer free of charge tools for application in research and academics. In this chapter, the main source for gene retrieval used is the National Center for Biotechnology (NCBI) while analysis of the genes are carried out using tools available at the ExPASy Proteomics Server maintained by the Swiss Institute of Bionformatics.

7.2 MATERIALS AND METHOD

7.2.1 Plant Material

D Pompadour flowers were obtained from the glasshouse of University of Malaya. Flowers were hand-pollinated stigmas explained in Chapter 3. Individual flowers were cut at the proximal end of the peduncles in water and placed in 20 ml water vials containing distilled water.

7.2.2 Extraction of RNA

Total RNA was extracted using GeneTACG RNA extraction kit. Approximately 0.1 g of flower material was ground using a mortar and pestle under liquid nitrogen and 500µl of PRX extraction buffer was added. The lysate was mixed and vortexed thoroughly and transferred to a shearing tube and centrifuged at 13,000 rpm for two minutes. The flow through was collected and the volume measured. Approximately half volume of absolute ethanol was added and the solution was mixed by gentle pippetting. The mixture was then transferred to a mini column and centrifuged at 10,000 rpm for two minutes. The flow through was discarded and 500µl of WF wash buffer and tubes were centrifuged at 13,000 for 1 minute. Flow through was discarded and 700 µl of WS wash buffer was added and tubes were centrifuged at 13,000 rpm for one minute. Washing step with WS buffer was repeated. Tubes were then centrifuged at 13,000 rpm for three minutes to discard any left-over wash buffer. The mini column was then transferred to an elution tube and 30µl of nuclease free water was added to elute the RNA at full speed for two minutes. RNA was stored at -20 until further usage.

7.2.3 Primer Design

Primers were designed for the amplification of *ACCO, ACCS, ETR1, ERS1* and *ERS2.* When primers were designed from conserved regions of existing genes, multiple gene

sequences from a number of species were retrieved from the Genbank and subsequently aligned using the ClustalW tool available on the Expert Protein Analysis System (ExPASy) website :www.expasy.org. Primers were then designed based on the conserved nucleotides of the multiple aligned sequences. The following primers were used for gene amplification using RT-PCR.

i) *ACCO* primers

Forward primers
5’ CTGAATAACTATATGGCAGTACTG 3’

Reverse primers
5’CCTTATAGCCTCCAGGGGTAAT3’

ii) *ACCS* primers

Forward primers
5’ TGCTCTCGAAGATAGCTACTAACA 3’

Reverse primers
5’ GATTGTCAAGAGCCAGGATCT 3’

iii) *ETR1* primers

Forward primers
5’ GCGTGTTTGGCAGCAGCCGCGCGG 3’

Reverse primers
5’ TGAGCCTGATTGTGAGCCTGA 3’

iv) *ERS1* primers

Forward primers
5’ TTAAGGGGGGGGGGCTTGTT 3’

Reverse primers
5’ TTAAGGGGGGGGGGCTTGTT 3’

v) *ERS2* primers

Forward primers
5’ ATTTTAGCATTCCGCTGGAACTGA 3’

Reverse primers
5’ TGCTGATGTTTCCGAGCGATA3’

7.2.4 Preparation of RT-PCR Reaction Mixture

RT- PCR (PromegaAccessQuick™) was carried out using a reaction mixture summarized below:

Component	Final volume
AccessQuick™ Master Mix 2X	25 μl
Forward primer (5μg μl^{-1})	1 μl
Reverse primer (5μg μl^{-1})	1 μl
RNA template (1μg μl^{-1})	2 μl
Nuclease free water	20 μl

The reaction mixture above was thoroughly mixed and 1μl of AMV Reverse Transcriptase was added and mixed by gentle vortexing.

7.2.5 RT-PCR Amplification of Genes

RT-PCR was carried out using a GeneAmp 9600 Thermocycler (Perkin Elmer). Amplification cycles were carried out as described below:

Reverse transcription (1 cycle)	48^{o}C for 45 minutes
PCR amplification:	
Denaturation (1 cycle)	95^{o}C for two minutes
Annealing (40 cycles)	95^{o}C for 30 seconds
	55^{o}C for one minute
	72^{o}C for one minute
Extension (1 cycle)	72^{o}C for seven minutes
Soak cycle (hold)	4^{o}C overnight

cDNA was then analysed on a 1% electrophoresis and sized using a GeneMass DNA ladder (Fermentas). Purification of cDNA was carried out using the Wizard SV Gel and PCR Cleanup System (Promega).

7.2.6 Agarose Gel Preparation

Electrophoresis of all the RT-PCR products obtained was carried out on a 1% (w/v) agarose gel in 1x TAE buffer. A solution of 250 ml of 1 %(w/v) agarose was made by adding 2.5 g of agarose to 5 ml of 50x TAE (242.1 g of Tris, 100 ml of 0.5 M EDTA, pH 8.0 adjusted with glacial acetic acid) boiled and made up with sterile deionized water to its final volume. The gel solution was then cooled down to 50°C and mixed with 2.5 µl of 10 µg μl^{-1}ethidium bromide before pouring into a gel casting tray to solidify. The gel was electrophoresed in 2 L of 1x TAE running buffer containing 80-100µl ethidium bromide (10µg ul^{-1}).

7.2.7 Purification of RT-PCR product

RT-PCR product was gel purified using the Wizard^{R} SV gel and PCR purification Clean-up system from Promega. The RT-PCR product was added to an equal volume of Membrane Binding Solution and the mixture was transferred to SV mini-column and centrifuged for five min at 12000 rpm. The column was then removed and transferred to a fresh Eppendorf tube and the purified product was eluted using the PCR Clean-Up system.

7.2.8 Sequencing and Characterization of Isolated cDNA

Purified cDNA obtained from RT-PCR was commercially sequenced by First Base Laboratories Sdn Bhd, Selangor, Malaysia.

7.2.9 Bioinformatic Analysis Tools

DNA sequences were analysed using bioinformatic tools available online.

7.2.9.1 Translate

Translated amino acid sequence was obtained using the Translate tool available at the ExPASy site. The output format included amino acid and nucleotide sequences.

7.2.9.2 Prosite

The finding of biologically relevant sites and signatures in a sequence was done using Prosite programme which allows scanning of protein sequences against PROSITE databases (Castro *et al.*, 2006). Detection of these similar sites gives clues to the function of a sequence or part of a sequence, as highly conserved regions are likely to be active sites.

7.2.9.3 Pfam

This software allows for the detection of homologous regions between knowns sequences stored in the database and the query sequence. Pfam gives information regarding important domains of a sequence and subsequently the family it belongs to.

7.2.9.4 BLAST NCBI

BlAST NCBI was accessed via the NCBI (National Center for Bioinformatics USA) website; www.ncbi.com. The programme is widely used for searching DNA and protein databases for sequence similarities. Protein blast or blastp is used in this study to search a number of non- redundant protein databases (Swiss Prot, PDB, PIR, PRF) for similarities with the query sequence.

7.2.9.5 ProtParam

ProtParam is a software used in this study to analyze the physico-chemical properties of the sequence of interest (Gasteiger *et al.,* 2005). In this chapter, the parameter analysed were molecular weight, theoretical PI, three major amino acids and protein stability.

7.2.9.6 ClustalW

ClustalW was used to perform multiple sequence alignment, which allows identification of conserved amino acid regions, and detects relationships between the proteins of interest (Chenna *et al.,* 2003). ClustalW by default uses a number of symbols to denote the degree of conversion; "*" (all residues are identical), ":" (conserved substitutions observed) and "." (semi-conserved substitutions observed). A distance matrix phylogenetic tree was generated from the resultant sequence alignment using the neighbour joining method.

7.2.9.7 Swiss Model

This is a server for automated comparative modelling of three dimensional (3D) protein structures (Schwede *et al.,* 2003). This involved generating a set of structures representative of most of the possible folds for specific protein domains and then solving the structures for new proteins based on known fold- structure relationships.

7.2.9.8 Rasmol

This programme interactively displays the molecule on the screen in a variety of colour schemes and molecular representations. Molecules can be shown as wire frame bonds, cylinder *Dreiding* stick bonds, alpha –carbon trace, space filling, spheres and macromolecular ribbons.

7.3 Results

7.3.1 *ACCO* gene of *D.* Pompadour

7.3.1.1 RT-PCR results of *ACCO* gene of *D.* Pompadour

RT-PCR using specific primers for the *ACCO* gene resulted in a 928bp product. A band indicating the presence of the product was visible on a 1% agarose gel (Fig 7.1).

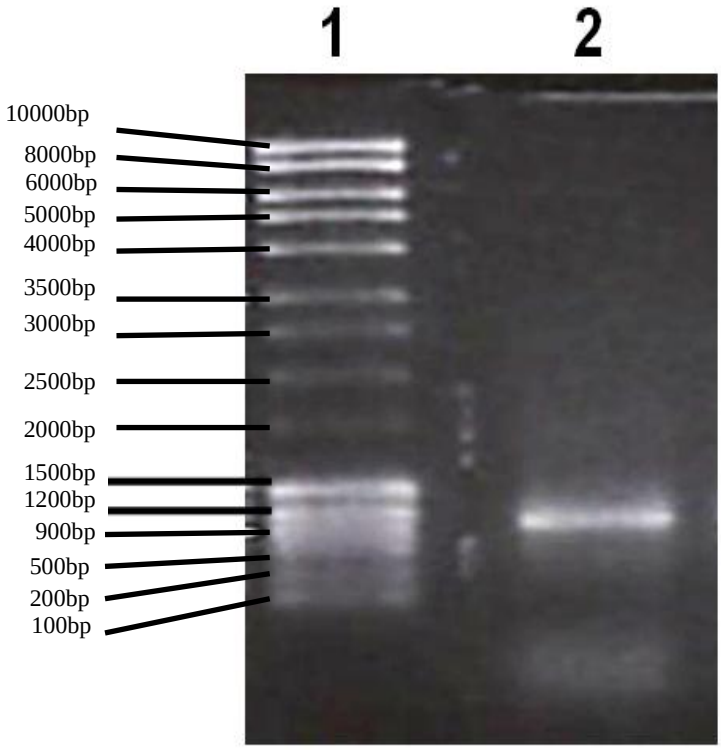


Fig 7.1 RT-PCR results with *ACCO* primers. In Lane 1 is the GeneRuler DNA ladder Mix (Fermentas). Lane 2 is the PCR product of 928bp *ACCO* gene analysed through 1% agarose gel.10 µl of ACCO was loaded into the well while 5 µl of ladder was used.

7.3.1.2 Nucleotide and amino acid sequence of *ACCO* gene of *D.* Pompadour

Fig 7.2 shows the nucleotide sequence of *ACCO* and the corresponding amino acid. The forward and reverse primers are italicized in red.

```
1       CTGAATAACTATATGGCAGTACTGAGGGACGCCTGTGAGAACTGGGGCTT
51      TTTTTCGCACTTTCAGGCGACACTCTCTCACGAGTTGATGGACAAAGTGG
101     AGCATATTAGTGAAAGGTATATCAAAAAATTCCGAGAGCAAAAATTCAAG
151     GATATATGTAGCCATAACTCAGGAAATTTCTGTAACACACAGGTGAATGA
201     TAGACTGGGAATCCACGTTTTACCTTAGACACCGACCAACTTCAAACATC
251     TCGAAGAGCCCCGAAGATCTCGACGATCAATACAGAAAACTCATGAAAGA
301     ATTCGCGGCACAAATCGAACGCTTAAGTGAACAGCTATTAGACTTACTTT
351     GCGAGAATCTTGGGCTAGAGAAGGCATACCTCAAGAATGCATTTTATGGC
401     GCTAATGGATGCAATCCAACTTTCGGAACCAAGGTGTCTAATTACCCCCC
451     ATGCCCGAAGCCTGAACTGATAAAGGGTCTGCGGGCGCATACGGACGCGG
501     GTGGAATTATTCTATTGTTTCAGGACGATACAGTCTCCGGGTTACAATTG
551     TTGAAGGATGAAGAGTGGATTGATGTTCCCCCAATGCGTCACAGCGTTGT
601     TGTAAACCTTGGCGACCAACTAGAGGTCATTACCAATGGAAAGTATAAGA
651     GTGTCATGCATCGTGTAATCGCTCAAACAGATGGTAACCGCATGTCAATA
701     GCCTCCTTTTATAACCCTGGGTCGGATGCAGTTATATATCCGGCTCCCAC
751     ATTAGTGGAGAAAGAAAAGGAGACCTACCCGAAATTCGTATTTGAGGATT
801     ATATGAAATTATACGTAAGACAGAAATTCGAAGCCAAAGAACCTCGGTTC
851     GAAGCTATGAAGACTATGGACGCCGTCATATCTAGTCAGCCTATCCCGAC
901     GGCACCATTACCCCTGGAGGCTATAAGG
```

```
MEKLNNYMAVLRDACENWGFFSHFQATLSHELMDKVEHISERYIKKFREQKFKDICSH
NSGNFCNTQVNDIDWESTFYLRHRPTSNISKSPEDLDDQYRKLMKEFAAQIERLSEQL
LDLLCENLGLEKAYLKNAFYGANGCNPTFGTKVSNYPPCPKPELIKGLRAHTDAGGII
LLFQDDTVSGLQLLKDEEWIDVPPMRHSVVVNLGDQLEVITNGKYKSVMHRVIAQTDG
NRMSIASFYNPGSDAVIYPAPTLVEKEKETYPKFVFEDYMKLYVRQKFEAKEPRFEAM
KTMDAVISSQPIPTAPLPLEAIRGGGYPGDL
```

Fig 7.2 Nucleotide sequence of *ACCO* from *D.*Pompadour and the corresponding amino acid as translated using the Translate tool accessed via www.expasy.org. The forward and reverse primers are printed in red.

7.3.1.3 Homology of *ACCO* protein of *D.* Pompadour

BLAST homology search at www.ncbi.nlm.nih.gov/blast revealed the following match in the database:

1. 567/799 positive identities (71%) (4e-103) with *ACCO* of *Dianthus caryophylus*(a)
2. 482/681 positive identities (71%) (3e-85) with *ACCO* of *Nicotiana tabacum* (b)
3. 486/691 positive identities (70%) (2e-80) with *ACCO* of *Hevea brasiliensis* (c)
4. 489/698 positive identities (70%) (7e-81) with *ACCO* of *Prunus persica* (d)
5. 492/698 positive identities (70%) (2e-80) with *ACCO* of *F*ragaria ananassa (e)
6. 478/681 positive identities (70%) (8e-80) with *ACCO* of *Nicotiana suaveolens*(f)
7. 477/675 positive identities (71%) (3e-78) with *ACCO* of *Lactuca sativa* (g)
8. 487/698 positive identities (70%) (3e-78) with *ACCO* of *Prunus domestica* (h)
9. 448/629 positive identities (71%) (4e-77) with *ACCO* of *Antirrhinum majus* (i)
10. 470/686 positive identities (69%) (3e-73) with *ACCO* of *Solanium lycopersicum* (j)

(a) Accession M62380	Wang and Woodson, 1993.
(b) Accession HQ418208	Moniuszko *et al.*, 2010.
(c) Accession AY207387	Kongsawadworakul and Chrestin, 2002.
(d) Accession AF319166	Callahan *et al.*,1993.
(e) Accession AY706156	Kiss *et al.*, 2006.
(f) Accession DQ984136	Roeder *et al*., 2009.
(g) Accession AB158345	Takahashi *et al.*, 2010.
(h) Accession AJ890086	Fernandez *et al.*, 2002.
(i) Accession AY333926	Woltering and Nijenhuis, 2003.
(j) Accession AK324411	Aoki *et al.*, 2010.

7.3.1.4 Features of *ACCO* protein of *D.* Pompadour

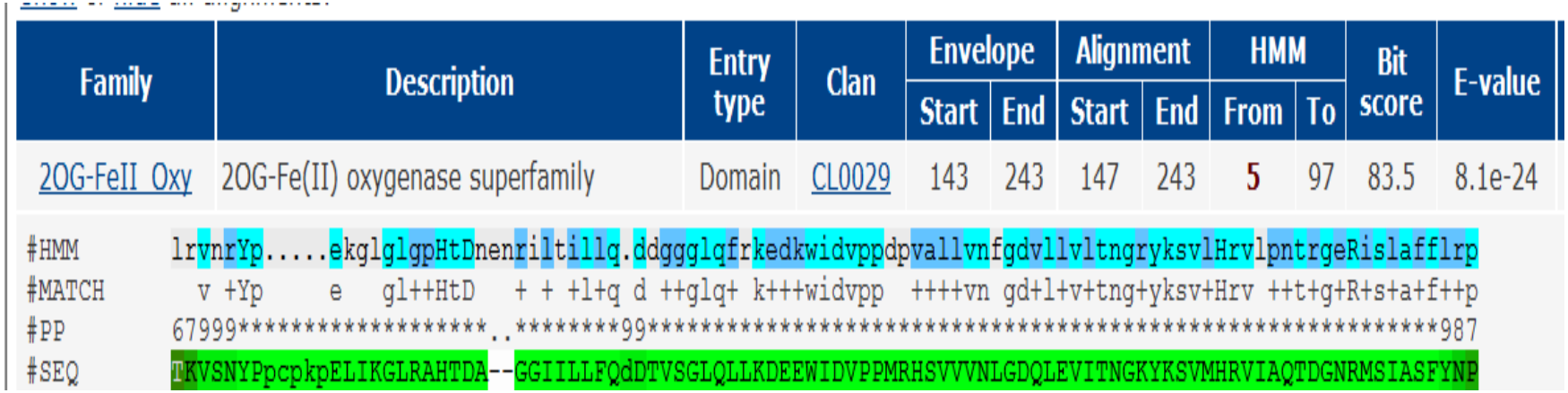
Analysis with Prosite resulted in the detection of four different sites; N-glycosylation, protein kinase C phophorylation, Casein kinase II phosphorylation and Myristoylation. These four sites are found in almost all *ACCO* protein sequences found in the data bank, the difference being only in the number of each pattern site. Table 7.1 summarizes the sites, position of residues corresponding to the sites, the consensus pattern and the number of sites detected in *ACCO*.

Table 7.1 Protein sites detected in *ACCO* using the Prosite tool accessed via www.expasy.org.

Site	Residue	Consensus Pattern	Number of sites
N-glycosylation	86 - 89 NISK	of N-{P}-[ST]-{P}	One
Protein kinase C phosphorylation	40 - 42 SeR	[ST]-x-[RK]	One
Casein kinase II phosphorylation	90 - 93 SpeD 254 - 257 TlvE	[ST]-x(2)-[DE]	Two
Myristoylation	61 - 66 GNfcNT 125 - 130 GLekAY 137 - 142 GAngCN 140 - 145 GCnpTF 146 - 151 GTkvSN 232 - 237 GNrmSI 315 - 320 GgypGD	G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	Seven

7.3.1.5 Pfam analysis of *ACCO* protein of *D.* Pompadour

Pfam analysis of *ACCO* protein showed the presence of Fe(2+) 2-oxoglutarate dioxygenase domain profile from positions 143 – 243 amino acids. Enzymes with the Fe(2+) and 2-oxoglutarate (2OG)-dependent dioxygenase domain typically catalyse the oxidation of an organic substrate using a dioxygen molecule, mostly by using ferrous iron as the active site cofactor and 2OG as a cosubstrate which is decarboxylated to succinate and CO_2. Fig 7.4 shows the amino acid sequence of ofFe(2+) 2-oxoglutarate dioxygenase domain in *ACCO* protein sequence. Residues conserved in the domain were detected which are Y, H, D, H and R.



Family	Description	Entry type	Clan	Envelope		Alignment		HMM		Bit score	E-value
				Start	End	Start	End	From	To		
2OG-FeII_Oxy	2OG-Fe(II) oxygenase superfamily	Domain	CL0029	143	243	147	243	5	97	83.5	8.1e-24

```
#HMM      lrvnrYp.....ekglglgpHtDnenriltillq.ddggglqfrkedkwidvppdpvallvnfgdvllvltngryksvlHrvlpntrgeRislafflrp
#MATCH      v +Yp     e   gl++HtD   + + +l+q d ++glq+ k+++widvpp  ++++vn gd+l+v+tng+yksv+Hrv ++t+g+R+s+a+f++p
#PP       67999******************..********99**********************************************************987
#SEQ       TKVSNYPpcpkpELIKGLRAHTDA--GGIILLFQdDTVSGLQLLKDEEWIDVPPMRHSVVVNLGDQLEVITNGKYKSVMHRVIAQTDGNRMSIASFYNP
```

MEKLNNYMAVLRDACENWGFFSHFQATLSHELMDKVEHISERYIKKFREQKFKDICSHNSGNFCNTQVNDIDWESTFYLRHRPTSNISKSPED
LDDQYRKLMKEFAAQIERLSEQLLDLLCENLGLEKAYLKNAFYGANGCNPTFGTKVSNYPPCPKPELIKGLRAHTDAGGIILLFQDDTVSGLQ
LLKDEEWIDVPPMRHSVVVNLGDQLEVITNGKYKSVMHRVIAQTDGNRMSIASFYNPGSDAVIYPAPTLVEKEKETYPKFVFEDYMKLYVRQK
FEAKEPRFEAMKTMDAVISSQPIPTAPLPLEAIRGGGYPGDL

Fig 7.3 Results from Pfam analysis show the information on the Fe(2+) and 2-oxoglutarate (2OG)-dependent dioxygenase domain. The #HMM line shows the consensus of the model, with capital letters representing the most conserved (high information content) positions, and dots (.) indicating insertions in the query sequence. Identical residues are coloured cyan, and similar residues are coloured dark blue. The #MATCH line indicates matches between the model and the query sequence, where a + indicates a "conservative substitution"; the #PP line represents the expected accuracy (posterior probability) of each aligned residue, where a 0 means 0–5%, 1 means 5–15%, and so on to 9 meaning 85–95% and a * meaning 95–100% posterior probability (pp); the #SEQ line is the query sequence, coloured according to the pp for each residue match on a scale from bright green for * through paler green and pale red down to bright red for 0. The Fe(2+) and 2-oxoglutarate (2OG)-dependent dioxygenase in the deduced *ACCO* protein is highlighted in yellow.

7.3.1.6 ProtParam analysis of *ACCO* protein of *D.* Pompadour

Table 7.2 summarises the physical and chemical properties of *ACCO* as analysed using the ProtParam programme accessed via the ExPASy website. The results show that *ACCO* protein of *D. Pompadour* share similar properties with other species. These include molecular weight within the range of 30-37 kDa and a theoretical PI that is acidic. The major amino acids in the *ACCO* protein of *D.* Pompadour, leucine and lysine are also major amino acids in *Dtps sp, P. sp, D. caryophyllus and Rosa* cultivar.

Table 7.2 Summary of protein parameters of *ACCO* protein analysed using ProtParam

Species	Amino Acid	Molecular Weight	PI	Major amino acid composition			Stability
Dendrobium Pompadour	321	36696.8	5.7	Leu (9%)	Lys (7.5%)	Asp (6.2%)	Stable
Doritaenopsis sp	317	36171.3	5.8	Leu (11%)	Glu (10.1%)	Lys (7.6%)	Unstable
Phalaenopsis sp	325	37131.4	5.2	Leu (10.5%)	Glu (9.5%)	Lys (6.8%)	Unstable
Dianthus caryophyllus	321	36819.2	6.03	Leu (9.3%)	Lys (8.7%)	Glu (7.5%)	Stable
Rosa cultivar	277	31663.1	5.12	Leu (10.5%)	Glu (9.0%)	Lys (8.7%)	Stable

7.3.1.7 Multiple sequence analysis of *ACCO* protein of *D.* Pompadour

Multiple sequence alignment of *D.* Pompadour *ACCO* protein with *ACCO* protein from Rosa hybrid, *D. caryophylus*, *Dtps. sp*, and *P.* sp. and *D. crumenatum* using the ClustalW program (Fig 7.4) revealed the presence of four conserved regions, each region containing 10 or more consecutive conserved residue. Within these conserve regions, a Fe(2+) and 2-oxoglutarate (2OG)-dependent dioxygenase domain was present. Five conserved residues were also present in this domain.

A phylogenetic analysis of the relationship of *D.* Pompadour *ACCO* protein with ACCO proteins from Rosa hybrid, *D. caryophylus*, *Dtps sp*, and *P.* True Lady and *D. crumenatum* was also carried out using the ClustalW program and is presented in Fig 7.5. The phylogenetic tree derived showed that *D.* Pompadour *ACCO* is clustered with *Dianthus caryophyllus and Rosa* hybrid. However, the *ACCO* from *D.* Pompadour is more closely related to that of *D. caryophyllus*. *ACCO* proteins of *Dtps. sp* and *P.* True Lady share a very close relationship while *D.* Anna and *D. crumenatum* also share the same cluster. The *ACCO* protein from *D.* Pompadour seems to diverge in evolution as it belongs to a different cluster than the orchids.

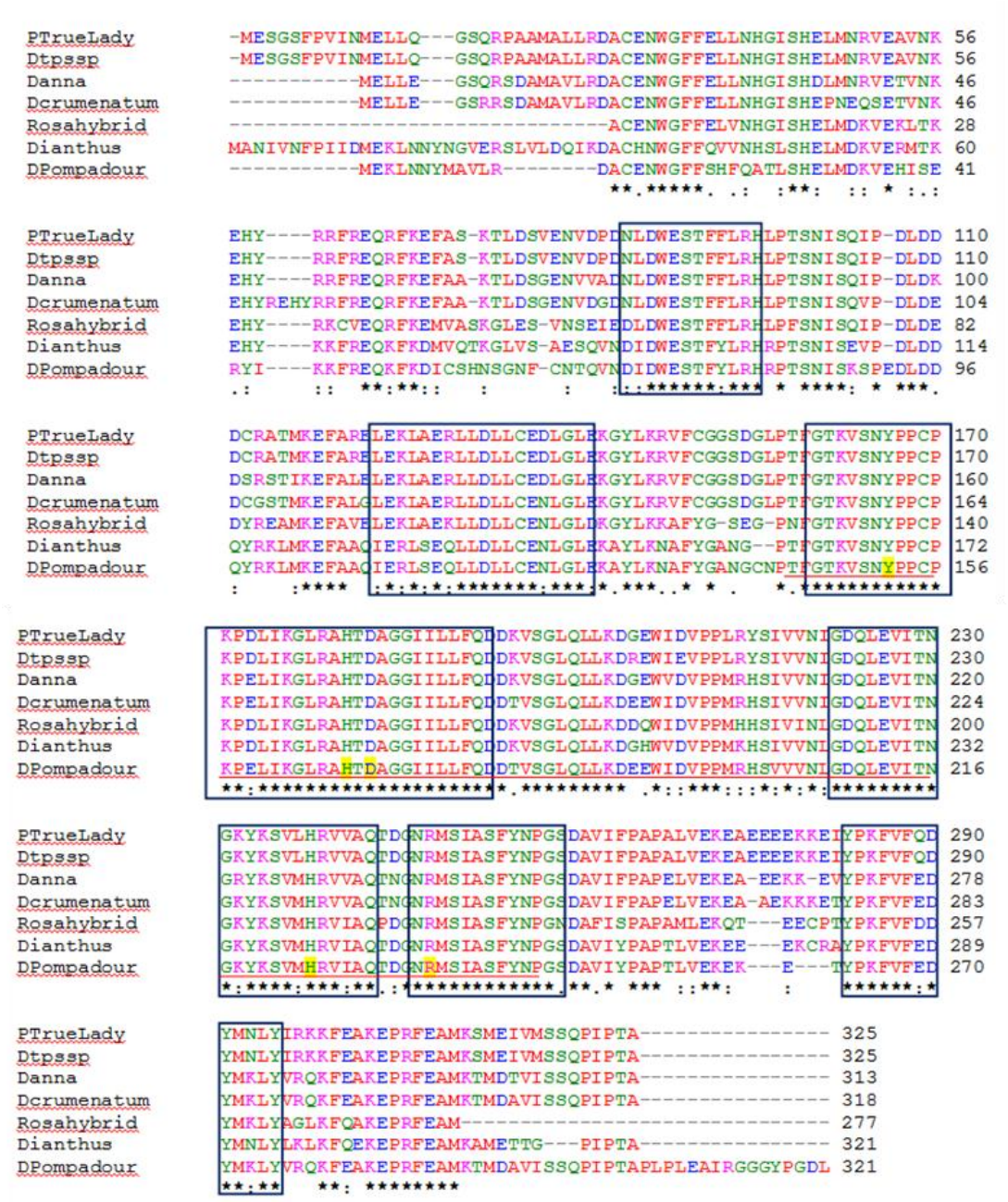


Fig 7.4 Comparison of the deduced amino acid sequence of *D.* Pompadour *ACCO* and sequences encoding *ACCO* from *P.* True Lady (AF004662) *Dtps. Sp* (L37103), *D.* crumenatum (Q9ZQZ1), *D.* Anna (GQ332400) *Rosa* (AF441282), and *D. caryophyllus* (M62380) using ClustalW. Fe(2+) and 2-oxoglutarate (2OG)-dependent dioxygenase domain are highlighted while conserved regions are boxed. Residues conserved in the Fe(2+) and 2-oxoglutarate (2OG)-dependent dioxygenase domain are highlighted in yellow.

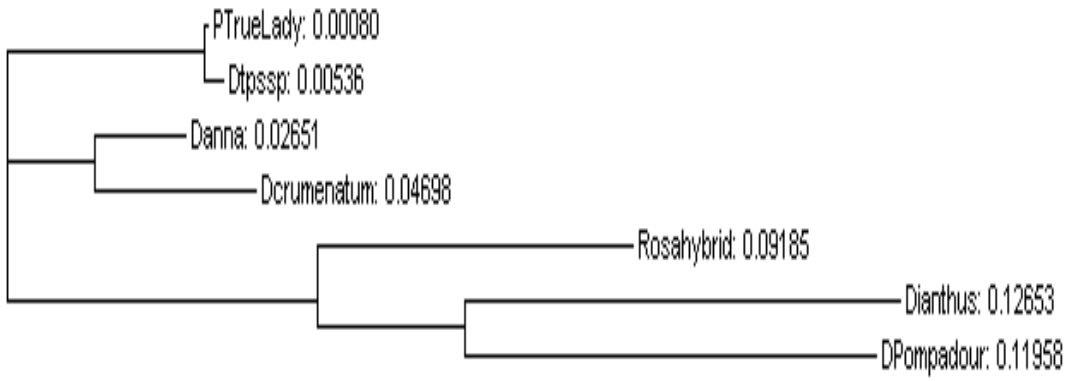


Fig 7.5 Phylogenetic tree generated from the CLUSTALW multiple alignment programme. Figure shows relationship between *D.* Pompadour *ACCO* protein and *ACCO* proteins from *D.* Anna, *D.* crumenatum, *Rosa* hybrid, *Dtps.* sp, *P.* Lady and *D. caryophyllus.* The numbers indicate the distance matrix.

7.3.1.8 Three dimensional (3D) structure of *ACCO* protein of *D.* Pompadour

Analysis using Swiss Model and Rasmol indicated that the *ACCO* protein sequence had enough functional domains to generate a 3D protein structure. The structure consisted of 286 groups, 2322 atoms and 2377 bonds. Furthermore, the programme also identified 205 H bonds, 11 helices, 15 strands and 23 turns.

Fig 7.6 3D Structure of the *ACCO* protein as generated by Swiss Model and Rasmol programmes. The structure is viewed as ribbons where secondary structures are identified. The alpha helix is represented as a coiled structure and the beta strand is represented as a pleated structure. The colour is a smooth spectrum from blue through green, yellow and orange to red. The N termini of proteins are coloured blue and the C termini, red. This shows folding from one end of the ribbon to the other.

7.3.2 *ACCS* gene of *D.* Pompadour

7.3.2.1 RT-PCR results of *ACCS* gene of *D.* Pompadour

RT-PCR using specific primers for the *ACCS* gene resulted in a 1.32Kb product. A clear band indicating the presence of the product was visible on a 1% agarose gel (Fig 7.7).

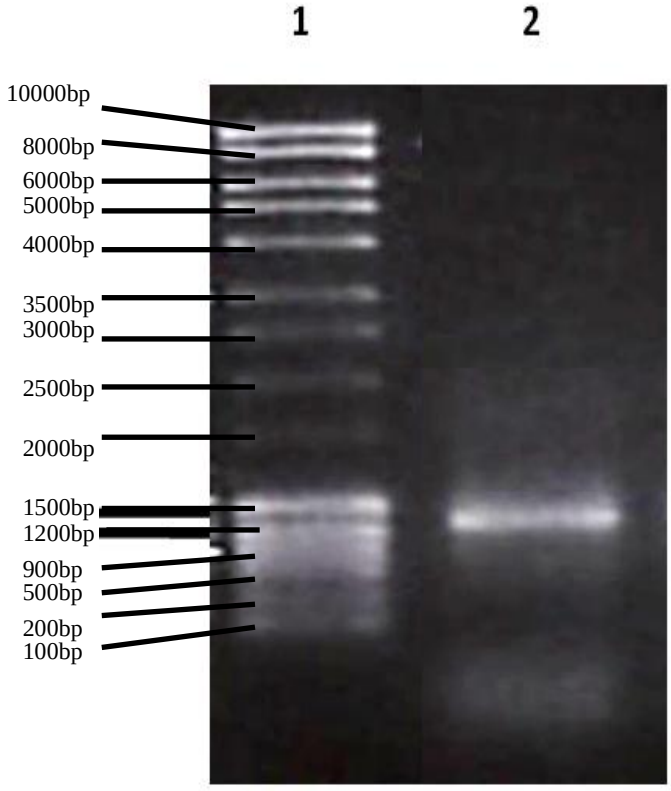


Fig 7.7 RT-PCR results with ACCS primers. In Lane 1 is the GeneRuler DNA ladder Mix (Fermentas). Lane 2 is the PCR product of 1.32Kb *ACCS* gene analysed through 1% agarose gel.10 µl of *ACCS* was loaded into the well while 5µl of ladder was used.

7.3.2.2 Nucleotide and amino acid sequence of ACCS gene of *D.* Pompadour

Fig 7.8 shows the nucleotide sequence of *ACCS* and the corresponding amino acid. The forward and reverse primers are italicized in red.

```
1     GCTCTCGAAGATAGCTACTAACATTCAACAGCTTGCGCATTCGTCTCCAA 51
      TATATATTCATCACTTGAAACTCATCGAGACAATTCGAAGGTGCCATTGC
101   CGAAAATGGCTGTGTCTAATGCTCATGGAGAGCGCTCTCGATACTTCGCT
151   GCCTGGAAAGCGTATGAAGAGATTCCTTAGGGTGTTGTAGGACATCCTGA
201   TGGAGTTAATTCATCAGTTTAGCGACATCGCTAATTTCCAAGACTACCAT
251   GGTCTTAAGAAGTTTAGACAGGCAATTGCACATTTCATGCTAAAAGCTAG
301   AGGTGGAAGAGTGACTTTTGATCCGGAGAGCGTGGTTATGAGCGCAGGAG
351   CCACCGGAGCTAATGAAACAGTCATGTTCTGCATTGCGGATCCCGGCGAC
401   GTTTTCCTCATTCCCTCCCCGTACTATGCCGCATTTGATAGAGACTTGAG
451   GTGCCGGACAGGTGCCGAGATAGTCCCGGTTCGTTGTTCATGCTCCGACA
501   ATCTCAAAATAGCCGTTGACGCGGCGGAATGCGCTTATAATTAAGCCCAA
551   GAGTCCAATAAAAAAGTCAACCGTCTGATTTTGACCAACCCATCAAATCC
601   ACTCGGTACAATGTTGGATAAGGACACACTCACGAACTTGGTCCGGTTTG
651   TCACGAGGAAGAACATTCACCTAGTCGTCGACGAGATCTACGCCGCCACA
701   GTCTTCGCCGGAGGAGATTTCGTGAGCGTTGCTGAGGTGGTCAATGATGT
751   GGACATCTCCGAAGTCAACGTTGACTTGATTCACATTGTCTATAGTCTTT
801   CTAAAGATATGGGACTTCCTGGTTTTAGAGTCGGGATAGTCTATTCTTTC
851   AATGACTCGGTCGTGTCTACGACAGAGTTGTTAAAACAGCGAGAAGAATG
901   TCAAGTTTCAGTTTGTTTCTTCTCAGACTCAAAAGTTGCTGTCTTTTATG
951   CTGTCAGATGAGGATTTACAGTGAGATATATAGAGAAGAATAGAGAGAGA
1001  CTGAGAGAGAGATATGAATTAGTTGTTAATGGGTTGAAGGAAGCAGAGCA
1051  ACGCTGGTTTATTTGCGTGGATGGATTTGAGACATCTACTGAGAGATCGT
1101  AACTCGGTCGAATCTGAGATCGAGCTTGCGCATATAATCATCGATAGAGT
1151  TAAAATCAATGTGTCTCCTGGCTCTTCCTTCCGTTGCACGGAACCTGTCT
1201  GGTTTAGGATTTGCTTTGCCAACATGATGTCATTATGTAGAGTGATTATA
1251  AACGATATGAGACTTAACGTTTCGCCTGGATCTTCATTTGATTGTCAAGA
1301  GCCAGGATCTGAAAAT
```

```
MCLPGRNIGAVLSKIATNIQQLAHSSPIYIHHLKLIETIRRCHCRKWLCLMLMESALD
TSLPGKRMKRFLRVLDILMELIHQFSDIANFQDYHGLKKFRQAIAHFMLKARGGRVTF
DPESVVMSAGATGANETVMFCIADPGDVFLIPSPYYAAFDRDLRCRTGAEIVPVRCSC
SDNLKIAVDAAECAYNAQESNKKVNRLILTNPSNPLGTMLDKDTLTNLVRFVTRKNIH
LVVDEIYAATVFAGGDFVSVAEVVNDVDISEVNVDLIHIVYSLSKDMGLPGFRVGIVY
SFNDSVVSTTELLKQREECQVSVLFLLRLKSCCLLCCQMRSLQDIRRIERDERDMNLL
MGRKQSNAGLFAWMDLRHLLRDRNSVESEIELAHIIIDRVKINVSPGSSFRCTEPVWF
RICFANMMSLCRVIINDMRLNVSPGSSFDCQEPGSEN
```

Fig 7.8 Nucleotide sequence of *ACCS* from *D.* Pompadour and the corresponding amino acid as translated using the Translate tool accessed via www.expasy.org Forward and reverse primers are printed in red.

7.3.2.3 Homology of *ACCS* protein of *D.* Pompadour

BLAST homology search at www.ncbi.nlm.nih.gov/blast revealed the following match in the database:

1. 918/1041 positive identities (88%) (0.0) with *ACCS* of *Arabidopsis thaliana*(a)
2. 797/1042 positive identities (87%) (0.0) with *ACCS* of *Brassica oleracea* var. italic (b)
3. 231/269 positive identities (97%) (2e-57) with *ACCS* of *Doritaenopsis* sp (c)
4. 230/269 positive identities (96%) (9e-56) with *ACCS* of *Phalaenopsis* True Lady (d)
5. 232/295 positive identities (88%) (2e-39) with *ACCS* of *Dendrobium* Karen (e)
6. 356/500 positive identities (71%) (9e-56) with *ACCS* of *Pyrus pyrifolia* (f)
7. 215/271 positive identities (79%) (4e-41) with *ACCS* of *Dendrobium* Anna (g)
8. 214/270 positive identities (79%) (9e-37) with *ACCS* of *Dendrobium crumenatum* (h)
9. 339/478 positive identities (71%) (3e-29) with *ACCS* of *Pyrus communis* (i)
10. 428/637 positive identities (67%) (1e-48) with *ACCS* of *Rosa* Kardinal (j)

(a) Accession NM_179241	Liang *et al.,* 1993.
(b) Accession AF338652	Gonzalez and Botella, 2001.
(c) Accession L07882	Bui and O'Neill, 1998.
(d) Accession AF004663	Huang *et al.,* 1997.
(e) Accession EF488014	Warin *et al.,* 2010.
(f) Accession EF566865	Qiao *et al.,* 2007.
(g) Accession GU138671	Nagtong *et al.,* 2009.
(h) Accession DCU64031	Yang *et al.,* 1996.
(i) Accession AF386518l	Sharkawi *et al.,* 2004.
(j) Accession AY061946	Lei *et al.,* 2002.

7.3.2.4 Features of *ACCS* of *D.* Pompadour

Analysis with Prosite resulted in the detection of seven different sites Aminotransferases class-I pyridoxal-phosphate attachment, N-glycosylation, cAMP- and cGMP-dependent protein kinase phosphorylation, Protein kinase C phosphorylation, Amidation, Casein kinase II phosphorylation site and N-myristoylation. These four sites are found in almost all *ACCS* gene sequences found in the data bank, the difference being only in the number of each pattern site. Table 7.3 summarizes the sites, position of residues corresponding to the sites, the consensus pattern and the number of sites detected in *ACCS*.

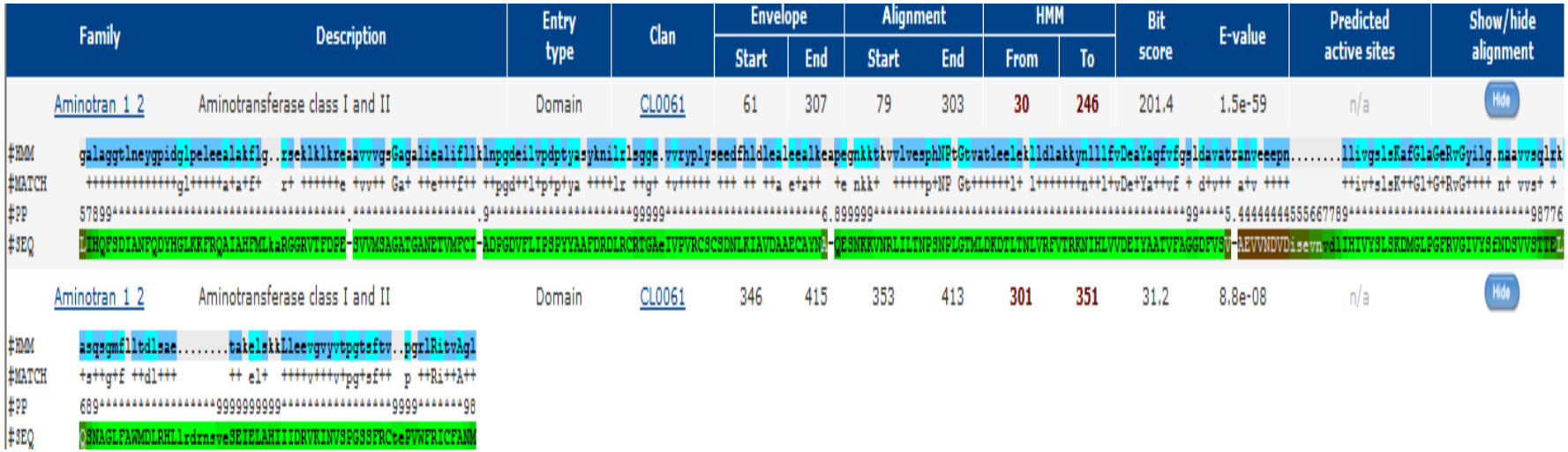
Table 7.3 Protein sites detected in ACCS using the Prosite tool accessed via www.expasy.org.

Site	Residue	Consensus Pattern	Number of sites
Aminotransferases class-I pyridoxal-phosphate attachment	274 - 287: SLSKdmGLpGFRVG	[GS] - [LIVMFYTAC] - [GSTA] - K - x(2) - [GSALVN] - [LIVMFA] - x - [GNAR] - {V} - R - [LIVMA] - [GA]	One
N-glycosylation	131 to 134 NETV 293 to 296 NDSV	N - {P} - [ST] - {P}	Two
cAMP- and cGMP-dependent protein kinase phosphorylation	351 to 354 RKQS	[RK](2) - x - [ST]	One
Protein kinase C phosphorylation	38 to 40 TIR 194 to 196 SNK 227 to 229 TRK 397 to 399 SFR	[ST] - x - [RK]	Four
Amidation	62 to 65 PGKR 349 to 352 MGRK	**x - G -** [RK] - [RK]	Two
Casein kinase II phosphorylation site	55 to 58 SALD 163 to 166 TGAE 173 to 176 SCSD	**[**ST] - x(2) - [DE]	Ten

	212 to 215 TMLD 251 to 254 SVAE 291 to 294 SFND 298 to 301 STTE 331 to 334 SLQD 376 to 379 SEIE 432 to 435 SSFD		
N-myristoylation	9 to 14 GAVLSK 111 to 116 GGRVTF 126 to 131 GATGAN 129 to 134 GANETV 287 to 292 GIVYSF	**G** - {EDRKHPFYW} - x(2) - [STAGCN] - {P}	Five

7.3.2.5 Pfam analysis of *ACCS* protein of *D.* Pompadour

Pfam analysis of *ACCS* protein showed the presence of two Aminotransferases class-I and II pyridoxal-phosphate attachment site at positions 30-246 and 301-351 amino acids. Aminotransferases share certain mechanistic features with other pyridoxal-phosphate dependent enzymes, such as the covalent binding of the pyridoxal-phosphate group to a lysine residue. All residues conserved in the domains were detected which are G, N, P, G, D, Y, K, G, G, R and G.



MCLPGRNIGAVLSKIATNIQQLAHSSPIYIHHLKLIETIRRCHCRKWLCLMLMESALDTSLPGKRMKRFLRVLDILMELIHQFSDIANFQDYH
GLKKFRQAIAHFMLKARGGRVTFDPESVVMSAGATGANETVMFCIADPGDVFLIPSPYYAAFDRDLRCRTGAEIVPVRCSCSDNLKIAVDAAE
CAYNAQESNKKVNRLILTNPSNPLGTMLDKDTLTNLVRFVTRKNIHLVVDEIYAATVFAGGDFVSVAEVVNDVDISEVNVDLIHIVYSLSKDM
GLPGFRVGIVYSFNDSVVSTTELLKQREECQVSVLFLLRLKSCCLLCCQMRSLQDIRRIERDERDMNLLMGRKQSNAGLFAWMDLRHLLRDRN
SVESEIELAHIIIDRVKINVSPGSSFRCTEPVWFRICFANMMSLCRVIINDMRLNVSPGSSFDCQEPGSEN

Fig 7.9 Results from Pfam analysis show the information on the two Aminotransferases class-I and II pyridoxal-phosphate domains. The #HMM line shows the consensus of the model, with capital letters representing the most conserved (high information content) positions, and dots (.) indicating insertions in the query sequence. Identical residues are coloured cyan, and similar residues are coloured dark blue. The #MATCH line indicates matches between the model and the query sequence, where a + indicates a "conservative substitution"; the #PP line represents the expected accuracy (posterior probability) of each aligned residue, where a 0 means 0–5%, 1 means 5–15%, and so on to 9 meaning 85–95% and a * meaning 95–100% posterior probability (pp); the #SEQ line is the query sequence, coloured according to the pp for each residue match on a scale from bright green for * through paler green and pale red down to bright red for 0. The aminotransferases class-I and II pyridoxal-phosphate domains in the deduced ACCS protein is highlighted in yellow.

7.3.2.6 ProtParam analysis of *ACCS* protein of *D.* Pompadour

The table below summarizes the physical and chemical properties of *ACCS*. The results show that *ACCS* protein of *D. Pompadour* share similar properties with other species. These include molecular weight within the range of 44-51 kDa. The theoretical PI of *ACCS* that is basic is similar to *ACCS* from *A. thaliana*. The major amino acids in the *ACCS* protein of *D.* Pompadour, Leu and Ser also major amino acids in *A. thaliana, D. crumenatum, D.* Karen, *D.* Anna and *P.* True Lady.

Table 7.4 Summary of protein parameters of *D.* Pompadour *ACCS* protein analysed using ProtParam.

Species	Amino Acid	Molecular Weight	PI	Major amino acid composition			Stability
Dendrobium Pompadour	443	49995.1	7.82	Leu (10.6%)	Val (8.1%)	Ser&Arg (7.7%)	Unstable
Arabidopsis thaliana	496	55531.5	7.20	Leu (9.1%)	Ser (8.3%)	Asp & Val (7.3%)	Stable
Dendrobium crumenatum	435	48699.6	5.47	Leu (11.5%)	Ser (9.2%)	Glu (7.6%)	Unstable
Dendrobium Karen	435	48793.8	5.98	Leu (11.0%)	Ser (9.2%)	Ala (8.0%)	Unstable
Dendrobium Anna	435	48528.7	5.90	Leu (11.5%)	Ser (9.0%)	Ala (8.0%)	Stable
Phalaenopsis True Lady	445	49666.7	6.53	Leu (9.9%)	Ser (9.7%)	Ala (7.0%)	Unstable

7.3.2.7 Multiple sequence analysis of ACCS protein of *D.* Pompadour

Multiple sequence alignment of *D.* Pompadour ACCS protein with ACCS protein from *D. crumenatum, D.* Karen, *D.* Anna, *Dtps.* Sp, *P.* true Lady, *D. caryophyllus* and *A. thaliana* using the ClustalW program (Fig 7.10) revealed the presence of Aminotransferases class-I pyridoxal-phosphate attachment site.

A phylogenetic analysis of the relationship of *D.* Pompadour *ACCS* protein with *ACCS* proteins *D. crumenatum, D.* Karen, *D.* Anna, *Dtps.* Sp, *P.* true Lady, *D. caryophyllus* and *A. thaliana* was also carried out using the ClustalW program and is presented in Fig 7.11. The phylogenetic tree derived showed that the *ACCS* proteins were divided into three clusters. *D.* Pompadour *ACCS* was clustered with *D. crumenatum, Dtps sp, P.* True Lady, *D. caryophyllus and A. thaliana.* However, *ACCS* of *D.* Pompadour had the closest relationship with *A. thaliana.*

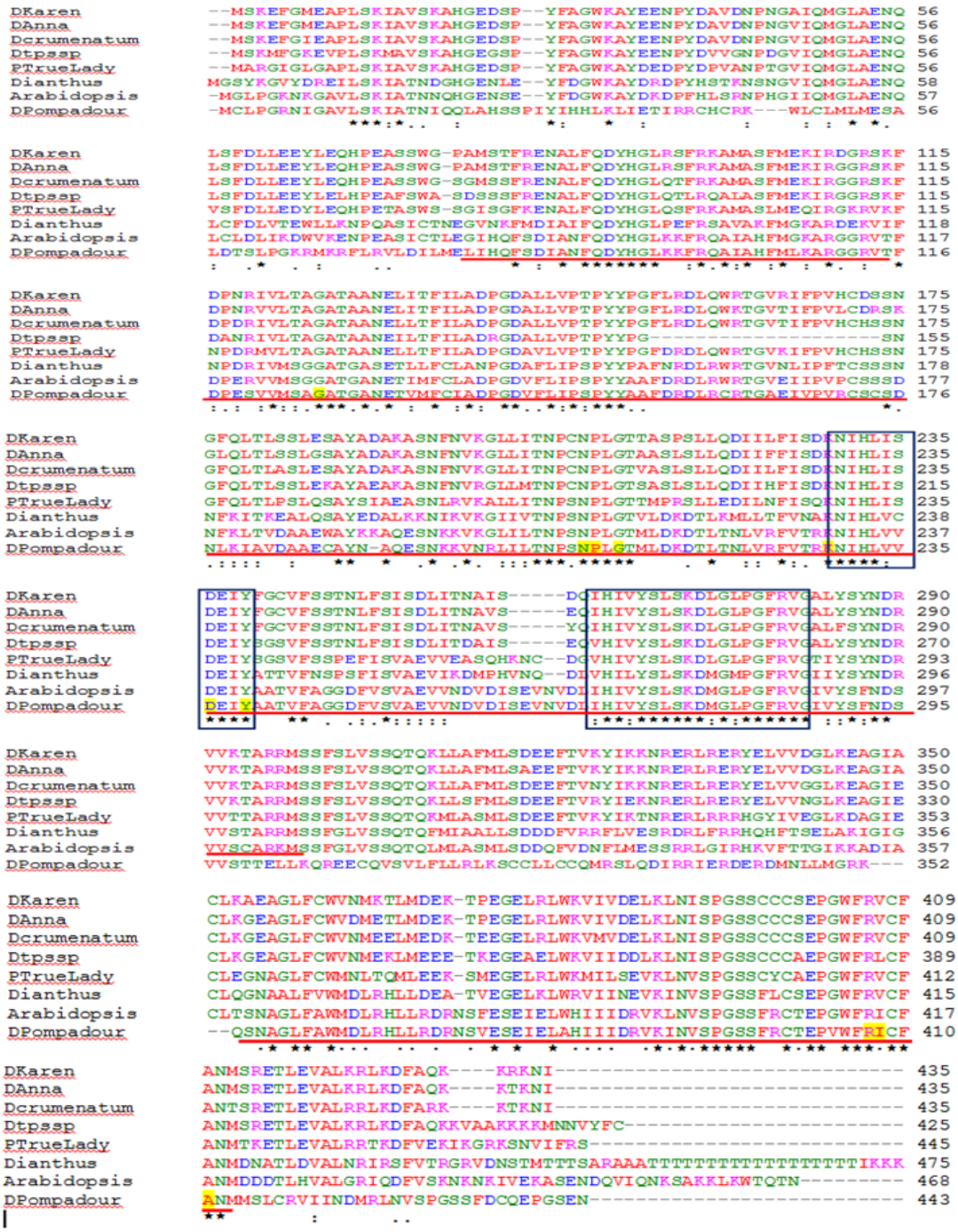


Fig 7.10 Comparison of the deduced amino acid sequence of *D.* Pompadour *ACCS* and sequences encoding *ACCS* from *D.* Karen (EF488014), *D.* Anna (GU138671) *D. crumenatum* (DCU64031), *Dtps. sp* (L07882), *P.* True Lady (Z77854), *D. caryophyllus* (Q43753) and *A. thaliana* (Q06402). The aminotransferases class-I pyridoxal-phosphate attachment site is underlined. Residues conserved in aminotransferases are highlighted in yellow.

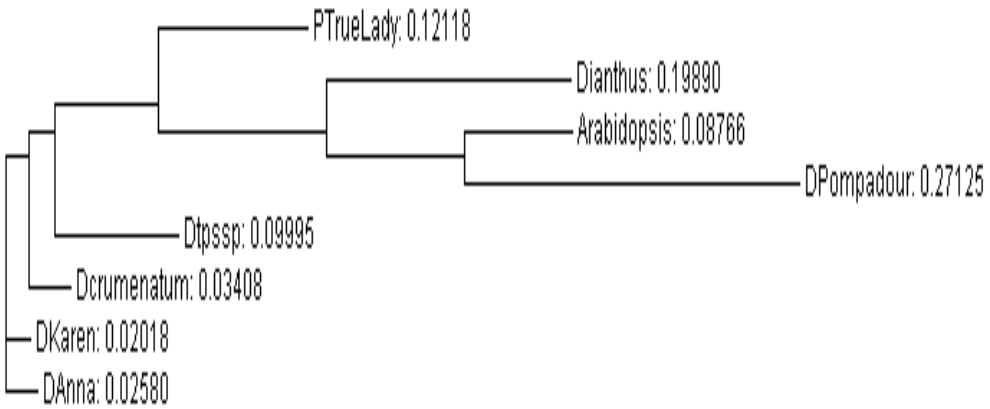


Fig 7.11 Phylogenetic tree generated from the CLUSTALW multiple alignment programme. Figure shows the relationship between *D.* Pompadour *ACCS* protein and *ACCS* proteins from *D.* Anna, *D.* Karen, *D. crumenatum*, *Dtps. sp*, *P.* True Lady, *D. caryophyllus* and *A. thaliana*. The numbers indicate the distance matrix.

7.3.2.8 Three dimensional (3D) structure of *ACCS* protein of *D.* Pompadour

Analysis using Swiss Model and Rasmol indicated that the *ACCS* protein sequence had enough functional domains to generate a 3D protein structure. The structure consisted of 420 groups, 3330 atoms and 3386 bonds. Furthermore, the programme also identified 270 H bonds, 18 helices, 14 strands and 43 turns.

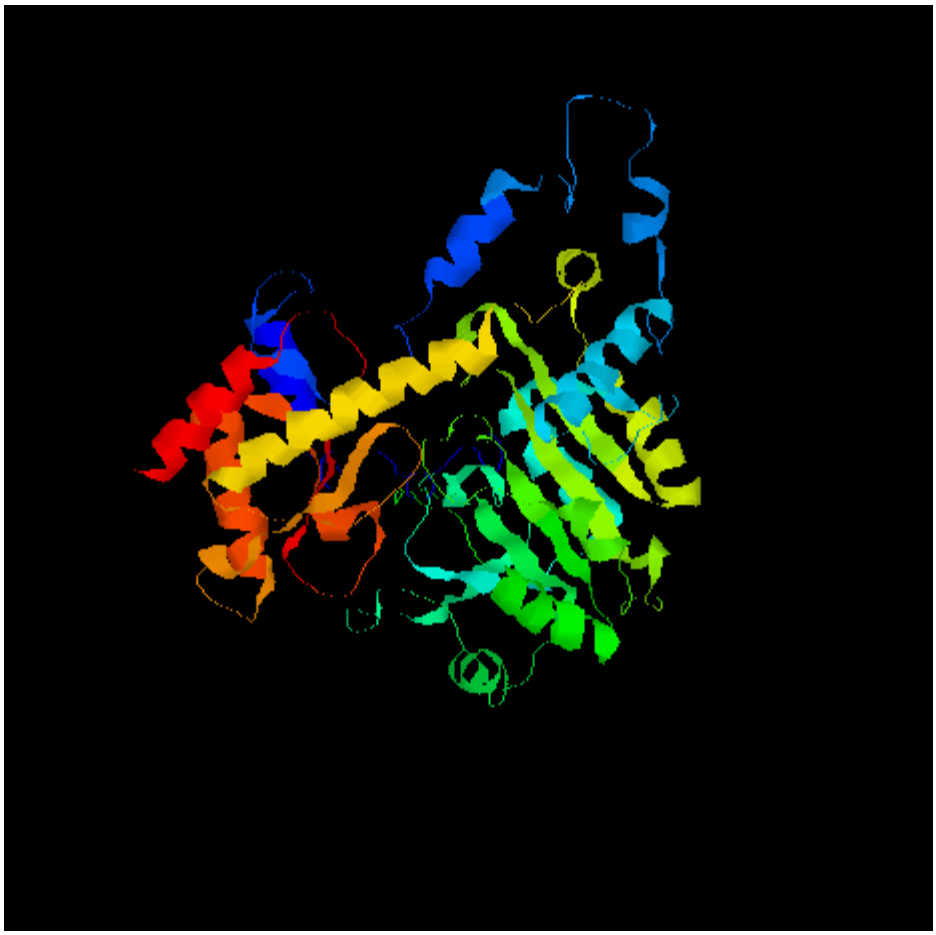


Fig 7.12 3D Structure of the *ACCS* protein as generated by Swiss Model and Rasmol programmes. The structure is viewed as ribbons where secondary structures are identified. The alpha helix is represented as a coiled structure and the beta strand is represented as a pleated structure. The colour is a smooth spectrum from blue through green, yellow and orange to red. The N termini of proteins are coloured blue and the C termini, red. This shows folding from one end of the ribbon to the other.

7.3.3 *ETR1* gene of *D.* Pompadour

7.3.3.1 RT-PCR results of *ETR1* gene of *D.*Pompadour

RT-PCR using specific primers for the *ETR1* gene resulted in a 692bp product. A clear band indicating the presence of the product was visible on a 1% agarose gel (Fig 7.13).

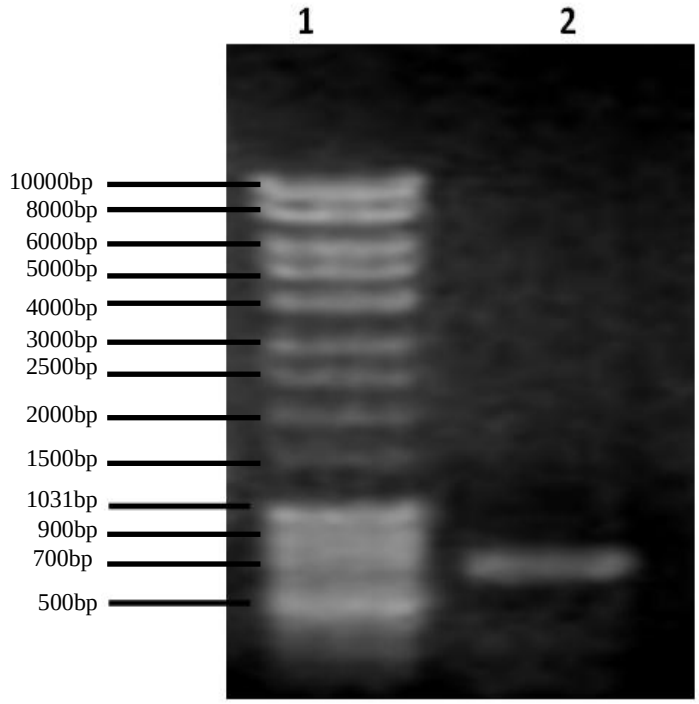


Fig 7.13 RT-PCR results with *ETR1* primers. In Lane 1 is the MassRuler DNA ladder Mix (Fermentas). Lane 2 is the PCR product of 692bp gene analysed through 1% agarose gel. 10µl of *ETR1* was loaded into the well while 5 µl of ladder was used.

7.3.3.2 Nucleotide and amino acid sequence of *ETR1* gene of *D.* Pompadour

Fig 7.14 shows the nucleotide sequence of *ETR1* and the corresponding amino acid sequence. The forward and reverse primers are italicized in red.

```
1      GCGTGTTTGGCAGCAGCCGCGCGGTGCGCATTACCCCGCATCAGCTGGCG
51     CGCAGCCAGCCGCATACCGGCCCGCAGGATATGCCGATTATTTTTCGCAA
101    ATTTACCCAGTTTAACCAGCTGAGCAACTTTCAGATTAACAACAGCAACT
151    GGCCGAGCGCGAAAAACTTTGCGGTGATGGTGCTGATGCTGCCGAGCCAT
201    AGCCGCGATAAATGGCATGTGTATGAACTGGAACTGGTGGAAGTGGTGGC
251    GCAGGCGGTGGCGGTGCATCTGAGCCATGCGGCGATTCTGGAACTGAGCA
301    TGCGCGAACAGGATCGCCAGCTGATGGAACAGAACGTGGCGCTGGATCTG
351    GCGCGCCCGAAGCGGAAATGGCGATTCGCGCGCGCAACGATTTTCTGGCG
401    GTGATGAACCATGAAATGCGCACCCCGATGCATGCGATTATTGCGCTGAG
451    CAGCCTGCTGCAGGAAACCGAACTGACCCCGGAACAGCGCCTGATGGTGG
501    AAACCATTCTGGAAAGCAGCAACCTGCTGGCGACCCTGATTAACGATGTG
551    CTGGATCTGAGCCGCCTGGAAGATGGCAGCTTTGAACTGGAAGTGACCGT
601    GTTTAACCTGCATACCGTGTTTCGCCTGGTGGCGGTGAACCTGATTAAAC
651    CGATTGCGGCGGTGAAAAAAC TGAGCCTGATTGTGAGCCTGA

ILGSVQSINLPVVNRVFNSSRAVRITPHQLARSQPHTGPQDMPIIFRKFTQFNQLSNF
QINNSNWPSAKNFAVMVLMLPSHSRDKWHVYELELVEVVADQVAVALSHAAILEESMR
ERDQLMEQNVALDLARREAEMAIRARNDFLAVMNHEMRTPMHAIIALSSLLLETELTP
EQRLMVETILKSSNLLATLINDVLDLSKLEDGSFELEVTVFNLHTVFLMVVNLIKPIA
AVDPLSLIVSLSPD
```

Fig 7.14 Nucleotide sequence of ETR1 from *D.* Pompadour and the corresponding amino acid as translated using the Translate tool accessed via www.expasy.org. Forward and reverse primers are printed in red.

7.3.3.3 Homology of *ETR1* protein of *D.* Pompadour

BLAST homology search at www.ncbi.nlm.nih.gov/blast revealed the following match in the database:

1. 369/533 positive identities (69%) (1e-50) with *ETR1* of *Phalaenopsis* sp. 'True Lady' (a)
2. 274/380 positive identities (72%) (2e-49) with *ETR1* of *Brassica* rapa subsp. Chinensis (b)
3. 275/382 positive identities (72%) (5e-49) with *ETR1* of *Pyrus* pyrifolia (c)
4. 367/533 positive identities (69%) (7e-48) with *ETR1* of *Phalaenopsis* equestris (d)
5. 274/382 positive identities (72%) (7e-48) with *ETR1* of *Malus* x domestica (e)
6. 274/383 positive identities (72%) (8e-47) with *ETR1* of *Arabidopsis thaliana* (f)
7. 306/434 positive identities (71%) (3e-46) with *ETR1* of *Prunus* persica (g)
8. 300/425 positive identities (71%) (1e-45) with *ETR1* of *Petunia* x hybrid (h)
9. 349/510 positive identities (68%) (2e-42) with *ETR1* of *Oncidium* Gower Ramsey (i)
10. 354/517 positive identities (68%) (6e-42) *ETR1* of *Dendrobium* Khao Sanan (j)

(a) Accession AF055894 Do *et al.,* 1999.

(b) Accession GU296679 Liu and Zhu, 2010.

(c) Accession AB042108 Itai *et al.,* 2004.

(d) Accession AJ563284 Chang, 2003.

(e) Accession AY821544 Asif and Solomos, 2004.

(f) Accession NM_105305 Swarbreck *et al.,* 2011.

(g) Accession AF124527 Basset *et al.,* 2002.

(h) Accession DQ154118 Wang and Kumar, 2007.

(i) Accession AF276237 Liu *et al.,* 2000.

(j) Accession FJ628419 Thongkum *et al.,* 2009.

7.3.3.4 Features of *ETR1* protein of *D.* Pompadour

Analysis with Prosite resulted in the detection of four different sites;N-glycosylation, protein kinase C Phosphorylation, Casein Kinase C Phosphorylation and N-Myristoylation. Table 7.5 summarizes the sites, position of residues corresponding to the sites, the consensus pattern and the number of sites detected in ETR1.

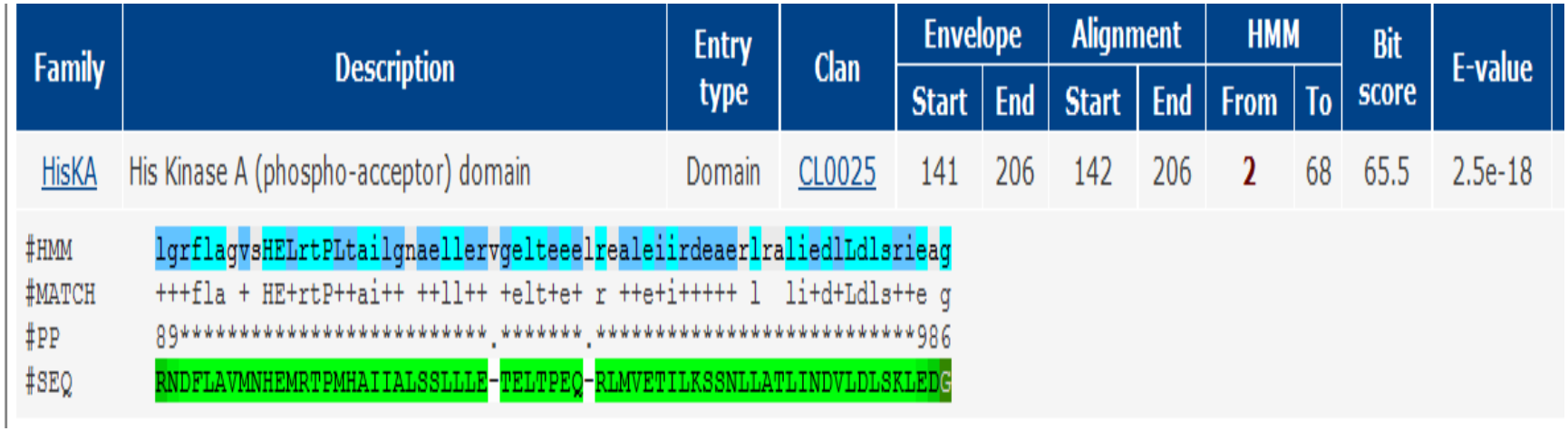
.

Table 7.5: Protein sites detected in *ETR1* using the Prosite tool accessed via www.expasy.org.

Site	Residue	Consensus Pattern	Number of sites
N-glycosylation	211 - 214 NSSR	N-{P}-[ST]-{P}	One
Protein kinase C phosphorylation	115 - 117 SvK 143 - 145 TgR 212 - 214 SsR	[ST]-x-[RK]	Three
Casein kinase II phosphorylation	78 - 81 TahE 139 - 142 TqeE	[ST]-x(2)-[DE]	Two
N-myristoylation	135 - 140 GLirTQ 196 - 201 GSvqSI	G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	Two

7.3.3.5 Pfam analysis of *ETR1* protein of *D.* Pompadour

Pfam analysis of *ETR1* protein showed the presence of histidine kinase domain at positions 148-208 amino acids. Histidine kinase domains are signature domains present in all ethylene receptor genes and have similar characteristics to a two-component bacteria system. Residues conserved in this domain were detected which are H, E, P and L. A conservative substitution of L/M was also occurred.



Family	Description	Entry type	Clan	Envelope		Alignment		HMM		Bit score	E-value
				Start	End	Start	End	From	To		
HisKA	His Kinase A (phospho-acceptor) domain	Domain	CL0025	141	206	142	206	2	68	65.5	2.5e-18

```
#HMM     lgrflagvsHELrtPLtailgnaellervgelteeelrealeiirdeaerlraliedlLdlsrieag
#MATCH   +++fla + HE+rtP++ai++ ++ll++ +elt+e+ r ++e+i+++++ l  li+d+Ldls++e g
#PP      89**************************.*******.*****************************986
#SEQ     RNDFLAVMNHEMRTPMHAIIALSSLLLE-TELTPEQ-RLMVETILKSSNLLATLINDVLDLSKLEDG
```

ILGSVQSINLPVVNRVFNSSRAVRITPHQLARSQPHTGPQDMPIIFRKFTQFNQLSNFQINNSNWPSAKNFAVMVLMLPSHSRDKWHVYELE

LVEVVADQVAVALSHAAILEESMRERDQLMEQNVALDLARREAEMAIRARNDFLAVMNHEMRTPMHAIIALSSLLLETELTPEQRLMVETIL

KSSNLLATLINDVLDLSKLEDGSFELEVTVFNLHTVFLMVVNLIKPIAAVDPLSLIVSLSPD

Fig 7.15 Results from Pfam analysis show the information on the histidine kinase domain. The #HMM line shows the consensus of the model, with capital letters representing the most conserved (high information content) positions, and dots (.) indicating insertions in the query sequence. Identical residues are coloured cyan, and similar residues are coloured dark blue. The #MATCH line indicates matches between the model and the query sequence, where a + indicates a "conservative substitution"; the #PP line represents the expected accuracy (posterior probability) of each aligned residue, where a 0 means 0–5%, 1 means 5–15%, and so on to 9 meaning 85–95% and a * meaning 95–100% posterior probability (pp); the #SEQ line is the query sequence, coloured according to the pp for each residue match on a scale from bright green for * through paler green and pale red down to bright red for 0. Histidine kinase domain in the ETR1 protein is highlighted in yellow.

7.3.3.6 ProtParam analysis of *ETR1* protein of *D.* Pompadour

The table below summarizes the physical and chemical properties of *ETR1*. The results show that *ETR1* protein of *D.* Pompadour share similar properties with other species. The size of *ETR1* fragment that is around 28 kDa is much smaller compare to *ETR1* protein from the other species which range from 180 kDa to 22 kDa. The theoretical PI of *D.* Pompadour *ETR1* that is acidic is similar to *ETR1* from *P.* hybrid, *D.* Sonia, *A. thaliana* and *O.* Gower Ramsey . The major amino acids in the *ETR1* protein of *D.* Pompadour, Thr is also the major amino acid in all the species analysed.

Table 7.6: Summary of protein parameters of *D.* Pompadour *ETR1* protein analysed using ProtParam

Species	Amino Acid	Molecular Weight	PI	Major amino acid composition			Stability
Dendrobium Pompadour	248	28374.9	6.44	Leu (10.5%)	Val (8.9%)	Thr (7.7%)	Unstable
Arabidopsis thaliana	2574	213796.8	4.86	Thr (30.3%)	Ala (28.3%)	Gly (22.6%)	Unstable
Dendrobium Sonia	2234	184879.2	4.88	Thr (28.8%)	Ala (27.8%)	Gly (23.6%)	Stable
Phalaenopsis hybrid	2564	212449.3	4.91	Thr (29.8%)	Ala (27.3%)	Gly (23.7%)	Stable
Oncidium Gower Ramsey	2409	199702.5	4.92	Thr (29.5%)	Ala (28.2%)	Gly (23.0%)	Stable

7.3.3.7 Multiple sequence analysis of *ETR1* protein of *D.* Pompadour

Multiple sequence alignment of *D.* Pompadour *ETR1* protein with *ETR1* from *P.* hybrid, *D.* Sonia, *O.* Gower Ramsey and *A. thaliana* using the ClustalW program (Fig 7.17 revealed the presence of a histidine kinase domain and three conserved regions. Residues conserved in a histidine kinase domain were also present.

A phylogenetic analysis of the relationship of *D.* Pompadour *ETR1* protein with *ETR1* proteins of *P.* hybrid, *D.* Sonia, *O.* Gower Ramsey and *A. thaliana* was also carried out using the ClustalW program and is presented in Fig 7.17. The phylogenetic tree derived showed that the proteins were divided into three clusters. *D.* Pompadour *ETR1* appeared to diverge in evolution since it shares the same cluster with and had the closest relationship to *A. thaliana.*

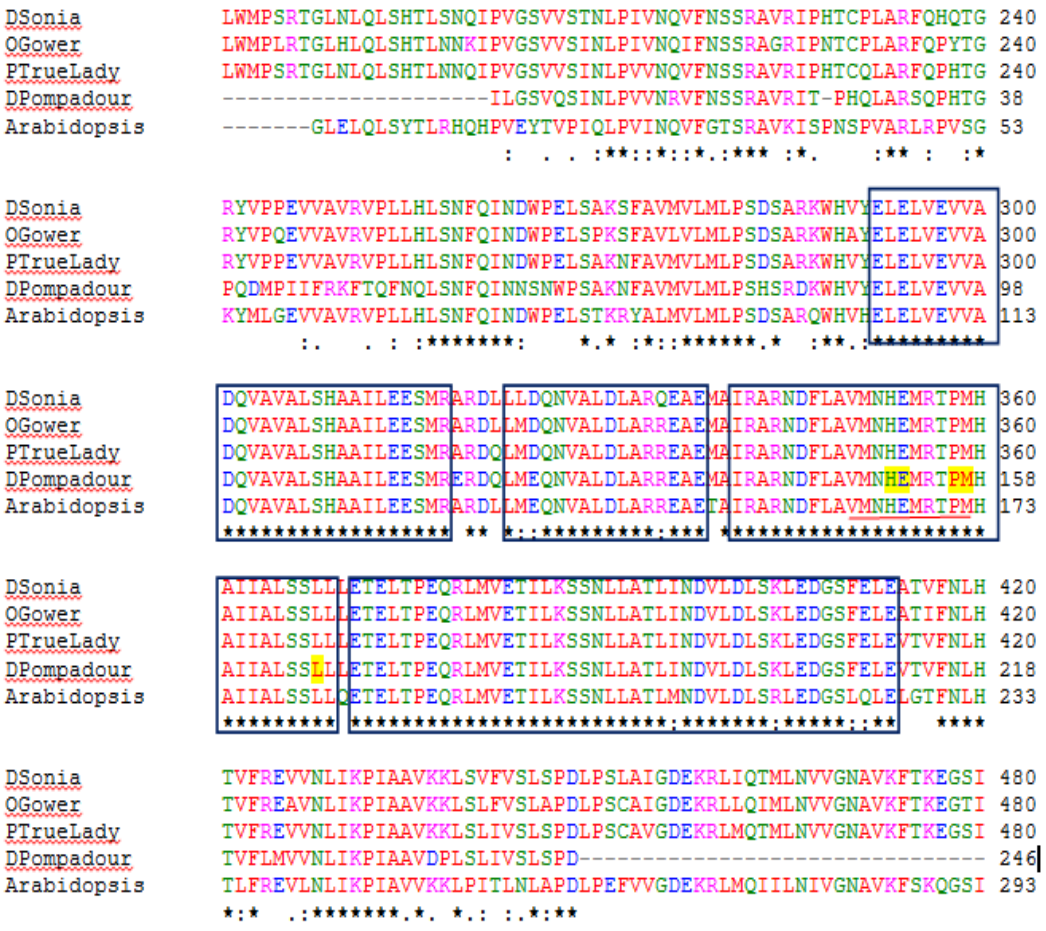


Fig 7.16 Comparison of the deduced amino acid sequence of *D.* Pompadour ETR1 and sequences encoding ETR1 from *O.* Gower Ramsey (AF276233), *D.* Sonia (ABJ91124.1), *P.* True Lady (AF055894) and *A. thaliana* (AY174554). Histidine kinase domain is underlined while conserved regions are boxed. Conserved residues of histidine kinase domain is highlighted in yellow.

Fig 7.17 Phylogenetic tree generated from the CLUSTALW multiple alignment programme. Figure shows the relationship between *D.* Pompadour ETR1 protein and ETR proteins from *P.* hybrid, *D.* Sonia, *O.* Gower Ramsey and *A. thaliana***.** The numbers indicate the distance matrix.

7.3.3.8 Three dimensional (3D) structure of *ETR1* protein of *D.* Pompadour

Analysis using Swiss Model and Rasmol indicated that the *ETR1* protein sequence had enough functional domains to generate a 3D protein structure. The structure consisted of 110 groups, 854 atoms and 864 bonds. Furthermore, the programme also identified 84 H bonds, 4 helices, and 4 turns.

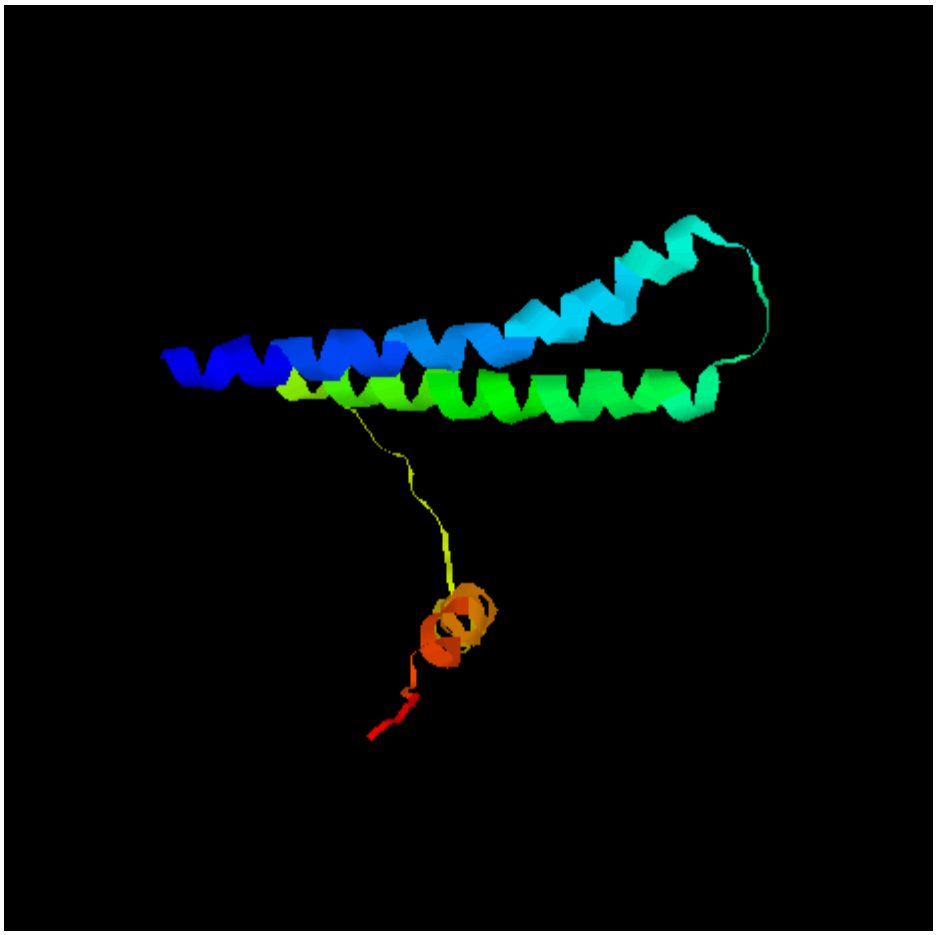


Fig 7.18 3D Structure of the *ETR1* protein as generated by Swiss Model and Rasmol programmes. The structure is viewed as ribbons where secondary structures are identified. The alpha helix is represented as a coiled structure and the beta strand is represented as a pleated structure. The colour is a smooth spectrum from blue through green, yellow and orange to red. The N termini of proteins are coloured blue and the C termini, red. This shows folding from one end of the ribbon to the other.

7.3.4 *ERS1* gene of *D.* Pompadour

7.3.4.1 RT-PCR results of *ERS1* gene of *D.* Pompadour

RT-PCR using specific primers for the *ERS1* gene resulted in a 955bp product. A clear band indicating the presence of the product was visible on a 1% agarose gel (Fig 7.19).

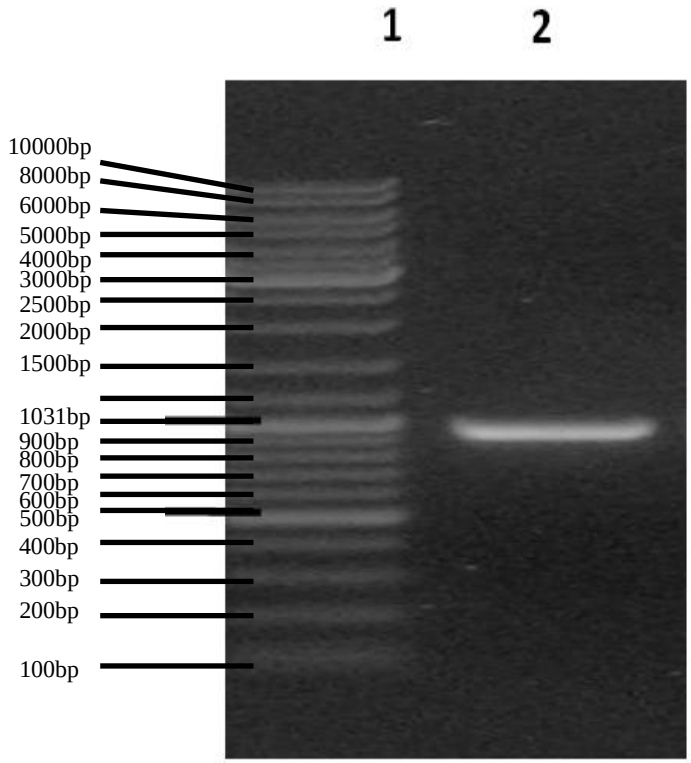


Fig 7.19 RT-PCR results with *ERS1* primers. In Lane 1 is the GeneRuler DNA ladder Mix (Fermentas). Lane 2 is the PCR product of 955bp gene analysed through 1% agarose gel.10µl of *ERS1* was loaded into the well while 5µl of ladder was used.

7.3.4.2 Nucleotide and amino acid sequence of *ERS1* gene of *D.* Pompadour

Fig 7.20 shows the nucleotide sequence of *ERS1* and the corresponding amino acid. Both the forward and reverse primers are italicized in red.

```
1     TTAAGGGGGGGGGGGCTTGTTTTTTCCATGCTGCCATCTTGGAGGAATCC  51
      ATGCGGGCACGAGATCTCCTTCTTGGATCAGAATGTTGCTTTAGATTTAG
101   CACGACAGGAGGCAGAGATGGCCATTCGTGCACGCAATGATTTTTTAGCT
151   GTCATGAACCATGAGATGCGGACTCCCATGCATGCAATCATTGCCCTTTC
201   CTCCCTGCTTCTTGAAACTGAACTGACTCCAGAGCAACGTTTGATGGTAG
251   AAACCATCTTAAAGAGTAGTAACTTGCTAGCAACCCTAATCAATGATGTT
301   TTAGACCTTTCTAAGCTTGAGGATGGCAGCTTCGAGTTAGAGGCCACAGT
351   TTTCAATCTTCATACTGTCTTCAGAGAGGTAGATGTGTTTATACTTTTTT
401   CATACTGTCTTCAGCCACAGTTATGTTTATACTTATATTTTTTTCCTTTG
451   GCAAATTTAATTTTCTACAAGTTTTGGTTAGCTGATTATTTTTTTATGTT
501   ATAATATAGGTCGTAAATTTGATAAAGCCAATAGCGGCTGTCAAAAAGTT
551   GTCAGTGTTCGTGTCTCTTTCTCCGGACTTGCCATCATTTGCCATTGGAG
601   ATGAGAAACGGCTTATACAAACTATGCTTAATGTTGTTGGCAATGCTGTT
651   AAGTTTACAAAGGAGGGTAGTATATCTATTACTGCGACTATTGCAAAATC
701   CGATTCCTTGAGAGATTCGCGAGACCCAGAGTTCCACCCTATCCCAAGCG
751   ATGGGCATTTCTATTTACGAGTACAGGTAACTTGATGACCCGGAACTGTA
801   TTTACAACTCTGATAGTTCTTCAATTATTCTTAGCAGATTTAAGCGAGGA
851   CTTAAAAATTGATTTGACATGTTTAGAGCATATGGAAATTTCCTGATTTG
901   ATGGTTCTTTTACAAAATTTGTGTAGAATAACCGAGACTACTCTTTCTTT
951   GTCCC
```

```
FLRGGGLVFSMLPSWRNPCGHEISFLDQNVALDLARQEAEMAIRARNDFLAVMNHEMR
TPMHAIIALSSLLLETELTPEQRLMVETILKSSNLLATLINDVLDLSKLEDGSFELEA
TVFNLHTVFREVDVFILFSYCLQPQLCLYLYFFPLANLIFYKFWLADYFFMLYRSISQ
RLSKSCQCSCLFLRTCHHLPLEMRNGLYKLCLMLLAMLLSLQRRVVYLLLRLLQNPIP
EIRETQSSTLSQAMGISIYEYRLDDPELYLQLFFNYSQIARTKLIHVSIWKFPDLMVL
LQNLCRITETTLSLSRFGAPDPDNPPGX
```

Fig 7.20 Nucleotide sequence of *ERS1* from *D.* Pompadour and the corresponding amino acid as translated using the Translate tool accessed via www.expasy.org. Forward and reverse primers are printed in red.

7.3.4.3 Homology of *ERS1* protein of *D.* Pompadour

BLAST homology search at www.ncbi.nlm.nih.gov/blast revealed the following match in the database:

1. 900/917 positive identities (98%) (0.0) with ERS1 of *Dendrobium* hybrid KhaoSanan (a)
2. 360/365 positive identities (99%) (2e-176) with ERS1 of*Dendrobium* cv. 'Sonia' (b)
3. 337/364 positive identities (93%) (5e-146) with ERS1 of *Oncidium* cv. 'Gower Ramsey' (c)
4. 331/359 positive identities (92%) (3e-142) with ERS1 of *Phalaenopsis* equestris (d)
5. 332/359 positive identities (92%) (3e-142) with ERS1 of *Phalaenopsis* sp. 'KCbutterfly' (e)
6. 299/355 positive identities (84%) (6e-101) with ERS1 of *Gladiolus* hybrid cultivar (f)
7. 290/359 positive identities (81%) (3e-86) with ERS1 of *Musa acuminata* AAA (g)
8. 282/354 positive identities (80%) (7e-81) with ERS1 of *Lilium* formosanum x *Lilium* longiflorum (h)
9. 282/354 positive identities (70%) (5e-76) with ETR1 of *Prunus persica* (i)
10. 274/348 positive identities (79%) (8e-74) with ERS1 of *Actinidia* deliciosa (j)

(a) Accession FJ644936 Thongkum *et al.*, 2009.

(b) Accession AY746972 Suwanagul *et al.*, 2004.

(c) Accession AF276234 Liu *et al.*, 2000.

(d) Accession AJ563284 Chang *et al.*, 2005.

(e) Accession , AF113541 Chai *et al.*, 1998.

(f) Accession , AB180247 Arora *et al.*, 2006.

(g) Accession AB266315 Yamane *et al.*, 2008.

(h) Accession DQ408428 Pan and Chen, 2006.

(i) Accession AF396830 Bonghi *et al.*, 2001.

(j) Accession EU170627 Yin *et al.*, 2008.

7.3.3.4 Features of *ERS1* protein of *D.* Pompadour

Analysis with Prosite resulted in the detection of four different sites; N-Myristolation, ,protein kinase C Phosphorylation, N-glycosylation and Casein Kinase C Phosphorylation. Table 7.7 summarizes the sites, position of residues corresponding to the sites, the consensus pattern and the number of sites detected in *ERS1*.

.

Table 7.7: Protein sites detected in *ERS1* using the Prosite tool accessed via www.expasy.org.

Site	Residue	Consensus Pattern	Number of sites
N-myristoylation	2-7: GIrmCI 69-74: GCspSA	G-{EDRKHPFYW}-x(2)-[STAGCN	Two
Protein kinase C phosphorylation	28 - 30:ScR 53 - 55:ScK 59 - 61:SdR 65 - 67:TtR 90 - 92:SaK 103 - 105:SdR 162 - 164:SdK	[ST]-x-[RK].	Seven
N-glycosylation	47 - 50:NSTY 160 - 163:NRSR	N-{P}-[ST]-{P}	Two
Casein kinase II phosphorylation	57 - 60:StsD 132 - 135:SleE 162 - 165:SrkD	[ST]-x(2)-[DE]	Three

7.3.4.5 Pfam analysis of *ERS1* protein of *D.* Pompadour

Pfam analysis of *ERS1* protein showed the presence of histidine kinase domain at positions 32-101amino acids. Residues conserved in this domain were also present which are H, P, E and L. A conservative substitution of L/M was detected.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		Bit score	E-value
				Start	End	Start	End	From	To		
HisKA	His Kinase A (phospho-acceptor) domain	Domain	CL0025	45	110	46	110	**2**	68	64.9	4e-18

```
#HMM      lgrflagvsHELrtPLtailgnaellervgelteeelrealeiirdeaerlraliedlLdlsrieag
#MATCH    +++fla + HE+rtP++ai++ ++ll++ +elt+e+ r ++e+i+++++ l  li+d+Ldls++e g
#PP       89************************.*******.*****************************986
#SEQ      RNDFLAVMNHEMRTPMHAIIALSSLLLE-TELTPEQ-RLMVETILKSSNLLATLINDVLDLSKLEDG
```

FLRGGGLVFSMLPSWRNPCGHEISFLDQNVALDLARQEAEMAIRARNDFLAVMNHEMRTPMHAIIALSSLLLETELTPEQRLMVETILKSSNL
LATLINDVLDLSKLEDGSFELEATVFNLHTVFREVDVFILFSYCLQPQLCLYLYFFPLANLIFYKFWLADYFFMLYRSISQRLSKSCQCSCLF
LRTCHHLPLEMRNGLYKLCLMLLAMLLSLQRRVVYLLLRLLQNPIPEIRETQSSTLSQAMGISIYEYRLDDPELYLQLFFNYSQIARTKLIHV
SIWKFPDLMVLLQNLCRITETTLSLSRFGAPDPDNPPGX

Fig 7.21: Results from Pfam analysis show the information on the histidine kinase domain. The #HMM line shows the consensus of the model, with capital letters representing the most conserved (high information content) positions, and dots (.) indicating insertions in the query sequence. Identical residues are coloured cyan, and similar residues are coloured dark blue. The #MATCH line indicates matches between the model and the query sequence, where a + indicates a "conservative substitution"; the #PP line represents the expected accuracy (posterior probability) of each aligned residue, where a 0 means 0–5%, 1 means 5–15%, and so on to 9 meaning 85–95% and a * meaning 95–100% posterior probability (pp); the #SEQ line is the query sequence, coloured according to the pp for each residue match on a scale from bright green for * through paler green and pale red down to bright red for 0. The histidine kinase domain in the deduced ERS1 protein is highlighted in yellow.

7.3.4.6 ProtParam analysis of *ERS1* protein of *D.* Pompadour

The table below summarizes the physical and chemical properties of *ERS1*. The results show that *ERS1* protein of *D. Pompadour* share similar properties with other species. The size of *ERS1* fragment that is around 22 kDa is much smaller compared to *ERS1* protein from the other species which range from 69 kDa to 71 kDa. The theoretical PI of *D.* Pompadour *ERS1* that is basic is similar to *ERS1* from *D.* Khao Sanan and Sonia. The major amino acid in the *ERS1* protein of *D.* Pompadour, Leu is also the major amino acid in *D.* Khao Sanan, Sonia, *P.* Equestris and KcButterfly.

Table 7.8: Summary of protein parameters of *D.* Pompadour *ERS1* protein analysed using ProtParam

Species	Amino Acid	Molecular Weight	PI	Major amino acid composition			Stability
Dendrobium Pompadour	199	22389.8	7.06	Ser (12.6%)	Leu (9.0%)	Ala (6.5%)	Unstable
Dendrobium Khao Sanan	622	69876.5	7.29	Leu (13.5%)	Ser (7.6%)	Ala (6.9%)	Stable
Phalaenopsis Equestris	633	70997.7	6.61	Leu (12.6%0	Val (8.2%)	Ala (7.3%)	Stable
Dendrobium Sonia	621	69763.4	7.29	Leu (13.5%)	Val (8.1%)	Ser (7.6%)	Stable
Phalaenopsis KcButterfly	633	71086.8	6.88	Leu (13.0%)	Val (8.1%)	Ala (7.4%)	Stable

7.3.4.7 Multiple sequence analysis of *ERS1* protein of *D.* Pompadour

Multiple sequence alignment of *D.* Pompadour *ERS1* protein with *ERS1* protein from *D.* Khao Sanan, *P.* Equestris, *D.* Sonia and *P.* KCButterfly using the ClustalW program (Fig 7.22) revealed the presence of a histidine kinase domain and three highly conserved regions where each region has ten or more consecutive conserved residue. Residues conserved in a histidine kinase domain were also present.

A phylogenetic analysis of the relationship of *D.* Pompadour *ERS1* protein with *ERS1* proteins *D.* Khao Sanan , *P.* Equestris and *P.* KC butterfly and *D.* Sonia was also carried out using the ClustalW program and is presented in Fig 7.23. The phylogenetic tree derived showed that these proteins were divided into three clusters. *D.* Pompadour *ERS1* did not share the same cluster with any of the other orchids.

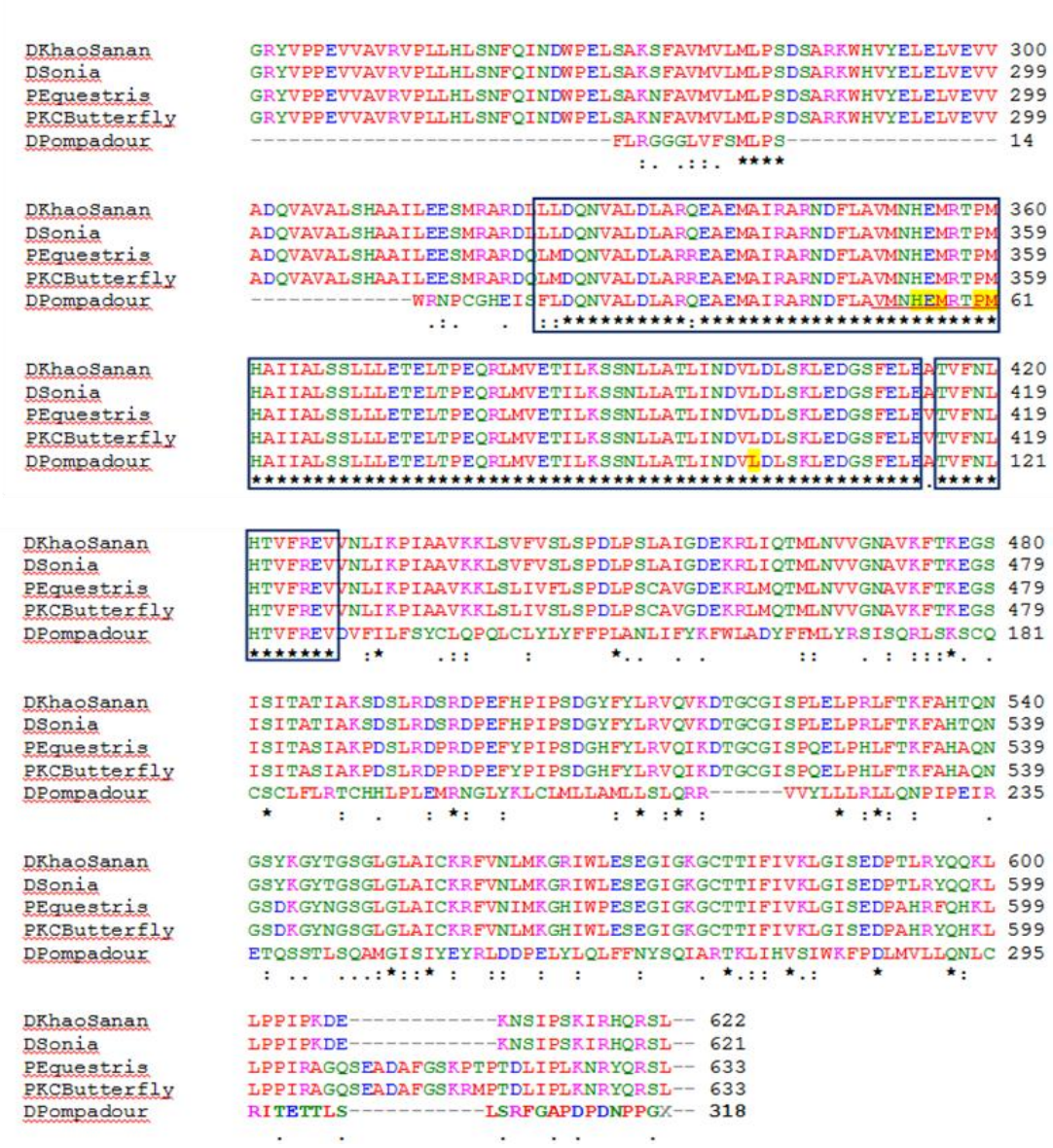


Fig 7.22 Comparison of the deduced amino acid sequence of *D.* Pompadour ERS1 and sequences encoding *ERS1* from *P.* KCButterfly (AF113541.1),*P.* Equestris (AJ563284.1), *D.* Sonia (AY746972.2) and *D.* Khao Sanan (FJ628419.1). The histidine kinase domain is underlined while conserved regions are boxed. Residues conserved in the histidine kinase domain are highlighted in yellow.

Fig 7.23 Phylogenetic tree generated from the CLUSTALW multiple alignment programme. Figure shows the relationship between *D.* Pompadour *ERS1* protein and proteins from *P.* Equestris and KCbutterfly , *D.*Sonia and *D.* Khao Sanan. The numbers indicate the distance matrix.

7.3.4.8 Three dimensional (3D) structure of *ERS1* protein of *D.* Pompadour

Analysis using Swiss Model and Rasmol indicated that the *ERS1* protein had enoufg functiona domains to generate a 3D protein structure. The structure consisted of 109 groups, 1131 atoms and 1157 bonds. Furthermore, the programme also identified 109 H bonds, 6 helices, 2 strands and 4 turns.

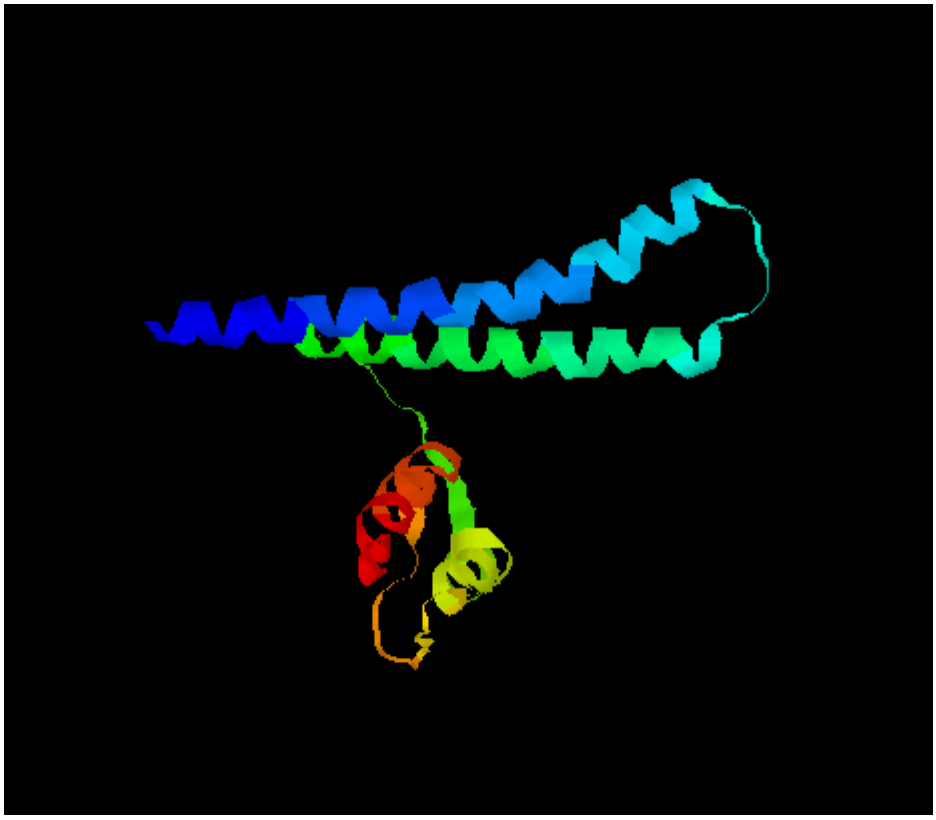


Fig 7.24 3D Structure of the *ERS1* genes as generated by Swiss Model and Rasmol programmes. The structure is viewed as ribbons where secondary structures are identified. The alpha helix is represented as acoiled structure and the beta strand is representedas a pleated structure. The colour is a smooth spectrum from blue through green, yellow and orange to red. The N termini of proteins are coloured blue and the C termini, red. This shows folding from one end of the ribbon to the other.

7.3.5 *ERS2* gene of *D.* Pompadour

7.3.5.1 RT-PCR results of *ERS2* gene of *D.* Pompadour

RT-PCR using specific primers for the *ERS2* gene resulted in a 756bp product. A clear band indicating the presence of the product was visible on a 1% agarose gel (Fig 7.25).

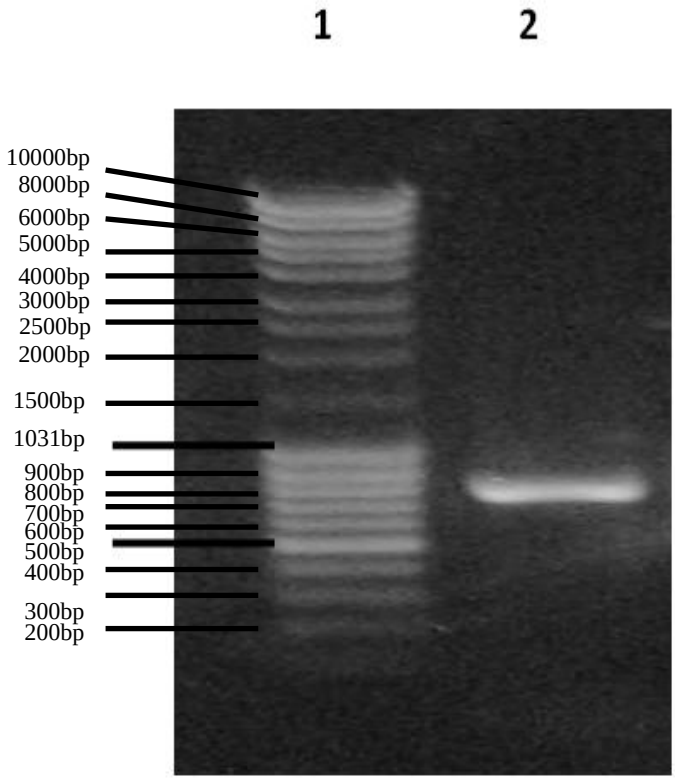


Fig 7.25 RT-PCR results with *ERS2* primers. In Lane 1 is the MassRuler DNA ladder Mix (Fermentas). Lane 2 is the PCR product of 756 bp gene analysed through 1% agarose gel. 10µl of *ERS2* was loaded into the well while 5µl of ladder was used.

7.3.5.2 Nucleotide and amino acid sequence of *ERS2* gene of *D.* Pompadour

Fig 7.26 shows the nucleotide sequence for *ERS2* and the corresponding amino acids. The forward and reverse primers are italicized in red.

```
1      ATTTTAGCATTCCGCTGGAACTGATTTATTTTGTGAAAAAAAGCAGCTTT
51     TTTCCGTATCGCTGGGTGCTGATTCAGTTTGGCGCGTTTATTGTGCTGTG
101    CGGCGCGACCCATCTGATTAACCTGTGGACCTTTACCATGCATAGCCGCA
151    CCCTGGCGATTGTGATGACCGTGGCGAAAGTGAGCACCGCGGTGGTGAGC
201    TGCGCGACCGCGCTGATGCTGGTGCATATTATTCCGGATCTGCTGAGCGT
251    GAAAACCCGCGAACTGTTTCTGCGCAACAAAGCGGAAGAACTGGATCGCG
301    AAATGGGCCTGATTCGCACCCAGGAAGAAACCGGCCGCCATGTGCGCATG
351    CTGACCCATGAAATTCGCAGCACCCTGGATCGCCATACCATTCTGAAAAC
401    CACCCTGGTGGAACTGGGCCGCACCCTGGATCTGGCGGAATGCGCGCTGT
451    GGATGCCGAGCCGCACCGGCCTGAACCTGCAGCTGAGCCATACCCTGAGC
501    AACCAGATTCCGGTGGGCAGCGTGGTGAGCACCAACCTGCCGATTGTGAA
551    CCAGGTGTTTAACAGCAGCCGCGCGGTGCGCATTCCGCATACCTGCCCGC
601    TGGCGCGCTTTCAGCATCAGACCGGCCGCTATGTGCCGCCGGAAGTGGTG
651    GCGGTGCGCGTGCCGCTGCTGCATCTGAGCAACTTTCAGATTAACGATTG
701    GCCGGAACTGAGCGCGAAAAGCTTTGCGGTGATGGTGCTGATGTTTCCGA
751    GCGATA
```

```
FIALAYFSIPLELIYFVKKSSFFPYRWVLIQFGAFIVLCGATHLINLWTFTM
HSRTLAIVMTVAKVSTAVVSCATALMLVHIIPDLLSVKTRELFLRNKAEELD
REMGLIRTQEETGRHVRMLTHEIRSTLDRHTILKTTLVELGRTLDLAECALW
MPSRTGLNLQLSHTLSNQIPVGSVVSTNLPIVNQVFNSSRAVRIPHTCPLAR
FQHQTGRYVPPEVVAVRVPLLHLSNFQINDWPELSAKSFAVMVLMFPSDTAS
SPHEH
```

Fig 7.26 Nucleotide sequence of *ERS2* from *D.* Pompadour and the corresponding amino acid as translated using the Translate tool accessed via www.expasy.org. Forward and reverse primers are printed in red.

7.3.4.3 Homology of *ERS2* protein of *D.* Pompadour

BLAST homology search at www.ncbi.nlm.nih.gov/blast revealed the following match in the database:

1. 579/781 positive identities (74%) (6e-137) with *ERS* of *Lilium* formosanum x longiflorum (a)
2. 553/760 positive identities (73%) (4e-114) with *ERS* of *Dendrobium* cv Sonia (b)
3. 552/763 positive identities (72%) (9e-110) with *ERS* of *Dendrobium* Khao Sanan (c)
4. 535/762 positive identities (70%) (8e-85) with *ERS* of *Delphinium* Magic fountain (d)
5. 530/760 positive identities (70%) (8e-85) with *ERS* of *Phalaenopsis* sp. 'KCbutterfly' (e)
6. 533/766 positive identities (70%) (4e-83) with *ER2* of *Oncidium* Gower Ramsey (f)
7. 529/760 positive identities (70%) (1e-83) with *ERS* of *Phalaenopsis equestris* (g)
8. 529/760 positive identities (70%) (1e-83) with *ERS* of *Phalaenopsis*'True lady' (h)
9. 530/762 positive identities (70%) with (1e-76) *with* ERS of *Delphinium* Belladonna (i)
10. 525/760 positive identities (69%) (2e-79) with *ERS* of *Gladiolus* hybrid cultivar (j)

11.

(a) Accession DQ408428 Pan and Chen, 2006.

(b) Accession AY746972 Suwanagul *et al.*, 2004.

(c) Accession FJ628419 Thongkum *et al.*, 2009.

(d) Accession AB055430 Kuroda *et al.*, 2003.

(e) Accession AF113541 Chai *et al.*, 1998.

(f) AccessionAF276234 Liu *et al.*, 2000.

(g) AccessionAJ563284 Chang, 2005.

(h) Accession AF055894 Do *et al.*, 1999.

(i) Accession AB201245 Tanase and Ichimura 2006.

(j) Accession AB180248 Arora *et al.*, 2006.

7.3.5.4 Features of *ERS2* protein of *D.* Pompadour

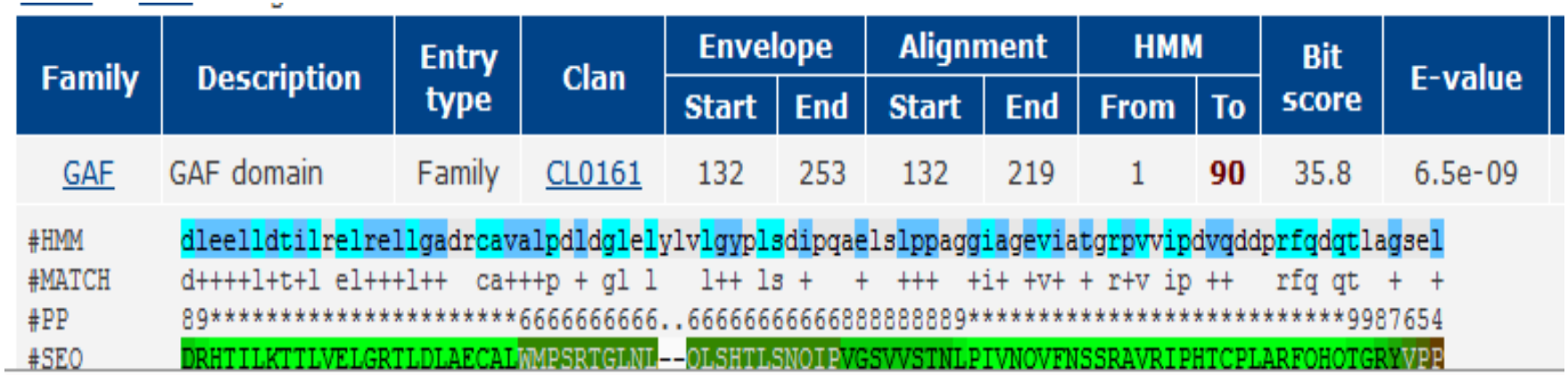
Analysis with Prosite showed that *ERS2* isolated from *D.* Pompadour contained five different protein site, cAMP- and cGMP-dependent protein kinase phosphorylation, Protein kinase C phosphorylation, *N*-myristoylation and Casein kinase II phosphorylation N-glycosylation. These sites are all common sites found in ethylene receptors.

Table 7.9: Protein sites detected in *ERS2* using the Prosite tool accessed via www.expasy.org.

Site	Residue	Consensus Pattern	Number of sites
cAMP- and cGMP-dependent protein kinase phosphorylation	18-21 KKsS	[RK](2)-x-[ST]	One
Protein kinase C phosphorylation	88-90 SvK 116 – 118 TgR 194 – 196 SsR 213 – 215 TgR 243 – 245 SaK	[ST]-x-[RK]	Five
N-myristoylation	108 - 113:GlirTQ 178 - 183:GSvvST	G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}	Two
Casein kinase II phosphorylation	112 – 115 TqEe 129 – 132 StlD 140 – 143 TlvE 261 – 264 SphE	[ST]-x(2)-[DE]	Four
N-glycosylation	193 – 196 NSSR	N-{P}-[ST]-{P}	One

7.3.5.5 Pfam analysis of *ERS2* protein of *D.* Pompadour

Pfam analysis of *ERS2* protein showed the presence of GAF domain at positions 132 - 220 amino acids. The GAF domain has been reported to be present in ethylene receptors, though its function in ethylene receptors is unknown.



Family	Description	Entry type	Clan	Envelope		Alignment		HMM		Bit score	E-value
				Start	End	Start	End	From	To		
GAF	GAF domain	Family	CL0161	132	253	132	219	1	**90**	35.8	6.5e-09

```
#HMM    dleelldtilrelrellgadrcavalpdldglelylvlgyplsdipqaelslppaggiageviatgrpvvipdvqddprfqdqtlagsel
#MATCH  d++++l+t+l el+++l++  ca+++p + gl l   l++ ls +   +  +++  +i+ +v+ + r+v ip ++   rfq qt  +  +
#PP     89**********************6666666666..6666666666888888889**************************9987654
#SEQ    DRHTILKTTLVELGRTLDLAECALWMPSRTGLNL--QLSHTLSNQIPVGSVVSTNLPIVNQVFNSSRAVRIPHTCPLARFQHQTGRYVPP
```

FIALAYFSIPLELIYFVKKSSFFPYRWVLIQFGAFIVLCGATHLINLWTFTMHSRTLAIVMTVAKVSTAVVSCATALMLVHI
IPDLLSVKTRELFLRNKAEELDREMGLIRTQEETGRHVRMLTHEIRSTLDRHTILKTTLVELGRTLDLAECALWMPSRTGLN
LQLSHTLSNQIPVGSVVSTNLPIVNQVFNSSRAVRIPHTCPLARFQHQTGRYVPPEVVAVRVPLLHLSNFQINDWPELSAKS
FAVMVLMFPSDTASSPHEH

Fig 7.27 Results from Pfam analysis show the information on GAF domain. The #HMM line shows the consensus of the model, with capital letters representing the most conserved (high information content) positions, and dots (.) indicating insertions in the query sequence. Identical residues are coloured cyan, and similar residues are coloured dark blue. The #MATCH line indicates matches between the model and the query sequence, where a + indicates a "conservative substitution"; the #PP line represents the expected accuracy (posterior probability) of each aligned residue, where a 0 means 0–5%, 1 means 5–15%, and so on to 9 meaning 85–95% and a * meaning 95–100% posterior probability (pp); the #SEQ line is the query sequence, coloured according to the pp for each residue match on a scale from bright green for * through paler green and pale red down to bright red for 0. The GAF domain in the deduced ERS2 protein is highlighted in yellow.

7.3.5.6 ProtParam analysis of *ERS2* protein of *D.* Pompadour

The table below summarizes the physical and chemical properties of *ERS2*. The results show that *ERS2* protein of *D. Pompadour* share similar properties with other species. The size of *ERS2* fragment that is around 29 kDa is much smaller compared to *ERS2* protein from the other species which range from 69 kDa to 71 kDa. The theoretical PI of *D.* Pompadour *ERS2* that is basic is similar to *ERS2* from *D.* Khao Sanan, *D.* Sonia, *P.* True Lady, *P.* KcButterfly and *O.* Gower Ramsey. The major amino acids in the *ERS2* protein of *D.* Pompadour, Leu and Val are also the major amino acids in the other species.

Table 7.10 Summary of protein parameters of *D.* Pompadour ERS2 protein analysed using ProtParam

Species	Amino Acid	Molecular Weight	PI	Major amino acid composition			Stability
Dendrobium Pompadour	265	29955.1	9.21	Leu (13.1%)	Val (9.4%)	Thr (8.3%)	Stable
Dendrobium Khao Sanan	622	69876.5	7.29	Leu (13.7%)	Val (8.9%)	Ser (7.6%)	Stable
Dendrobium Sonia	621	69763.4	7.29	Leu (13.4%)	Val (8.1%)	Ser (7.6%)	Stable
Phalaenopsis True Lady	633	71084.7	6.88	Leu (13.0%)	Val (8.4%)	Ala (7.4%)	Stable
Phalaenopsis KcButterfly	633	71086.8	6.88	Leu (13.0%)	Val (8.1%)	Ala (7.4%)	Stable
Oncidium Gower Ramsey	631	70721.6	6.82	Leu (13.6%)	Ala (7.9%)	Val (7.1%)	Stable

7.3.5.7 Multiple sequence alignment of *ERS2* protein of *D.* Pompadour

Multiple sequence alignment of *D.* Pompadour *ERS2* protein with *ERS* protein from *D.* Khao Sanan, *P.*True Lady, *D.* Sonia and *P.* KCButterfly using the ClustalW program (Fig 7.28) revealed the presence of a GAF domain and eight highly conserved regions where each region has ten or more consecutive conserved residue.

A phylogenetic analysis of the relationship of *D.* Pompadour *ERS2* protein with *ERS* proteins from from *D.* Khao Sanan, , *D.* Sonia, *P.* True Lady and *P.* KCButterfly was also carried out using the ClustalW program and is presented in Fig 7.29. The phylogenetic tree derived showed *D.* Pompadour *ERS2* is more closely related to *D.* Khao Sanan and Sonia compared to the other orchid hybrids.

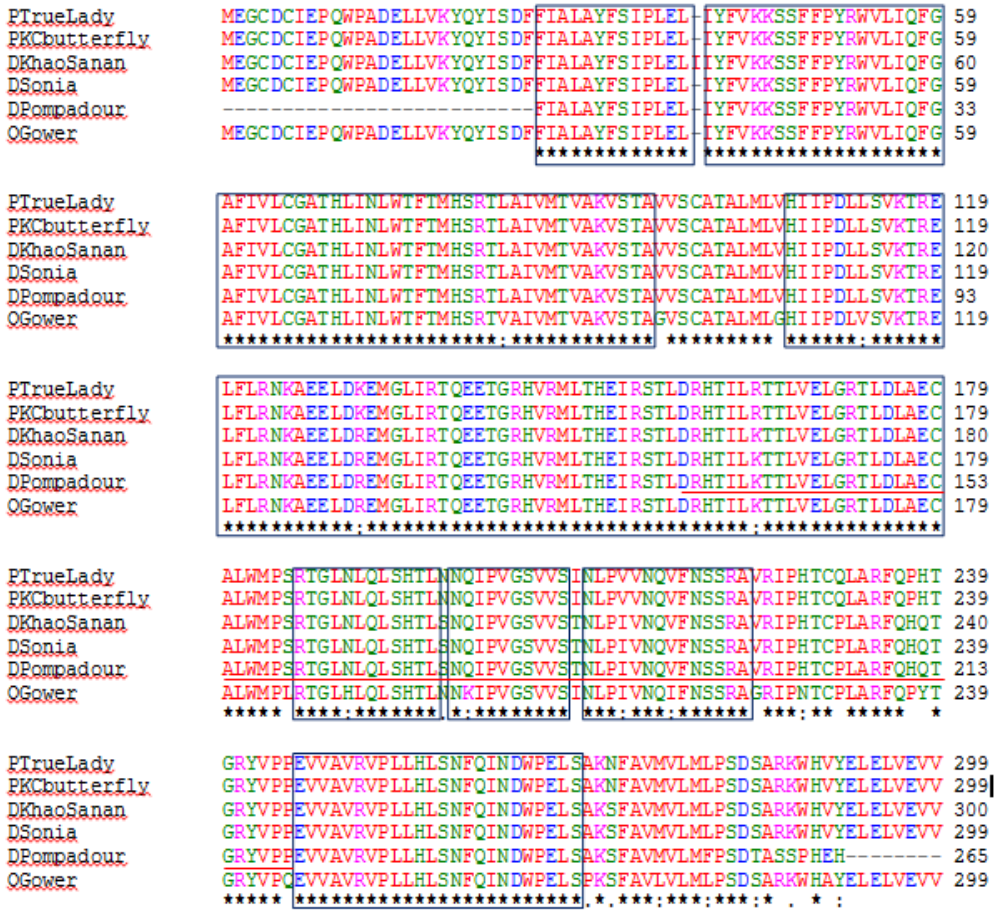


Fig 7.28 Comparison of the deduced amino acid sequence of *D.* Pompadour ERS2 and sequences encoding *ERS2* from *D.* Khao Sanan ethylene (FJ628419), *D.* Sonia (AY746972), *P.* True Lady (AF055894), *P.* KCbutterfly (AF113541) and *O.* Gower Ramsey (AF276233). The GAF domain is underlined while conserved regions are boxed.

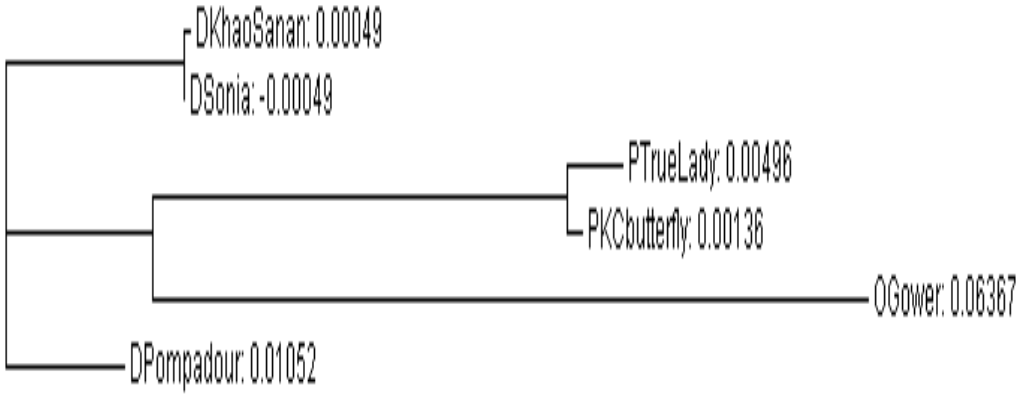


Fig 7.29 Phylogenetic tree generated from the CLUSTALW multiple alignment programme.Figure shows the relationship between *D.* Pompadour *ERS2* protein and *ERS* proteins from *D.* Khao Sanan, *P.* True Lady, *D.* Sonia, *P.* KCButterfly and *O.* Gower Ramsey. The numbers indicate the distance matrix.

7.3.5.8 Three dimensional (3D) structure of *ERS2* protein of *D.* Pompadour

Analysis using Swiss Model and Rasmol indicated that the *ERS2* protein had enough functional domains to generate a 3D protein structure. The protein structure consisted of 117 groups, 1219 atoms and 1107 bonds. Furthermore, the programme also identified 86 H bonds, 5 helices, 6 strands and 11 turns.

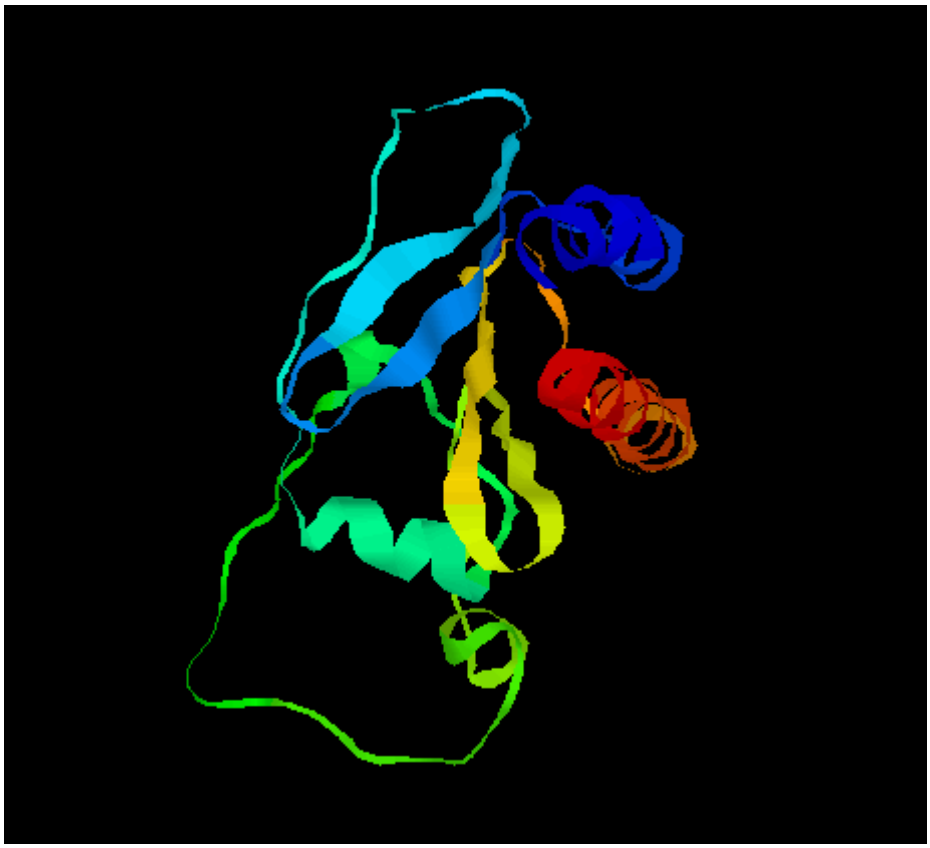


Fig 7.30 3D Structure of the ERS2 genes as generated by Swiss Model and Rasmol programmes. The structure is viewed as ribbons where secondary structures are identified. The alpha helix is represented as acoiled structure and the beta strand is representedas a pleated structure. The colour is a smooth spectrum from blue through green, yellow and orange to red. The N termini of proteins are coloured blue and the C termini, red. This shows folding from one end of the ribbon to the other.

7.4 DISCUSSION

The regulation of pollination-induced senescence in flowers by ethylene has long been an accepted dogma (Arditti, 1979). In the earlier section (3.2), results confirm the climacteric nature of pollinated *D.* Pompadour flowers, producing a burst of ethylene and eventually senescing within 48 after pollination. In Arditti's review on orchids, this climacteric nature was found to be present in all the orchids studied in literature. However, this is not unique to orchids, as other flowers such carnations, rose, petunia, hibiscus and *Arabidopsis* have also been established as climacteric flowers.

ACCO requires Fe(II) and ascorbate as cofactors for enzymatic activity (McGarvey and Christoffersen 1992), and 12 amino acid residues of ACCO participate in the interaction with these cofactors (Zarembinski and Theologis 1994). The 321 amino acid protein encoded by the gene isolated from *D.* Pompadour in this study contains a highly conserved Fe(2+) 2-oxoglutarate dioxygenase domain. This domain is also present in *ACCO* isolated in all the other orchids reported in the NCBI GeneBank which are *C.* hybrid, *,Dtps. sp*, *P.* True Lady, *O.* Gower Ramsey, *D. Crumenatum,D.* Sonia, *D. Anna* and *D.* Khao Sanan as well as other flowers such as carnations, *Arabidopsis* , *Pelargonium* and *Rosa* hybrid. This further confirms the Fe(2+) 2-oxoglutarate dioxygenase domain as the signature domain which is not only unique to orchid *ACCO* but also other climacteric flowers as well. In plants, Fe(II) 2OG dioxygenase domain enzymes catalyse the formation of plant hormones, such as ethylene, gibberellins, anthocyanidins and pigments such as flavones. The protein encoded also contained five conserved residues with seven conservative substitutions. These substitutions may not affect the function of the *ACCO* protein as has been demonstrated in petunia where

substitution of alanine with glycine still resulted in a functional *ACCO*, suggesting that amino acids such as alanine may play the same role (Do *et al.,* 2005).

Analysis with ScanProsite showed the presence of two histidine and an aspartate residues which are conserved in all the Fe(II) requiring families of enzymes (Tang *et al.,* 1993). It also contains the sequence motifs of 2OG-dependent oxygenases, [RK], which was found to bind to the terminal carboxylate of 2OG. The histidine and aspartate residues catalytically bind a metal ion, generally iron, and are directly involved in catalysis. Mutagenesis and spectroscopic investigations found that [HX--H] motif of the conserved domain is employed in the binding of iron in *ACCO* (Schofield and Zhang, 1999). Moreover, the substitution of histidine or aspartate in tomato *ACCO* with glutamate resulted in a lower rate of ACC conversion to ethylene (Nakatsuka *et al.,* 1998).

A conserved glycosylation site [NISK] was also detected in the deduced *ACCO* protein. This site allows for oligosaccharide attachment to allow for protein folding. Correctly folded proteins will eventually be exported from the endoplasmic reticulum. A study done by Meli *et al., (*2010) showed that when N-glycoproteins (precursors for glycosylation) were suppressed in transgenic Tomato (cv. Pusa ruby), expression of ACCO was down regulated and vase life of the fruits was extended. The results from that study further validated the importance of glycosylation sites in ensuring that *ACCO* is able to function. Analysis using ProtParam showed that the *ACCO* of *D.* Pompadour is soluble. This is similar to *ACCO* of *Phalaenopsis* that was readily extracted with water (Chung *et al.,* 2002). The author suggested that the association of *ACCO* with the pellet fraction as was reported in apple and pear fruit might be an artifact.

Homology analysis showed that although *D.* Pompadour belonged to the same genus and family of *D. Crumenatum* and *Doritaenopsis sp.* respectively, the degree of homology was closer to that of the *ACCO* protein isolated from *Dianthus caryophyllus* and *Rosa* hybrid, both with more than 70% homology compared to 65-67% homology with the other orchid species. The close relationship with *Dianthus caryopyllus*is further confirmed by the phylogram where the *ACCO* from *D.* Pompadour and *Dianthus* caryophyllus share the same branch. Since *D.* Pompadour is a hybrid line, the different genotype may contribute to its variation from the other orchid species. Similar findings have been observed in *D.* Anna. The size of *D.* Pompadour *ACCO* was similar to *ACCO* proteins found in the GenBank.

ACCS belongs to the family of pyridoxal 5'-phosphate- dependant enzyme. This enzyme is part of a multigene family and transcription of different forms is induced by different physiological and environmental conditions (Stearns and Glick, 2003). RT-PCR amplification with specific primers resulted in a 1.32Kb nucleotide that translates into 440 amino acids. The protein contains 10 invariant residues in this subgroup and a conserved lysine residue which has been postulated as the essential residue for ensuring the function of *ACCS* protein. It was reported that this lysine residue is responsible for the binding of PLP and covalent linking of 2-aminobutyrate portion of AdoMet during the inactivation (Yip *et al.,* 1992). The deduced amino acid contained one Aminotransferases class-I pyridoxal-phosphate attachment site, the signature domain for this enzyme within which lies the lysine residue. Alignment of *D.* Pompadour *ACCS* with other species within this site detected one amino acid substitution (L/I). This substitution is unique to *D.* Pompadour as the *ACCS* from *Cymbidium* (BAF 36561),*D. crumenatum*(AAB67882), *D.* Sonia (ACY 72181), *Doritaenopsis* sp (AAB05849) and *Oncidium* Gower Ramsey (AAQ07441)maintain the L residue within this site. Amino

acid substitution however is not unique to *D.* Pompadour as in other flowers such as *Arabidopsis* and *Rosa* hybrid, (G/S) and (A/V) substitutions were found.

A study on mutant *ACCS* in tomatoes found that mutations which involved variant residues between *ACCS* and subgroup 1 aminotransferase did not result in a loss of function, concluding that some of the residues may be tolerant to substitution. These residues include Y92 and Y240. This, according to the author may mean that these residues are not essential for the functioning of *ACCS* (Tarun *et al.*, 1998). Studies on the effects of residue substitution on the function of *ACCS* are crucial as they could give insight into the accurate role of these residues.

The alignment of *D.* Pompadour *ACCS* with other species also showed a conserved Thr-Asn-Pro (TNP), a tripeptide implicated to be a stability determinant in *ACCS*. Poor expression of *ACCS* was observed when no TNP was present in an *Arabidopsis ACCS1* isoform (Liang *et al.*, 1995) while mutation of the tripeptide in tomato ACCS resulted in inactivity in mutants (Tarun *et al.*,1998). *ACCS* isolated from *D.* Pompadour shared a high degree of homology with *ACCS* from other plants species, the highest being 81.9% with *ACCS* isolated from pelargonium. More than 70% homology was shared with *Petunia*, *Nicotiana, Dianthus caryophyllus* and *Arabidopsis thaliana.* Analysis with ProtParam further highlights the similarity between *ACCS* protein of the different species with Leu and Ser as the major amino acids. The *ACCS* protein of *Dendrobium* Pompadour was found to be unstable; a similar characteristic shared by *ACCS* isolated from *Petunia*, *Nicotiana* and *Dianthus caryophyllus*. Instability of *ACCS* coupled with its low concentrations in ethylene producing tissues has made purification of this enzyme a difficult task. When comparing *D.* Pompadour *ACCS* with proteins available in the GenBank, it was found that the size of the isolated *D.* Pompadour *ACCS* was

comparable as the average size of proteins submitted was approximately 450 aa. Multigene family for *ACCS* sequences has been reported for zucchini (Huang *et al.*, 1991), tomato (Sato and Theologis 1989; van Der Straeten *et al.*, 1990; Rottmann *et al.*, 1991), *Arabidopsi*s (Liang *et al.*, 1992; van Der Straeten *et al.*, 1992), mung beans (Botella *et al*., 1992), rice (Zarembinski and Theologis 1993), potato (Beltran Destefano *et al*., 1995) and carnation (ten Have and Woltering 1997).

Ethylene receptors have been widely studied at both biochemical and molecular level in many flowers including *Arabidopsis thaliana, Dianthus caryophyllus*a nd *Petunia* hybrida. In orchids, a number of ethylene receptors have been sequenced, mostly in *Phalaenopsis sp.* In this chapter, three partial cds of *ETR1, ERS1* and *ERS2* were isolated and further characterised using the bioinformatics tool. Using degenerate primers designed via alignment of ethylene receptors retrieved from the GenBank, four partial cds were isolated sized 692bp, 955bp and 756bp for *ETR1, ERS1* and *ERS2* respectively. Analysis of the deduced amino acids detected a histidine kinase domain in *ETR1* and *ERS1* while a GAF domain was detected in *ERS2*. A typical His protein kinase has five conserved signature motifs, H, N, G1, F, and G2, with the conserved His as the central feature in the H motif. The other four motifs define the nucleotide-binding cleft (Grebe and Stock, 1999).The GAF domain binds cyclic nucleotides in a number of bacterial proteins, and the chromophore in the plant photoreceptor phytochrome (Aravind and Ponting, 1997). However, the function of this domain in the ethylene receptors is unknown.

These sequence data suggests that amino acids sequenced could be putative ethylene receptors with the ability to bind ethylene.The histidine kinase domain detected in *ETR1* and *ERS1* was highly conserved at the H residue, presumed as autophosphorylation site

for the domain. The mechanism of histidine phosphorylation is iniated as a respond to a stimulus, received by the response regulator and followed through by the effector domain. Following autophosphorylation, a phosphoryl group is transferred by the response regulator which in turn catalyses the transfer to an aspartate residue in the receiver domain (Stock *et al,* 2000).

Early investigations on ethylene receptors assumed that the presence of a histidine kinase is compulsory in order for signal transduction to occur. This hypothesis was established as *ETR1* in both *Arabidopsis* and tomatoeswere found to have conserved histidine kinase domains (Chang *et al.,* 1993; Tieman and Klee, 1999). More recent published data however have come to a contradictory hypothesis where histidine kinase may not be the primary means for signal transduction (Wang *et al.*, 2003). The isolation of ethylene receptors from subfamily 2 that contain a degenerate histidine kinase domain but still allow for signal transduction suggest the less significant role of histidine kinase. Furthermore, serine substitutions in the ethylene receptors, except for *ETR* still allowed for autophosphorylation to occur. In fact, it was postulated that the receptors may have evolved into serine-phosphorylation proteins (Moussatche and Klee 2004).

Phylogenetic analysis of *ETR1* isolated from *D.* Pompadour resulted in a closer relationship with *Arabidopsis thaliana,* compared to three other orchid flowers, *P.* hybrid, *D.* Sonia and *O.* Gower Ramsey. *ERS1* on the other hand showed high degree of similarity with four other orchid species and hybrids, *D.* Khao Sanan, Sonia, *P. Equestris* and KCbutterfly. *ERS2* shared high degree of similarity with other orchids. The diversion from orchids, in the case of *ETR1* may be contributed by changes in genotype caused by hybridization of the flower. Amongst the three receptors isolated,

ETR1 showed different major amino acid composition. Where *ETR1* is concerned (in *Dendrobium* Pompadour and the other species, the major amino acids composed of Thr, whereas in *ERS1* and *ERS2*, Leu seemed to be the major amino acid for all the species analysed. The difference in the composition may affect the folding of secondary and tertiary structure of the proteins which in turn results in the different molecular weights.

Though the molecular weight of the ethylene receptors isolated is smaller compared to the complete sequence of the ethylene receptors retrieved, the average size of ethylene receptor proteins retrieved is within the range of 19kDa to 71kDA. *ETR1* proteins are within the range of 20kDA while the size of *ERS1* and *ERS2* reaches up to 71kDA. Analysis on solubilised ethylene receptors from *Arabidopsis* using gel filtration technique has shown that ethylene receptors exists as components of high mass protein complexes (Chen *et al.,* 2010).The data from that study suggests the existence of heterogeneity among the ethylene receptors in the form of their protein complexes. They also reported that upon ligand binding, different response were exhibited by *ETR1* and *ERS1*, which indicates a certain amount of independence between the protein complexes. Since the isolation of the five receptors from *Arabidopsis*, the number of ethylene receptors deposited in the GeneBank has increased tremendously. Members of the subfamily have also grown to include *ETR3- ETR7*.

In conclusion, this data obtained in this chapter has successfully characterise five genes that play are integral to the control of pollination induced senescence. The determination and analysis of these characteristics is important to ensure accurate identification of the role they play in both ethylene biosynthesis and perception. Moreover, this will also allow for effective downstream utilization which may include expression and manipulation of genes.