

Research Article

Bioinformatics Analysis of Phenylacetaldehyde Synthase (PAAS), a Protein Involved in Flower Scent Production, in Rose

Akbar Karami^{1*}, Samira Jandoust¹ and Esmaeil Ebrahimie^{2,3}

¹Department of Horticultural Science, College of Agriculture, Shiraz University, Shiraz, Iran ²Department of Crop Production and Plant Breeding, Faculty of Agriculture, Shiraz University, Shiraz, Iran ³Discipline of Genetics, School of Molecular and Biomedical Science, The University of Adelaide, Adelaide, Australia

Abstract

Rose flowers produce and emit the aromatic volatiles 2-phenylacetaldehyde (PAA) and 2-phenylethanol (2-PE), which have a distinctive flowery/rose-like scent. Previous studies of rose have shown that, similar to *Petunia* flowers, PAA is formed from L-phenylalanine via pyridoxal-5'-phosphate-dependent L-aromatic amino acid decarboxylase. *Rosa* phenylacetaldehyde synthase sequence (RhPAAS) is homologous to *Petunia* phenylacetaldehyde synthase (PhPAAS). Since there is not much experimental data available about different structural properties of that PAAS protein, in the present investigation, we studied the different structural properties of the PAAS protein in petunia and rose using bioinformatics tools. The features of the first, secondary and tertiary structures of this protein were compared between Petunia and Rose. The results indicated that the frequency of negatively charged, Leucine, and frequency of the Ser-Leu, Pro-Glu, Phe-Ser and the, Thr-Thr dipeptides in petunia are more than those in Rose. In contrast, in petunia, the frequencies of hydrophobic and hydrophilic residues, α-helix, β-sheet, β-strands of petunia are lower than those in rose. The features achieved in this study may also provide useful clues for designing scent production pathways using protein engineering techniques.

Keywords: Rose; Phenylacetaldehyde synthase; Bioinformatic; Protein; Feature; Prediction

Introduction

Phenyl acetaldehyde (PHA), 2-phenylethanol (2-PE), and its acetate ester are important scent compounds in numerous flowers, including petunias and roses. They also contribute to the aromas of tomato, grape, and tamarind, fruits and to the flavor of tea. PHA has been identified in some animals and fungi as well. Under aerobic conditions; phenylacetaldehyde synthase (PAAS) catalyze oxidative decarboxylation by a radical mechanism. PAAS is the first PLP enzyme to be described that, in its native state, catalyzes the stoichiometric oxidative decarboxylation of an L-amino acid substrate (Figure 1) [1-13]. PAAS by alignment is homologous with Aromatic L-amino acid decarboxylase (AADC) catalyzes the second enzymatic step in synthesis of the neurotransmitters dopamine and serotonin, which are found in neurons of all animals [4]. The *Arabidopsis thaliana* genome includes two genes, *At2g20340* and *At4g28680*, encoding pyridoxal 5'-phosphate-dependent AADCs with high homology to PhPAAS [8].

Since there is not much experimental data available about different structural properties of PAAS protein, in the present investigation, we studied different structural properties of PAAS protein by bioinformatics tools in *Petunia* and *Rosa*. The features of first, secondary and tertiary structure of this protein were compared between *Petunia* and *Rosa*.

Material and Methods

Amino acid sequence of the PAAS protein

The nucleotide sequence of PAAS was determined using Gen Bank (GenBank no:ABB72475.1 and ABB04522.1; http://www.ncbi.nih.gov/ gen bank/). According to Gen Bank, the PAAS protein is composed of 508 and 506 amino acid residues in RhPAAS and PhPAAS respectively (Table 1).

Analysis of the secondary structure of the PAAS protein

The protein sequence of the PAAS protein was input, and four

conformational states, including helices, sheets, turns and coils, were analyzed. Protein features (attributes) such as the counts and frequencies of each element (carbon, nitrogen, sulfur, oxygen, and hydrogen) and each amino acid, weight, isoelectric point, aliphatic index, N-terminal amino acids, half-life and some more features for double strand elements were extracted using various bioinformatics tools and software including CLC Bio software (CLC bio, Finlandsgade 10-12, Katrinebjerg 8200 Aarhus N Denmark). Prosite features describe domains, families and functional sites as well as associated patterns and profiles of each protein. ScanProsite site (http://www.expasy.org/tools/scanProsite/) was used as a web-based tool to compute Prosite signature matches in protein sequences of favorite proteins [14-17].

Prediction of the 3D structure of the PAAS protein

Predictive analysis of the PAAS protein tertiary structure was conducted using 3D Ligand Site, the online ligand-binding site prediction server (http://www.sbg.bio.ic.ac.uk). This web server automates the manual processes used for the prediction of ligandbinding sites in the eighth round of the critical assessment of protein structure prediction (CASP8), and is a useful tool for the analysis of protein tertiary structure [18].

Prediction of the protein-protein interactions (PPI)

The understanding protein-protein interaction (PPI) in sequence

*Corresponding author: Akbar Karami, Department of Horticultural Science, College of Agriculture, Shiraz University, Shiraz, Iran, Tel/Fax: +98(71)32286133; E-mail: akbarkarami@shirazu.ac.ir, akarami2004@gmail.com

Received: November 25, 2014; Accepted: January 09, 2015; Published: January 12, 2015

Citation: Karami A, Jandoust S, Ebrahimie E (2015) Bioinformatics Analysis of Phenylacetaldehyde Synthase (PAAS), a Protein Involved in Flower Scent Production, in Rose. J Phys Chem Biophys 5: 170. doi:10.4172/2161-0398.1000170

Copyright: © 2015 Karami A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Karami A, Jandoust S, Ebrahimie E (2014) Bioinformatics Analysis of Phenylacetaldehyde Synthase (PAAS), a Protein Involved in Flower Scent Production, in Rose. J Phys Chem Biophys 5: 170. doi:10.4172/2161-0398.1000170



Figure 1: 2-PE metabolic pathways. On the top, the plant metabolic pathway with two enzymatic steps is presented: the first step is the conversion of L-phe into phenylacetaldehyde by PAAS with PLP as the cofactor. The second step is the conversion of phenylacetaldehyde to 2-PE by adehydrogenase. On the bottom, the Ehrlich pathway abundant in yeast is shown including three well-described enzymatic steps. The scheme is based on the suggestions of Achmon et al. [1].

Protein Name	RhPAAS	Ph PAAS
AC. NO	Q0ZS27_ROSHC	Q0ZQX0_PETHY
PDB ID	1JS3	1JS3
Length	508	506
Weight (Kda)	56.371	57.077
Isoelectric point	7.03	5.77
Aliphatic index	91.358	87.688
N-terminal	Methionine	Methionine

 Table 1: Sequence information of PAAS in Rosa hybrid (RhPAAS) and Petunia hybrid (PhPAAS).

of PAAS was determined by using of STRING (http://STRING.embl. de) [17].

Results and Discussion

Protein features (attributes) such as the counts and frequencies of each element (carbon, nitrogen, sulfur, oxygen, and hydrogen) and each amino acid, weight, isoelectric point, aliphatic index, N-terminal amino acids, half-life and some more features for double strand elements of PAAS were extracted using various bioinformatics tool (Table S1-S6). The results indicated that, the major aromatic amino acid for PAAS (Phe) in Petunia is more than Rosa. The frequency of negatively charged, Leucin, and frequency of Ser-Leu dipeptid, Pro-Glu dipeptid, Phe-Ser dipeptid, Thr-Thr dipeptid in Petunia are more than Rosa. PAAS as a protein be a member of group II PLP-dependent amino acid decarboxylases with histidine, glutamate, serine, and aromatic L-amino-acid decarboxylases [11]. The content of Arg, Pro, His, Try were the most important features used to prediction of protein characteristics. Also the frequency of Asn-Gln, Gly-Gly and Asp-Pro were the most important features used to build the rest of tree. This tree was the best model to demonstrate the importance of dipeptides in protein function like thermostability [3].

In contrast, in *Petunia*, the frequency of hydrophobic and hydrophilic residues, α -helix, β -sheet, β -strand is lower than *Rosa*. Neither PhPAAS nor RhPAAS contains signal peptides at its N-terminal suggesting cytosolic localization [11]. Predictive analysis of the PAAS protein tertiary structure was conducted using 3D Ligand Site, are presented in Figure 2. PAL2 (L-Phe ammonia-lyase 2), PAL4 and COMT (caffeic acid 3-O-methyltransferase) were as protein-protein interaction in this pathway (Figure S1). Phenylalanine ammonia-lyase (PAL) catalyzes the first step of the phenylpropanoid pathway by the

deamination of L-phenylalanine to cinnamic acid, which synthesis the hundreds of most important phenolics group secondary metabolic products in plants comprising phenylpropanoids depravities, lignins, lignans, flavonoids and alkaloids [14]. In Arabidopsis thaliana have been annotated that PAL encoding by four genes including AtPAL1, 2, 3 and 4. In Arabidopsis thaliana have been shown that recombinant native AtPAL1, 2 and 4 were exhibited to be catalytically competent for L-phenylalanine deamination, whereas AtPAL3, achieved as a N-terminal His-tagged protein, was of very low activity and only measurable at high substrate concentrations. All four PALs demonstrated similar pH optima, but not temperature optima; AtPAL3 had a lower temperature optimum than the other three isoforms [2]. In plants, COMT an enzyme is belonging to lignin biosynthesis and converting caffeic acid to ferulic acid and 5-hydroxyferulic acid to sinapic acid. In Transgenic alfalfa plants have been shown that COMT cDNA sequences under control of the bean PAL2 promoter. This study demonstrated that strong downregulation of COMT resulted in decreased lignin content [7].

The transmembrane (TM) secondary structures of membrane proteins were predicted by using the method of preference functions (http://split.pmfst.hr/split/4). The predicted TM helix position of PhPAAS was 108-131 amino acids, however this trends was 251-265 in RhPAAS. Transmembrane domain usually denotes a single transmembrane α -helix of a transmembrane protein. Transmembrane domain is any three-dimensional protein structure which is thermodynamically stable in a membrane. The amino acid sequence or primary structure of a protein is the most important indication for its function. However, it is approved that prediction of protein characteristics from the primary amino acid sequence is not possible directly. Therefore, methods to predict protein characteristics have converged on tertiary and quaternary structures [3].

Understanding protein-protein interactions (PPI) is important for the analysis of intracellular signaling and metabolic pathways, and modeling of protein complex structures. When two or more than two proteins bind to each other protein-protein interaction occurs that frequently perform biological functions. Thus, knowledge of PPI can be significant insights into almost all biological processes within a cell. In addition to high-throughput experimental methods, computational approaches have been developed to predict functional associations between two proteins by extracting information from the genomic

Page 2 of 4

Page 3 of 4



context and evolutionary relationship. These computational methods include gene cluster, gene neighborhood, gene fusion, phylogenetic profile, in silico two-hybrid, and mirror tree methods [15].

Recently the paas synthetic gene was cloned into the pETDuet-1 vector that contains two multiple cloning sites. However, protein over expression or biological activity was not observed for PAAS. Inspection of the PAAS amino acid sequence revealed the presence of 13 cysteine residues, signifying the probability of disulfide bonds in the PAAS tertiary structure [1]. In unfavorable situations for disulfide bonds, configuration of the tertiary structure can cause damage and misfolding of the protein in E. coli. Such problems were explained for bovine β -lactoglobulin having five cysteine residues [12] and also for Nogo-A-specific exon 3 that has eight cysteine residues [6]. Export from the plasma membrane into the periclinal cell wall engages transfer of the comparatively unpolar scent molecules from a lipophilic into an aqueous section. The unpolar character of scent constituents can be explained using their octanol-water partition coefficients, a parameter evaluating their comparative solubilities in lipophilic and aqueous situations. This demonstrates the very low solubility of scent molecules in any aqueous location, which will significantly hinder their transfer across the cell wall. Noticeably both physicochemical processes, disconnection from the plasma membrane, and transport across the cell wall have to be assisted energetically, possibly by direct or indirect action of proteins [10].

Although a novel system showed for the first time the possibility of using the IMPACT system intein tag as a PAAS protein solubility supporter [1], intensifying information about the structure of the PAAS will be needed for the applying protein engineering.

Conclusion

In this study, various supervised and unsupervised tools were applied to identify the most important features contribute to the prediction of PAAS protein features. PAASs have been known only in scent-producing *P. hybrida* and *Rosa hybrida* [11]. PAASs are a member of group II of amino acid decarboxylases, and their general amino acid character makes it complicated to predict the actual function of the protein based on amino acid sequences. A major dispute in the postgenomic epoch is to find out protein function at the proteomic level. Most organisms contain a large number of proteins whose functions are presently unidentified. Normally, computational techniques to refer protein function have relied principally on sequence homology. However, the recent appearance of high-throughput techniques for determining protein interactions has facilitated a new line of study where protein function is predicted by utilizing interaction data [16]. Additionally to interaction networks that have been determined experimentally, there are a number of computational methods for building functional interaction networks, where two proteins are linked if they are predicted to complete a common natural mission. The molecular function of a protein demonstrated its biochemical activity, while its biological process specifies the role it plays in the cell or the pathway in which it contributed.

References

- Achmon Y, Ben-Barak Zelas Z, Fishman A (2014) Cloning Rosa hybrid phenylacetaldehyde synthase for the production of 2-phenylethanol in a whole cell Escherichia coli system. Appl Microbiol Biotechnol 98: 3603-3611.
- Cochrane FC, Davin LB, Lewis NG (2004) The Arabidopsis phenylalanine ammonia lyase gene family: kinetic characterization of the four PAL isoforms. Phytochemistry 65: 1557-1564.
- Ebrahimi M, Lakizadeh A, Agha-Golzadeh P, Ebrahimi E, Ebrahimi M (2011) Prediction of thermostability from amino acid attributes by combination of clustering with attribute weighting: a new vista in engineering enzymes. PLoS One 6: e23146.
- Facchini PJ, Huber-Allanach KL, Tari LW (2000) Plant aromatic L-amino acid decarboxylases: evolution, biochemistry, regulation, and metabolic engineering applications. Phytochemistry 54: 121-138.
- Farhi M, Lavie O, Masci T, Hendel-Rahmanim K, Weiss D, et al. (2010) Identification of rose phenylacetaldehyde synthase by functional complementation in yeast. Plant Mol Biol 72: 235-245.
- Fiedler M, Horn C, Bandtlow C, Schwab ME, Skerra A (2002) An engineered IN-1 F(ab) fragment with improved affinity for the Nogo-A axonal growth inhibitor permits immunochemical detection and shows enhanced neutralizing activity. Protein Eng 15: 931-941.
- Guo D, Chen F, Inoue K, Blount JW, Dixon RA (2001) Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa. impacts on lignin structure and implications for the biosynthesis of G and S lignin. Plant Cell 13: 73-88.
- Gutensohn M, Klempien A, Kaminaga Y, Nagegowda DA, Negre-Zakharov F, et al. (2011) Role of aromatic aldehyde synthase in wounding/herbivory response and flower scent production in different *Arabidopsis* ecotypes. The Plant Journal 66: 59-602.
- Hirata H, Ohnishi T, Ishida H, Tomida K, Sakai M, et al. (2012) Functional characterization of aromatic amino acid aminotransferase involved in 2-phenylethanol biosynthesis in isolated rose petal protoplasts. J Plant Physiol 169: 444-451.
- Jetter R (2006) Examination of the processes involved in the emission of scent volatiles from flowers. In: Dudareva N, Pichersky E (Eds) Biology of Floral Scent, Section III. Cell Biology and Physiology of Floral Scent, Taylor & Francis CRC Press, pp. 125-144.
- Kaminaga Y, Schnepp J, Peel G, Kish CM, Ben-Nissan G, Weiss D, Orlova I, Lavie O, Rhodes D, Wood K, Porterfield DM, Cooper AJL, Schloss JV, Pichersky E, Vainstein A, Dudareva N (2006) Plant phenylacetaldehyde synthase is a bifunctional homotetrameric enzyme that catalyzes phenylalanine decarboxylation and oxidation. J Biol Chem 281:23357-23366
- Ponniah K, Loo TS, Edwards PJ, Pascal SM, Jameson GB, et al. (2010) The production of soluble and correctly folded recombinant bovine betalactoglobulin variants A and B in Escherichia coli for NMR studies. Protein Expr Purif 70: 283-289.
- 13. Sakai M, Hirata H, Sayama H, Sekiguchi K, Itano H, et al. (2007) Production of 2-phenylethanol in roses as the dominant floral scent compound from L-phenylalanine by two key enzymes, a PLP-dependent decarboxylase and a phenylacetaldehyde reductase. Biosci Biotechnol Biochem 71: 2408-2419.

Citation: Karami A, Jandoust S, Ebrahimie E (2014) Bioinformatics Analysis of Phenylacetaldehyde Synthase (PAAS), a Protein Involved in Flower Scent Production, in Rose. J Phys Chem Biophys 5: 170. doi:10.4172/2161-0398.1000170

Page 4 of 4

- Shi R, Shuford CM, Wang JP, Sun YH, Yang Z, et al. (2013) Regulation of phenylalanine ammonia-lyase (PAL) gene family in wood forming tissue of *Populus trichocarpa*. Planta 238: 487-497.
- Shoemaker BA, Panchenko AR (2007) Deciphering protein-protein interactions. Part II. Computational methods to predict protein and domain interaction partners. PLoS Comput Biol 3: e43.
- 16. Singh M (2008) From protein interaction networks to protein function. In:

Panchenko A, Przytycka T (Eds) Protein-protein interactions and networks, Springer-Verlag London Limited.

- von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, et al. (2005) STRING: known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res 33: D433-437.
- Wass MN, Kelley LA, Sternberg MJ (2010) 3DLigandSite: predicting ligandbinding sites using similar structures. Nucleic Acids Res 38: W469-473.