Appl Microbiol Biotechnol (2014) 98:1165–1174
DOI 10.1007/s00253-013-4961-3

Escherichia coli protein YgiD produces the structural unit of plant pigments betalains: characterization of a prokaryotic enzyme with DOPA-extradiol-dioxygenase activity

Fernando Gandía-Herrero · Francisco García-Carmona

Received: 5 March 2013 / Revised: 26 April 2013 / Accepted: 27 April 2013 / Published online: 12 May 2013

Abstract Betalamic acid is the structural unit of all the natural pigments betalains. These are nitrogen-containing water-soluble compounds with high colorant and bioactive properties, characteristic of plants of the order Caryophyllales. The formation of betalamic acid from the precursor amino acid 3,4-dihydroxy-L-phenylalanine (L-DOPA) by the enzyme 4,5-DOPA-extradiol-dioxygenase was supposed to be restricted to plants of this order and two fungal species. Here, the first case of betalamic acid formation by an enzyme other than eukaryotes is reported with a homolog enzyme from *Escherichia coli*. The protein YgiD has been cloned, expressed, and purified to carry out its molecular and functional characterization. The enzyme was obtained as a monomeric active protein with a molecular mass of 32 kDa characterized by chromatography, electrophoresis, and MALDI-TOF analysis. Enzyme kinetic properties are characterized in the transformation of the relevant substrate L-DOPA. Reaction was analyzed spectrophotometrically and by HPLC-DAD, electrospray ionization mass spectrometry, and time-of-flight mass spectrometry.

Keywords Betalain · Betalamic acid · Dioxygenase · Extradiol · Plant pigment · Prokaryotic

Introduction

Betalains are water-soluble, nitrogen-containing pigments characteristic of plant species belonging to the order Caryophyllales (Gandía-Herrero and García-Carmona 2013). They are present in fruits, roots, stems, and leaves, where they fulfill the role played by anthocyanins in other plants (Brockington et al. 2011). Betalain-related pigments have also been described in two fungal species of the genera *Amanita* (Musso 1979; Stintzing and Schliemann 2007) and *Hygrocybe* (von Ardenne et al. 1974; Babos et al. 2011). Betalains are classified as either betacyanins, which exhibit a violet coloration, or betaxanthins, which are yellow pigments that in addition present green fluorescence (Gandía-Herrero et al. 2005a). Both groups share the molecule betalamic acid [4-(2-oxoethylidene)-1,2,3,4-tetrahydropyridine-2,6-dicarboxylic acid] as the same structural unit and differ in the molecule attached to it by condensation (Gandía-Herrero et al. 2010).

Betalains are molecules with a strong antiradical activity (Escribano et al. 1998; Gliszczyńska-Świgło et al. 2006; Gandía-Herrero et al. 2012). The betalamic acid resonance system is responsible for this activity, which is modulated by the condensed molecule in each individual pigment. Thus, betalamic acid is not only the structural backbone of betalains, but also its bioactive unit. Studies in vitro with different cell lines have revealed that betalain-containing extracts and pure pigments are active in the dose-dependent inhibition of cancer cell growth and proliferation (Sreekanth et al. 2007; Kapadia et al. 2011; Khan et al. 2012). Dietary betalains have also demonstrated a strong health-promoting potential in vivo by inhibiting the formation of tumors in model animals (Kapadia et al. 2003; Lu et al. 2009; Lechner et al. 2010).

All betalains are derived from betalamic acid by condensation, and this molecule is obtained by the specific activity of the enzyme 4,5-DOPA-extradiol-dioxygenase (4,5-DODA). Broadly, dioxygenases catalyze the incorporation of both atoms of molecular oxygen into catechol derivatives,

F. Gandía-Herrero (✉) · F. García-Carmona
Departamento de Bioquímica y Biología Molecular A, Unidad Docente de Biología, Facultad de Veterinaria, Regional Campus of International Excellence "Campus Mare Nostrum", Universidad de Murcia, Espinardo,
30100 Murcia, Spain
e-mail: fgandia@um.es

resulting in ring cleavage (Lipscomb 2008). The mechanism implies two histidine residues and a carboxylic group to activate molecular oxygen and the catecholic substrate through a reduced metal found in the active site. Depending on the position for the ring cleavage, the enzymes are classified as 1,2 (Ferraroni et al. 2012), 2,3 (Mbughuni et al. 2011), 3,4 (Orville et al. 1997), and 4,5-dioxygenases (Sugimoto et al. 1999). 4,5-DODA catalyzes the ring opening oxidation of 3,4-dihydroxy-L-phenylalanine (L-DOPA) to form the intermediate 4,5-seco-DOPA, which experiences a spontaneous cyclization to betalamic acid (Fig. 1). To date, only four enzymes have been described that perform this betalamic acid formation activity.

The first enzyme to be characterized was from the fungus *Amanita muscaria* (Girod and Zryd 1991; Mueller et al. 1997; Hinz et al. 1997). The fungal enzyme produces betalamic acid accompanied by another reaction product derived from L-DOPA, the pigment muscaflavin. This is due to a 2,3-DOPA-extradiol dioxygenase activity of the enzyme which leads to a 2,3-seco-DOPA intermediate in addition to the plant 4,5-seco-DOPA product (Fig. 1). The three other enzymes are characterized from betalain-producing plants and produce betalamic acid without the presence of muscaflavin. The function of the 4,5-*DODA* gene from *Portulaca grandiflora* responsible for betalain formation was confirmed by biolistic genetic complementation in white petals (Christinet et al. 2004). In this case, recombinant protein expression was not achieved. The 4,5-DODA protein from *Mirabilis jalapa* was successfully expressed in bacteria, allowing the formation of betalamic acid in vitro (Sasaki et al. 2009). Functional and molecular characterization was achieved for 4,5-DODA from *Beta vulgaris* after purification (Gandía-Herrero and García-Carmona 2012).

To date, the formation of betalamic acid from the precursor amino acid L-DOPA by the activity 4,5-DOPA-extradiol-dioxygenase has been supposed to be restricted to plants of the Caryophyllales and two fungal species. This paper describes the cloning, expression, purification, and molecular and functional characterization of the protein YgiD, a DODA homolog from *Escherichia coli* with known structure. For the first time an enzyme from a prokaryotic organism is described to produce the structural, chromophoric, and bioactive unit of the betalains.

Materials and methods

Chemicals, bacterial strains, plasmids, and enzymes

Solvents were from Merck Chemicals Ltd. (Dorset, England). HPLC-grade acetonitrile was purchased from Labscan Ltd. (Dublin, Ireland). *E. coli* Rosetta (DE3) cells and pET16b plasmid were from Novagen (Merck KGaA, Darmstadt, Germany). Restriction enzymes were obtained from New England BioLabs (Ipswich, MA, USA). T4 DNA ligase was from Roche Diagnostics (Basel, Switzerland). *Pfu* Turbo® DNA polymerase and its reaction buffer were from Stratagene (Agilent Technologies Inc., Santa Clara, CA, USA). QIAprep spin plasmid miniprep kit and QIAquick PCR purification kit were from QIAgen (Hilden, Germany). All other chemicals and reagents were obtained from Sigma (St. Louis, MO, USA).

ygiD sequence and cloning

The sequence for YgiD protein (Jw3007) from *E. coli* K-12 is deposited at National Center for Biotechnology

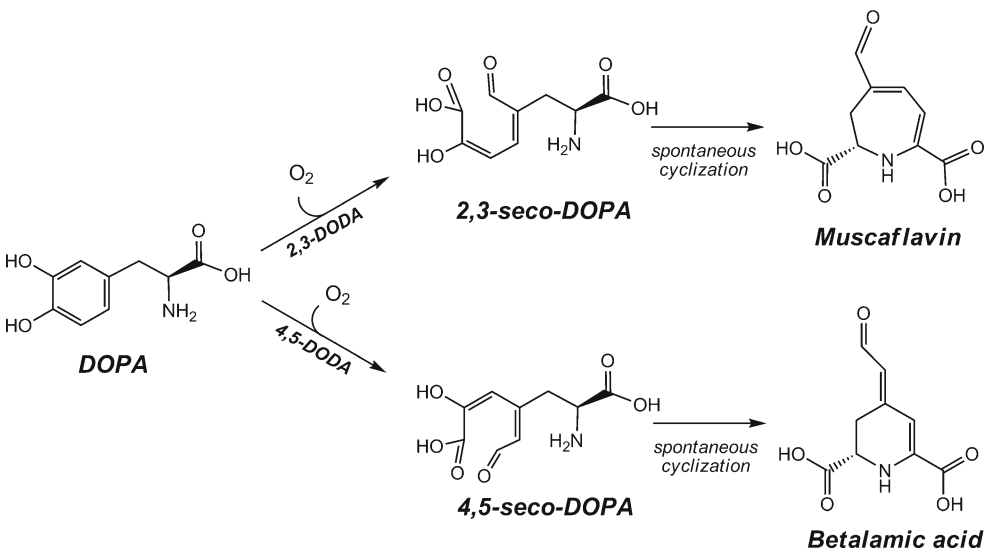

Fig. 1 Transformation of L-DOPA to 2,3- and 4,5-seco-DOPA compounds by 2,3- and 4,5-DOPA-extradiol-dioxygenase activities. Cyclization leads to muscaflavin and betalamic acid, respectively. The latter may experience further condensation with amine groups to form the pigments betalains

Information (NCBI, Bethesda, MD, USA) under the code gi:269849723 and at the Protein Data Bank (PDB Research Collaboratory for Structural Bioinformatics) under the code 2PW6. The gene for *ygiD* was cloned from *E. coli* strain K-12 substrain MG1655 obtained from the German collection of microorganisms and cell cultures (DSMZ, Braunschweig, Germany), code DSM 18039. PCR amplification was obtained using *Pfu* Turbo® DNA polymerase and the following oligonucleotides, which include the restriction sequences recognized by the enzymes *Nde*I and *Xho*I: EcDODA-F (5′-CGCATATGACACCTTTAGTTAAGGATA) and EcDODA-R (5′-ATCTCGAGTTAGCCTATCTGCAC). This amplification yielded a product of 829 bp, which represents the entire *ygiD* gene plus the sequences recognized by the restriction enzymes. The PCR product was digested with *Nde*I and *Xho*I, purified and inserted into the vector pET16b downstream of the T7 RNA polymerase promoter. This produced the recombinant plasmid pET16b-EcDODA, which encodes for an additional 22 amino acid N-terminal sequence containing a 10× His tag. Plasmid was transferred to *E. coli* XL-1 blue (Stratagene) and obtained from positive colonies. Sequence was checked by DNA sequencing of the plasmid and used in further experiments.

Expression and purification

The YgiD-His fusion protein was expressed in *E. coli* strain Rosetta (DE3)pLysS (Novagen), grown at 37ºC in LB containing ampicillin and chloramphenicol to an A_{600} of 0.8–1.0. Induction was then performed with different concentrations of the inducer isopropyl-1-thio-β-D-galactopyranoside (IPTG). Cells were harvested by centrifugation and resuspended in sodium phosphate buffer 50 mM, pH 8.0, with 0.3 M sodium chloride. Cell lysis was performed by sonication in a Cole-Parmer 4710 series ultrasonic homogenizer (Chicago, IL, USA). Chemical lysis was performed for trial scale and parameter optimization using Bugbuster protein extraction reagent (Novagen). Recombinant protein was purified by His-select nickel affinity gel (Sigma) according to the manufacturer's instructions and then desalted onto PD10 columns (General Electric Healthcare, Milwaukee, USA) using sodium phosphate buffer 20 mM, pH 8.5. Protein was quantified using the Bradford protein assay (Bio-Rad, Hercules, CA, USA) (Bradford 1976). Samples were analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Laemmli 1970), by application to 12 % polyacrylamide gels, and stained using a standard Coomassie Blue method.

Gel filtration

Samples of pure recombinant protein were applied to a Superdex 200 10/300 GL column equilibrated with 50 mM Tris–HCl buffer pH 7.5, containing 0.15 M NaCl. The protein was eluted with the same buffer at a flow rate of 0.5 mL min^{-1}. Elutions were performed in an Äkta purifier apparatus (General Electric Healthcare) and monitored at 280 nm. Column calibration was performed with the following protein markers (Sigma): cytochrome C (12.4 kDa), carbonic anhydrase (29 kDa), albumin (66 kDa), alcohol dehydrogenase (150 kDa), and β-amylase (200 kDa).

MALDI-TOF MS

Matrix solution for peptide analyses was α-cyano-4-hydroxycinnamic acid (20 mg/ml) in ACN/water/trifluoroacetic acid (TFA) (70:30:0.1). Peptide sample was dissolved in 0.1 % TFA and mixed with the matrix solution. One microliter of this mixture was applied to the atmospheric pressure matrix-assisted laser desorption ionization (AP-MALDI) target plate and allowed to dry. Experiments were carried out with an Agilent Time of Flight (TOF) Mass Spectrometer (Agilent Technologies, Santa Clara, CA, USA), equipped with an AP-MALDI Ion Source with a N_2 laser (337 nm). Samples were measured in reflectron mode to identify molecular formulas based on precise mass measurements in positive mode. External calibration of the spectrometer was performed with standard peptides from the ProteoMass™ Peptide MALDI-MS Calibration Kit (Sigma). Data were recorded and processed with Agilent MassHunter Workstation Software. Peptide Mass Fingerprint was analyzed using Agilent Spectrum Mill Software.

Trypsin digestion

The protein sample was prepared in 100 μl of buffer NH_4HCO_3 50 mM, pH 8.0, with 0.02 % ProteaseMAX™ Surfactant (Promega, Madison, WI, USA). Then, the sample was reduced with DTT 10 mM at 56 °C for 20 min and alkylated with iodoacetamide 50 mM at room temperature in the dark for 20 min. One microgram of proteomics grade trypsin (Promega) was added and the sample was incubated at 37 °C for 4 h. Finally, the sample was centrifuged at 15,000×*g* for 1 min to collect the condensate and 0.5 % TFA was added to stop the digestion. Peptides were cleaned up with C18 ZipTips (Millipore) and evaporated using an Eppendorf vacuum concentrator model 5301.

HPLC analysis

A Shimadzu LC-20AD apparatus (Kyoto, Japan) equipped with a SPD-M20A photodiode array detector was used for analytical HPLC separations performed with a 250×4.6-mm Kromasil 100 C-18 column (Teknokroma, Barcelona, Spain) (Gandía-Herrero et al. 2005b). Solvent A was water with 0.05 % TFA, and solvent B was acetonitrile with 0.05 % TFA. A linear gradient was performed for 25 min from 0 % B to 35 % B. Injection volume was 30 μL.

Standard betalamic acid was purified from beet root extracts by basic hydrolysis and further anionic exchange chromatography (Gandía-Herrero et al. 2012).

Elecrospray ionization mass analysis

An Agilent VL 1100 apparatus with LC/MSD Trap was used for HPLC-electrospray ionization mass spectrometry (ESI-MS) analyses. Elution conditions were analogous to those described above using the same column with a flow rate of 0.8 mL/min and gradient was performed for 30 min from 0 % B to 35 % B. Vaporizer temperature was 350 °C, and voltage was 3.5 kV. The sheath gas was nitrogen, operated at a pressure of 45 psi. Samples were ionized in positive mode. Ion monitoring mode was full scan in the range *m*/*z* 50–600. The electron multiplier voltage for detection was 1,350 V.

Accurate mass determination was carried out with a TOF/Q-TOF MS spectrometer Agilent 6220, equipped with a dual ESI-APCI interface. Samples were ionized in positive mode, using a capillary voltage of 3.5 kV. The gas temperature was 350 °C, nitrogen drying gas was set at 11 L min^{-1} and the nebulizer pressure was 40 psi. Data were processed through the software MassHunter (Agilent Technologies).

Absorbance spectroscopy

Enzyme activity was determined by a continuous spectrophotometric method by measuring the absorbance due to betalamic acid and muscaflavin appearance at λ=414 nm (Girod and Zryd 1991; Gandía-Herrero and García-Carmona 2012). Unless otherwise stated, the reaction medium contained 50 mM sodium phosphate buffer, pH 6.0, 2.5 mM L-DOPA, and 10 mM ascorbic acid (AA). Measurements were performed at 25 °C in 96-well plates in a Synergy HT plate reader (Bio-Tek Instruments, Winooski, USA). The final sample volume was 300 μL. The plate reader detector signal was calibrated with betalamic acid solutions of known concentration. For the quantification of betalamic acid, the molar extinction coefficient at 424 nm, ε=24,000 M^{-1} cm^{-1}, was taken (Trezzini and Zrÿd 1991). A Jasco V-630 spectrophotometer (Jasco Corporation, Tokyo, Japan) was used. Measurements were performed in triplicate and mean values and standard deviations were plotted. Errors associated with the results provided correspond to the residual standard deviations. Kinetic data analysis was carried out by using nonlinear regression fitting (Marquardt 1963).

Results

E. coli K-12 YgiD protein sequence

To date betalamic acid forming dioxygenases have only been described in plants belonging to the order Caryophyllales and the fungus *A. muscaria*, and can be considered the characteristic enzyme of the biosynthetic pathway of betalains. However, a close homolog was found in *E. coli*, the uncharacterized protein YgiD.

The sequence for YgiD protein (Jw3007) from *E. coli* K-12 is deposited at the databases under accession number gi:269849723 (NCBI). The crystal structure for the protein has been solved with a resolution of 2.27 Å by the high-throughput project Protein Structure Initiative (National Institutes of Health, Bethesda, MD, USA) and deposited at the PDB under the code 2PW6. The sequence in this database presents nine additional amino acids at the N-terminal end not present in gi:269849723. The nucleotides coding for them were retrieved from the whole genome for *E. coli* K-12 MG1655 (Blattner et al. 1997) as deposited at the NCBI (U00096.2 code). The full sequence for the YgiD protein as structurally characterized was considered in this work. The gene for *ygiD* was cloned from *E. coli* strain K-12 substrain MG1655 and sequenced. This confirmed the expected nucleotide sequence and the gene was amplified by PCR and inserted into the multiple cloning site of the expression vector pET16b, obtaining the plasmid pET16b-EcDODA.

YgiD expression and purification

The recombinant plasmid pET16b-EcDODA was used to transform *E. coli* Rosetta (DE3)pLysS cells. Optimization of the IPTG concentration and temperature of induction was performed using IPTG final concentrations of 0.1, 0.5, and 1.0 mM and temperatures of 20, 30, and 37 °C.

The highest soluble protein content was achieved after induction with 0.5 mM IPTG at 20 °C for 20 h. YgiD was expressed accounting for 13 % of the total soluble protein in the cell extract. Protein production for the following steps was performed in 0.5 L cultures of LB medium under the optimized conditions. YgiD was purified from *E. coli* cells by Ni^{2+}-chelating affinity chromatography. The His-tagged protein was eluted with buffer supplemented with 250 mM imidazole after non-tagged proteins were washed. The protein solution buffer was changed to sodium phosphate pH 8.5, 20 mM, and stored at 4 °C. The purification process was followed by denaturing SDS-PAGE, revealing that the *E. coli* DODA homolog protein was purified to homogeneity after the affinity separation, as can be seen in Fig. 2. Table 1 shows the results for the purification process.

Molecular characterization

The YgiD protein band obtained in SDS-PAGE yields an estimated molecular mass of 28 kDa, close to that calculated according to the protein sequence (32 kDa). In order to determine whether the prokaryotic DODA homolog is a monomer or whether it forms oligomers, purified YgiD

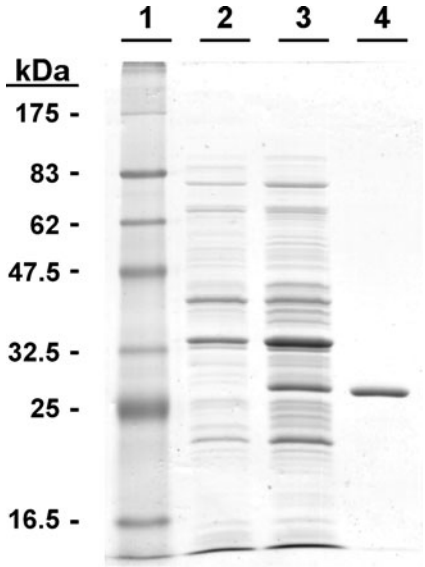

Fig. 2 Electrophoretic analysis for the expression and purification of YgiD protein in *E. coli* strain Rosetta (DE3) pLysS cells. *Lane 1*: molecular weight markers; *lane 2*: soluble protein content of cells harvested prior to IPTG induction; *lane 3*: soluble protein content of cells harvested 20 h after IPTG induction (0.5 mM); *lane 4*: eluted protein after affinity chromatography purification

protein was submitted to gel filtration under native conditions. As can be seen in Fig. 3, a single peak was eluted. Different amounts of starting sample were tested, ranging from 10 μg to 0.43 mg of purified protein, in order to consider a possible equilibrium among the monomer and one or more oligomers. A consistent single peak eluting at the same volume was obtained in all cases, with a molecular mass estimated at 18 kDa. This value is lower than the expected mass according to SDS-PAGE but points to the existence of YgiD as a monomer even at the highest concentrations assayed.

To determine with precision the absolute molecular mass of YgiD, the purified protein obtained from gel filtration was subjected to MALDI-TOF analysis. MALDI mass spectra of the sample showed a single peak with a molecular mass of 32,296 Da, consistent with the protein sequence. To fully characterize the protein, its peptide mass fingerprint was determined by MALDI-TOF analysis after trypsin digestion. The main peptides identified corresponded to the masses 1,238.86 *m/z* [(R)MPALFLGHGSPMNVLEDNLYTR(S)], 1,191.06 *m/z* [(K)LGMTLPRPQAIVVVSAHWFTR(G)], 1,174.42 *m/z* [(K)LAALRDEGIMLVASGNVVHNLR(T)], 1,137.06 *m/z* [(K)WHGDSSPYPWATSFNEYVK(A)], and 851.71 *m/z* [(K)EAWGFDHGSWGVLIK(M)].

Functional characterization

In order to evaluate the existence of an uncharacterized enzymatic activity for *E. coli* protein YgiD, the purified enzyme was added to a suitable reaction medium for the determination of dioxygenase activity. It contained the aromatic amino acid L-DOPA as a substrate and the reaction mixture was analyzed by HPLC after 30 min. This generated three peaks which were not visible in the absence of the enzyme (Fig. 4).

The existence of a betalamic acid forming 4,5-DOPA-extradiol-dioxygenase activity was explored by using real betalamic acid as a positive control. HPLC analysis resulted in the identification of betalamic acid in the reaction mixture by co-elution experiments. The retention time was 14.93 min (Fig. 4, peak 2), and the superimposable spectra showed a λ_{max}=405 nm under the analysis conditions, confirming the expected product for a 4,5-DODA activity.

In addition, a peak with a retention time of 7.99 min (Fig. 4, peak 1) and a spectrum with a maximum wavelength of λ_{max}=363 nm was obtained. This is in agreement with the results obtained for *M. jalapa* 4,5-DOPA-dioxygenase, where an analogous peak was identified as the intermediate 4,5-seco-DOPA (Sasaki et al. 2009). The reaction media of YgiD also present a peak with a retention time of 16.85 min (Fig. 4, peak 3) with a maximum wavelength of λ_{max}= 405 nm and a secondary maximum with wavelength of λ_{max}=236 nm. The compound might correspond to an additional activity present in YgiD distinct to the 4,5-DODA, as happens with the enzyme from *A. muscaria* with the 2,3-extradiol-dioxygenase activity (Mueller et al. 1997). The derived product from this activity is muscaflavin, for which an identical spectrum to peak 3 has been reported (λ_{max}= 404 nm, 237 nm) and also elutes after betalamic acid in reversed phase chromatography analyses (Stintzing and Schliemann 2007). Thus, peaks 1 and 3 were tentatively identified as 4,5-seco-DOPA and muscaflavin, respectively.

Table 1 Expression and purification of *E. coli* protein YgiD

	Volume (mL)	Protein (mg/mL)	Total protein (mg)	Activity[a] (μM min^{-1})	Specific activity (μmol min^{-1} mg^{-1})	Purification fold	Yield (%)
Crude extract[b]	6.0	16.4	98.4	0.843	0.307	1.0	100
Ni^{2+} chromatography	7.0	3.0	21.2	0.661	1.310	4.3	91

[a] Activity was determined using 50 μL of protein solution under the assay conditions described in the "Materials and methods" section

[b] Crude extract was obtained from a cellular paste coming from a 0.5-L culture

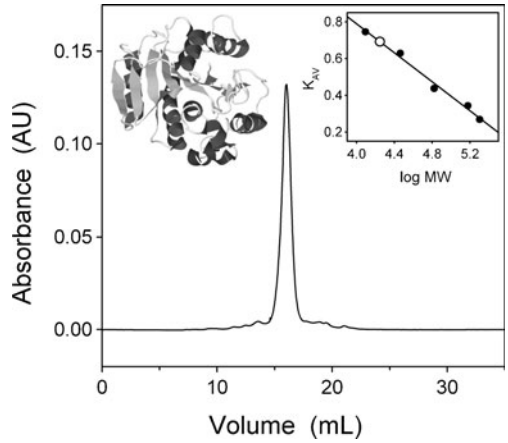

Fig. 3 Analysis of *E. coli* protein YgiD by gel filtration chromatography. The chromatogram was obtained by injecting 100 μL of a solution containing the purified protein at a concentration of 1.0 mg/mL and followed at λ=280 nm. The profile shows a single peak corresponding to a monomeric form. *Insets*: Structure for YgiD as deposited in PDB, entry 2PW6, and the calibration curve determined with the molecular mass markers described in "Materials and methods" section

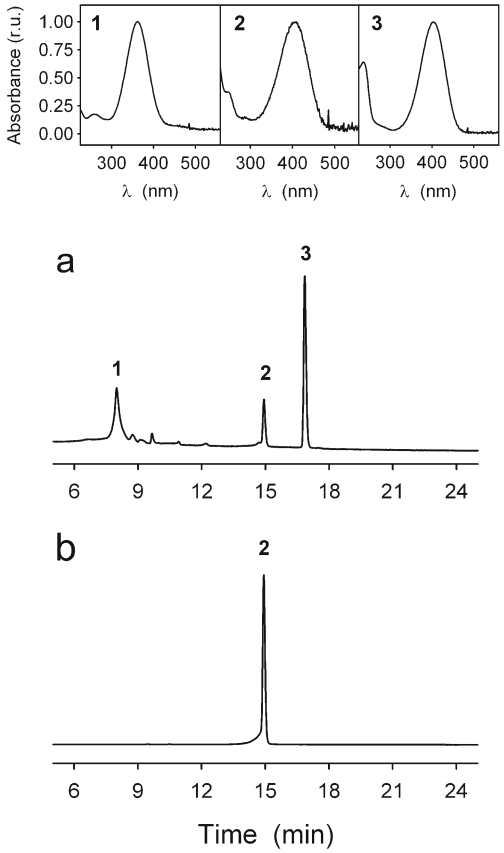

Fig. 4 HPLC analysis of reaction products formed by protein YgiD activity. **a** Chromatogram obtained at λ=405 nm for a reaction medium containing 2.5 mM L-DOPA and 10 mM AA in phosphate buffer 50 mM, pH 8.5 at 25 °C, 30 min after the addition of pure *E. coli* protein YgiD. Full scale is 25 mAU. *Insets*: Spectra obtained under the chromatographic separation conditions for peaks 1 (4,5-seco-DOPA), 2 (betalamic acid), and 3 (muscaflavin). **b** Chromatogram obtained at λ=405 nm for a real betalamic acid sample purified from *B. vulgaris* roots. Full scale is 500 mAU. In all cases, 30 μL was injected

Mass spectrometry analysis of reaction products

The nature of the products of YgiD activity was further characterized by ESI-MS. The mass values determined for peaks 2 and 3 was 212 *m*/*z*, which is in agreement with the expected masses of the protonated molecular ions $[M+H]^+$ of betalamic acid and muscaflavin. Both isomers possess a similar fragmentation pattern, with the main daughter ions corresponding to 166 *m*/*z*. This result was reproduced for the betalamic acid obtained from *B. vulgaris* roots and is in agreement with previous results for the MS analysis of betalamic acid and muscaflavin (Stintzing and Schliemann 2007; Gandía-Herrero et al. 2012). The same loss $[M+H]^+$ −46 *m*/*z* was found as the characteristic daughter ion in the related molecule dopachrome (Lemos-Amado et al. 2001). Our study also detected this as a secondary daughter ion of 152 *m*/*z* for the enzyme substrate L-DOPA, whose protonated molecular ion corresponded to a mass of 198 *m*/*z* (main daughter ion was 181 *m*/*z*).

In order to determine the exact mass of the protonated molecular ions for betalamic acid and muscaflavin, the reaction media were analyzed by time-of-flight mass spectrometry (TOF-MS). The ESI accurate mass spectra showed the detection of $[M+H]^+$ with an accurate mass of 212.0552 *m*/*z* (experimental mass), which matches the calculated betalamic acid and muscaflavin mass (*m*/*z* 212.0559) with an error of 3.3 ppm. This is below the accepted accuracy threshold for elemental composition analysis, established at 5 ppm (Ferrer et al. 2005) and confirms $C_9H_{10}O_5N$ as the molecular formula for the ions derived from peaks 2 (betalamic acid) and 3 (muscaflavin). The mass corresponding to peak 1 could not be determined in this study due to the low ionization of the chemical species.

Kinetic characterization

DODA activity of YgiD was characterized spectrophotometrically. The addition of the enzyme caused the appearance of yellow color in the reaction media with a λ_{max}= 414 nm (results not shown). Spectral changes were not observed in the absence of the enzyme and, therefore, they were considered to be the result of YgiD activity. The activity described is in accordance with previous observations on the absorbance characteristics of betalamic acid and muscaflavin extracted from plant and fungal material (Mueller et al. 1997; Trezzini and Zrÿd 1991; Gandía-Herrero et al. 2012) and of those formed in vitro by the *A. muscaria*, *M. jalapa*, and *B. vulgaris* dioxygenases (Girod and Zryd 1991; Sasaki et al. 2009; Gandía-Herrero and García-Carmona 2012).

The optimum pH for the DODA activity of YgiD was determined as pH=8.0, with a very small variation in the range 7.5–8.5. As can be seen in Fig. 5a, the pH effect on enzyme activity led to a marked increase of activity at pH values above 6.0. Dependence of the activity rate on substrate concentration was analyzed at the optimum pH. An increase in L-DOPA concentration resulted in increased enzymatic activity, as can be seen in Fig. 5b. Fitting the steady-state rates to the Michaelis–Menten equation yielded a value of 7.9±2.5 mM for the Km constant. Values for the maximum rate and the catalytic constant were determined as Vm=2.6±0.6 μM min^{-1} and k_{cat}=0.17±0.04 min^{-1}, respectively.

Betalamic acid-producing homologs

YgiD protein is a sequence homolog to the 4,5-extradiol-dioxygenases from plants of the order Caryophyllales which experimentally demonstrated to produce betalamic acid: *P. grandiflora* (gi:33636343), *M. jalapa* (gi:209976918), and *B. vulgaris* (gi:34850079) (Christinet et al. 2004; Sasaki et al. 2009; Gandía-Herrero and García-Carmona 2012). Identity ranges from 30.2 % (48.9 % similarity) for the enzyme from *M. jalapa* to 36.7 % (52.9 % similarity) for the enzyme obtained from *P. grandiflora*. The enzyme from *A. muscaria* (gi:2072624) produces betalamic acid but possesses less sequence analogy with the above-cited enzymes or YgiD (22.2 % identity). Figure 6 shows the multiple sequence alignment for the three active enzymes from plants with YgiD. As can be seen, the novel 4,5-DODA activity described in this work is supported by the conservation of amino acid blocks with the enzymes from betalain-producing plants. On analyzing the active site for the structure deposited for YgiD (PDB:2PW6), it is seen that among the highly conserved amino acids are those around the three strictly conserved histidine residues involved in metal coordination in the *E. coli* protein (His22, His57, and His234).

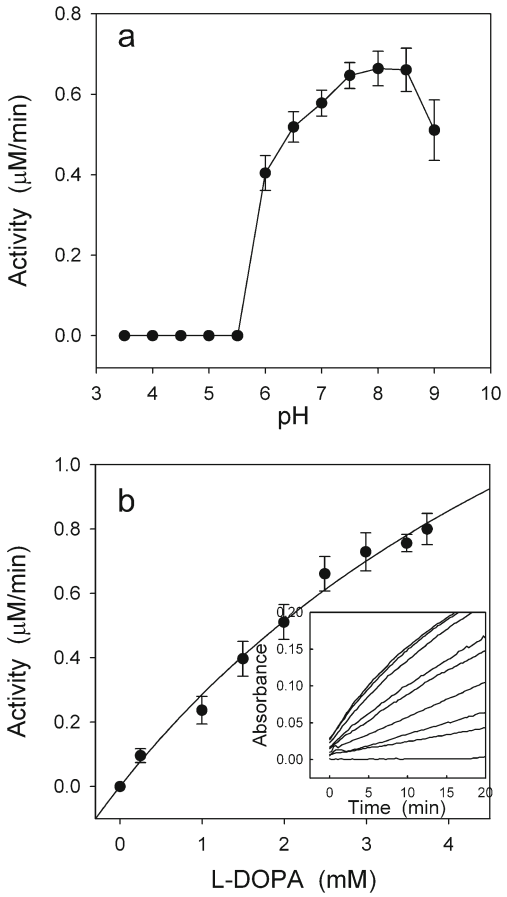

Fig. 5 Protein YgiD activity characterization. **a** Effect of pH on *E. coli* YgiD activity. Reactions were performed with 2.5 mM L-DOPA in 50 mM sodium acetate buffer for pH values ranging from 3.5 to 5.5 and 50 mM sodium phosphate for 5.5 to 9.0. **b** Enzyme activity dependence on L-DOPA concentration measured in 50 mM sodium phosphate buffer pH 8.0. *Inset*: Continuous recordings at λ=414 nm for the activity measurements presented in the main panel. In both cases, the total protein concentration was 0.5 mg/mL

From the sequence homology, and using the structural information provided by the YgiD crystal, the amino acids present in the plant enzymes can be assigned to specific structural motifs (Fig. 6). The residues that form α-helices and β-strands appear highly conserved in plants with respect to the information provided by YgiD, except for strands β8 and β10. By comparison with YgiD, the plant proteins possess seven α-helices and up to ten β-strands; these configure an α/β-sandwich with a mixed β-sheet in the core of the protein with helices α1 and α3 on one side and helices α2, α4, α5, α6, and α7 on the other. The metal coordination defines the active site which could be found in the crevice outside the carboxy ends of strands β2 and β6 (Lipscomb 2008). As can be seen in Fig. 6, α-helices and β-strands alternate along the polypeptide chain.

Discussion

Transformation of L-DOPA to 4,5-seco-DOPA and its further cyclization to betalamic acid constitute the key steps in the formation of the plant pigments betalains (Fig. 1). The reaction is catalyzed by the enzyme 4,5-DOPA-extradiol-dioxygenase, supposed to be specific of plants of the order Caryophyllales (Gandía-Herrero and García-Carmona 2013) and two fungal species (Stintzing and Schliemann 2007; Babos et al. 2011). Only one fungal (Girod and Zryd 1991) and three plant enzymes (Christinet et al. 2004; Sasaki et al. 2009; Gandía-Herrero and García-Carmona 2012) had previously been studied.

The uncharacterized protein YgiD from *E. coli* is a close homolog of known structure for the plant enzymes (Fig. 6). YgiD protein possesses a molecular mass of 32 kDa and it is present as an active monomer (Fig. 3). In plants, the

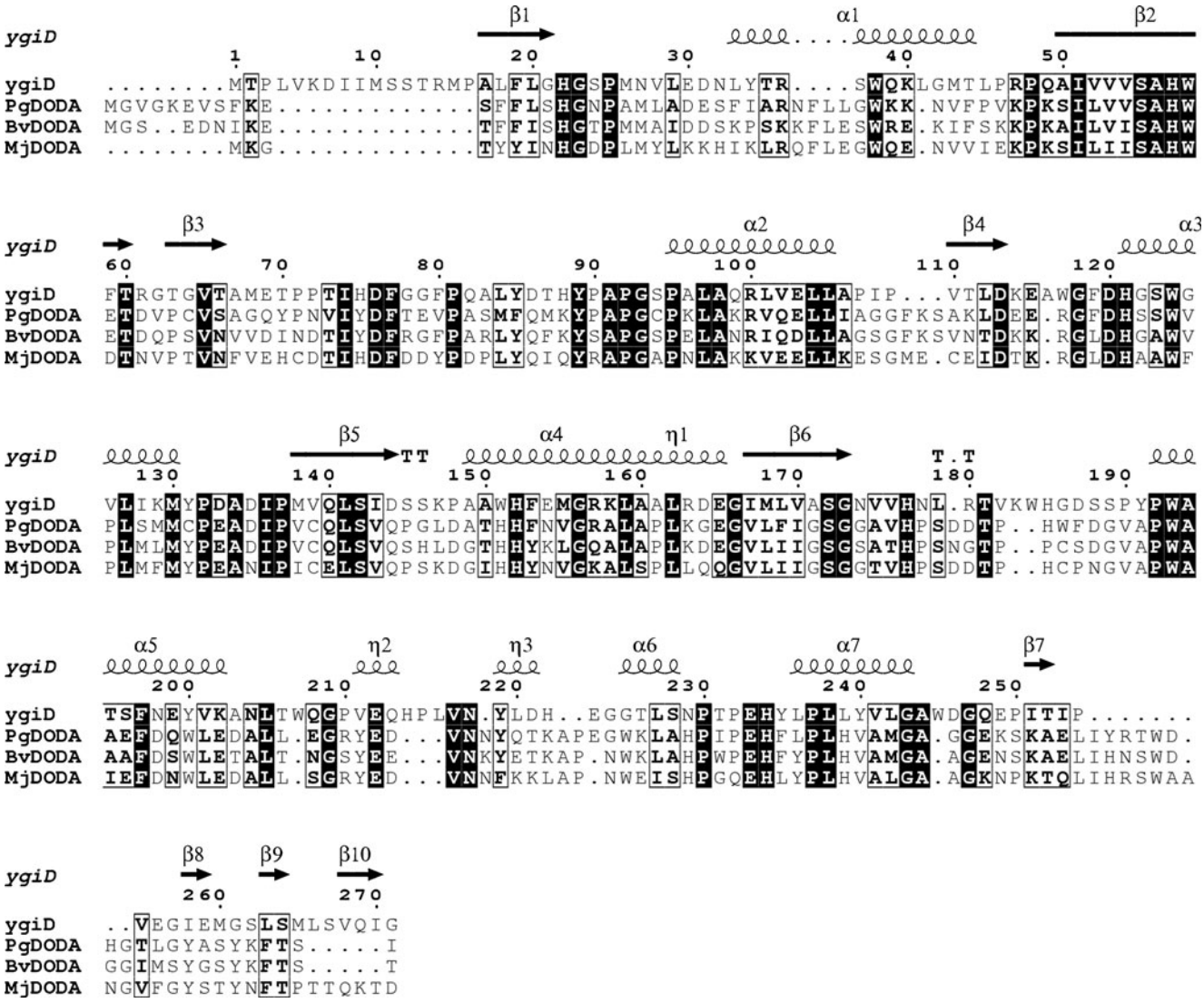

Fig. 6 Sequence comparison of YgiD protein with active plant 4,5-dioxygenases. Multiple sequence alignment using structural information was performed with Expresso (Armougom et al. 2006). Conserved blocks of amino acids are *squared* and strictly conserved residues are shown in *black*. Information is displayed with ESPript program (Gouet et al. 1999). Sequence accession codes (NCBI) for plant homologs are: gi:33636343 (*P. grandiflora*), gi:209976918 (*M. jalapa*), and gi:34850079 (*B. vulgaris*)

homolog enzyme from *B. vulgaris* was also found to be a monomer of 32 kDa (Gandía-Herrero and García-Carmona 2012). However, in fungi, the betalamic acid-forming enzyme was described as oligomers of a 22-kDa unit (Girod and Zryd 1991).

YgiD catalyzes the transformation of L-DOPA to colored compounds that were identified as betalamic acid and muscaflavin (Fig. 4). The analysis of the reaction products was carried out by HPLC-DAD, ESI-MS, and TOF-MS. Analogous peaks for betalamic acid formation have been described previously for the analysis of the reaction media of the 4,5-DODA enzymes from *M. jalapa* and *B. vulgaris* (Sasaki et al. 2009; Gandía-Herrero and García-Carmona 2012). YgiD constitutes the first case reported for the formation of betalamic acid by a prokaryotic dioxygenase enzyme.

The characterization of YgiD activity revealed an optimum pH of 8.0. This is in line with the data obtained for the other two betalamic acid-forming enzymes characterized at this level, the dioxygenases from *A. muscaria* and *B. vulgaris*. In both cases, an optimum pH value of 8.5 has been described (Girod and Zryd 1991; Gandía-Herrero and García-Carmona 2012). Kinetic analysis yielded a value of 7.9±2.5 mM for the Michaelis–Menten constant, Km. This value is similar to the results obtained for the plant and fungal enzymes previously characterized. A Km value of 6.9 mM was reported for *B. vulgaris* 4,5-DODA (Gandía-Herrero and García-Carmona 2012), and a Km value of

3.9 mM was found for the enzyme in *A. muscaria* (Girod and Zryd 1991). Although the actual role of YgiD in *E. coli* is unknown, dioxygenases are versatile enzymes that usually play a role in the metabolic degradation of aromatic compounds in bacterial species (Khomenkov et al. 2008).

The protein YgiD from *E. coli* has been expressed, purified to homogeneity, and characterized for the first time at molecular and kinetic levels. The protein possesses activity 4,5-DOPA-extradiol-dioxygenase and is found to be a homolog of the enzymes involved in the formation of betalains in plants. YgiD is a monomeric enzyme and the activity against the relevant substrate L-DOPA has been characterized. For the first time, a dioxygenase from a prokaryotic organism is described to produce muscaflavin and betalamic acid, the structural, chromophoric, and bioactive unit of the plant pigments betalains. This work opens up a new line in the efforts to obtain betalamic acid and its derived bioactive molecules by using enzymes of bacterial origin.

Acknowledgments This work was supported by the Ministerio de Ciencia e Innovación (MICINN, FEDER, Spain) (Project AGL2011-25023) and by Programa de Ayudas a Grupos de Excelencia de la Región de Murcia, of the Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia (Plan Regional de Ciencia y Tecnología 2007/2010). F. Gandía-Herrero holds a contract with the "Programa Ramón y Cajal" (MICINN, FEDER, Spain). The authors are grateful to Dr. Alejandro Torrecillas and to Dr. José Rodriguez (SAI, University of Murcia) for skillful technical assistance in MALDI-TOF and ESI-MS experiments, respectively.

References

Armougom F, Moretti S, Poirot O, Audic S, Dumas P, Schaeli B, Keduas V, Notredame C (2006) Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. Nucleic Acids Res 34:604–608. doi:10.1093/nar/gkl092

Babos M, Halasz K, Zagyva T, Zold-Balogh A, Szego D, Bratek Z (2011) Preliminary notes on dual relevance of ITS sequences and pigments in *Hygrocybe* taxonomy. Persoonia 26:99–107. doi:10.3767/003158511X578349

Blattner FR, Plunkett G III, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, Gregor J, Davis NW, Kirkpatrick HA, Goeden MA, Rose DJ, Mau B, Shao Y (1997) The complete genome sequence of *Escherichia coli* K-12. Science 277:1453–1462. doi:10.1128/JB.01695-07

Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein dye-binding. Anal Biochem 72:248–258. doi:10.1016/0003-2697(76)90527-3

Brockington SF, Walker RH, Glover BJ, Soltis PS, Soltis DE (2011) Complex pigment evolution in the Caryophyllales. New Phytol 190:854–864. doi:10.1111/j.1469-8137.2011.03687

Christinet L, Burdet FX, Zaiko M, Hinz U, Zrÿd J-P (2004) Characterization and functional identification of a novel plant 4,5-extradiol dioxygenase involved in betalain pigment biosynthesis in *Portulaca grandiflora*. Plant Physiol 134:265–274. doi:10.1104/pp.103.031914

Escribano J, Pedreño MA, García-Carmona F, Muñoz R (1998) Characterization of the antiradical activity of betalains from *Beta vulgaris* L. roots. Phytochem Anal 9:124–127. doi:10.1002/(SICI)1099-1565(199805/06)

Ferraroni M, Matera I, Steimer L, Bürger S, Scozzafava A, Stolz A, Briganti F (2012) Crystal structures of salicylate 1,2-dioxygenase-substrates adducts: a step towards the comprehension of the structural basis for substrate selection in class III ring cleaving dioxygenases. J Struct Biol 177:431–438. doi:10.1016/j.jsb.2011.11.026

Ferrer I, García-Reyes JF, Fernandez-Alba A (2005) Identification and quantitation of pesticides in vegetables by liquid chromatography time-of-flight mass spectrometry. Trends Anal Chem 24:671–682. doi:10.1016/j.trac.2005.04.004

Gandía-Herrero F, García-Carmona F, Escribano J (2005a) Floral fluorescence effect. Nature 437:334. doi:10.1038/437334a

Gandía-Herrero F, García-Carmona F, Escribano J (2005b) A novel method using high-performance liquid chromatography with fluorescence detection for the determination of betaxanthins. J Chromatogr A 1078:83–89. doi:10.1016/j.chroma.2005.05.013

Gandía-Herrero F, Escribano J, García-Carmona F (2010) Structural implications on color, fluorescence, and antiradical activity in betalains. Planta 232:449–460. doi:10.1007/s00425-010-1191-0

Gandía-Herrero F, Escribano J, García-Carmona F (2012) Purification and antiradical properties of the structural unit of betalains. J Nat Prod 75:1030–1036. doi:10.1021/np200950n

Gandía-Herrero F, García-Carmona F (2012) Characterization of recombinant *Beta vulgaris* 4,5-DOPA-extradiol-dioxygenase active in the biosynthesis of betalains. Planta 236:91–100. doi:10.1007/s00425-012-1593-2

Gandía-Herrero F, García-Carmona F (2013) Biosynthesis of betalains: yellow and violet plant pigments. Trends Plant Sci. doi:10.1016/j.tplants.2013.01.003

Girod PA, Zryd JP (1991) Biogenesis of betalains: purification and partial characterization of DOPA 4,5-dioxygenase from *Amanita muscaria*. Phytochemistry 30:169–174. doi:10.1016/0031-9422(91)84119-D

Gliszczyńska-Świgło A, Szymusiak H, Malinowska P (2006) Betanin, the main pigment of red beet: molecular origin of its exceptionally high free radical-scavenging activity. Food Add Cont 23:1079–1087. doi:10.1080/02652030600986032

Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: multiple sequence alignments in PostScript. Bioinformatics 15:305–308. doi:10.1093/bioinformatics/15.4.3

Hinz UG, Fivaz J, Girod PA, Zyrd JP (1997) The gene coding for the DOPA dioxygenase involved in betalain biosynthesis in *Amanita muscaria* and its regulation. Mol Gen Genet 256:1–6. doi:10.1007/s004380050539

Kapadia GJ, Azuine MA, Sridhar R, Okuda Y, Tsuruta A, Ichiishi E, Mukainake T, Takasaki M, Konoshima T, Nishino H, Tokuda H (2003) Chemoprevention of DMBA-induced UV-B promoted, NOR-1-induced TPA promoted skin carcinogenesis, and DEN-induced phenobarbital promoted liver tumors in mice by extract of beetroot. Pharmacol Res 47:141–148. doi:10.1016/S1043-6618(02)00285-2

Kapadia GJ, Azuine MA, Rao GS, Arai T, Iida A, Tokuda H (2011) Cytotoxic effect of the red beetroot (*Beta vulgaris* L.) extract compared to doxorubicin (adriamycin) in the human prostate (PC-3) and breast (MCF-7) cancer cell lines. Anti-Cancer Agents Med Chem 11:280–284. doi:10.2174/187152011795347504

Khan MI, Harsha PSCS, Giridhar P, Ravishankar GA (2012) Pigment identification, nutritional composition, bioactivity, and in vitro cancer cell cytotoxicity of *Rivina humilis* L. berries, potential source of betalains. LWT - Food Sci Technol 47:315–323. doi:10.1016/j.lwt.2012.01.025

Khomenkov VG, Shevelev AB, Zhukov VG, Zagustina NA, Bezborodov AM, Popov VO (2008) Organization of metabolic

pathways and molecular-genetic mechanisms of xenobiotic degradation in microorganisms: a review. Appl Biochem Microbiol 44:117–135. doi:10.1134/S0003683808020014

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227:680–685. doi:10.1038/227680a0

Lechner JF, Wang LS, Rocha CM, Larue B, Henry C, McIntyre CM, Riedl KM, Schwartz SJ, Stoner GD (2010) Drinking water with red beetroot food color antagonizes esophageal carcinogenesis in N-nitrosomethylbenzylamine-treated rats. J Med Food 13:733–739. doi:10.1089/jmf.2008.0280

Lemos-Amado F, Domingues P, Ferrer-Correia A, Remião F, Milhazes N, Borges F, Carvalho FD, Bastos ML (2001) Electrospray tandem mass spectrometry of aminochromes. Rapid Commun Mass Spectrom 15:2466–2471. doi:10.1002/rcm.498

Lipscomb JD (2008) Mechanism of extradiol aromatic ring-cleaving dioxygenases. Curr Op Struct Biol 18:644–649. doi:10.1016/j.sbi.2008.11.001

Lu X, Wang Y, Zhang Z (2009) Radioprotective activity of betalains from red beets in mice exposed to gamma irradiation. Eur J Pharmacol 615:223–227. doi:10.1016/j.ejphar.2009.04.064

Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. J Soc Ind Appl Math 11:431–441. doi:10.1137/0111030

Mbughuni MM, Chakrabarti M, Hayden JA, Meier KK, Dalluge JJ, Hendrich MP, Münck E, Lipscomb JD (2011) Oxy intermediates of homoprotocatechuate 2,3-dioxygenase: facile electron transfer between substrates. Biochemistry 50:10262–10274. doi:10.1021/bi201436n

Mueller LA, Hinz U, Zryd JP (1997) The formation of betalamic acid and muscaflavin by recombinant DOPA-dioxygenase from *Amanita*. Phytochemistry 44:567–569. doi:10.1016/S0031-9422(96)00625-5

Musso H (1979) Pigments of fly agaric, *Amanita muscaria*. Tetrahedron 35:2843–2853. doi:10.1016/S0040-4020(01)99498-0

Orville AM, Lipscomb JD, Ohlendorf DH (1997) Crystal structures of substrate and substrate analog complexes of protocatechuate 3,4-dioxygenase: endogenous Fe^{3+} ligand displacement in response to substrate binding. Biochemistry 36:10052–10066. doi:10.1021/bi970469f

Sasaki N, Abe Y, Goda Y, Adachi T, Kasahara K, Ozeki Y (2009) Detection of DOPA 4,5-dioxygenase (DOD) activity using recombinant protein prepared from *Escherichia coli* cells harboring cDNA encoding *DOD* from *Mirabilis jalapa*. Plant Cell Physiol 50:1012–1016. doi:10.1093/pcp/pcp053

Sreekanth D, Arunasree MK, Roy KR, Reddy TC, Reddy GV, Reddanna P (2007) Betanin, a betacyanin pigment purified from fruits of *Opuntia ficus-indica* induces apoptosis in human chronic myeloid leukemia cell line-K562. Phytomedicine 4:739–746. doi:10.1016/j.phymed.2007.03.017

Stintzing F, Schliemann W (2007) Pigments of fly agaric (*Amanita muscaria*). Z Naturforsch 62c:779–785

Sugimoto K, Senda T, Aoshima H, Masai E, Fukuda M, Mitsui Y (1999) Crystal structure of an aromatic ring opening dioxygenase LigAB, a protocatechuate 4,5-dioxygenase, under aerobic conditions. Structure 7:953–965. doi:10.1016/S0969-2126(99)80122-1

Trezzini GF, Zrÿd J-P (1991) Characterization of some natural and semisynthetic betaxanthins. Phytochemistry 30:1901–1904. doi:10.1016/0031-9422(91)85036-Y

von Ardenne R, Döpp H, Musso H, Steiglich W (1974) Über das Vorkommen von Muscaflavin bei Hygrocyben (Agaricales) und seine Dihydroazepin-Struktur. Z Naturforsch 29c:637–639