

A CAROTENOID DIOXYGENASE AND METHODS FOR THE BIOTECHNOLOGICAL PRODUCTION IN MICROORGANISMS AND PLANTS OF COMPOUNDS DERIVED FROM SAFFRON

DESCRIPTION

The present invention relates to a new carotenoid dioxygenase and methods for the biological production in microorganisms and plants of compounds with high added value derived from saffron.

The present invention relates in particular to genes, proteins and methods for the production of apocarotenoids of saffron in *in vitro* systems, in engineered strains of *E. coli* and in plant tissues. More specifically, the present invention relates to the isolation and characterisation of the saffron gene *CCD2*, which codes for an enzyme with dioxygenase activity that cleaves carotenoids (Carotenoid Cleavage Dioxygenase: **CCD**). The overexpression of this gene in an engineered microorganism that accumulates zeaxanthin, such as *E. coli* or in a plant tissue rich in zeaxanthin, such as corn endosperm, determines the production of crocetin dialdehyde and of crocetin, respectively. The invention also relates to the development of an *in vitro* assay to study the activity of said dioxygenase.

Crocus sativus L. (*Iridaceae*) is a perennial, triploid, sterile plant propagated vegetatively and cultivated in mild areas from Spain to Kashmir (1). Its dried stigmas constitute the spice saffron, which is commonly considered to be the most costly in the world with prices ranging from 2,000 to 7.000 0 €/Kg. The high costs of saffron are caused by the necessary manual work associated with the harvesting of the stigmas: 1 Kg of dried saffron requires the manual harvesting of approximately 150,000 flowers (<http://www.europeansaffron.eu/>) (1). The stigmas of saffron accumulate a series of apocarotenoid compounds: crocetin and the glycosylated forms thereof (crocins), which are responsible for the red colouration of the stigmas; picrocrocin, which is responsible for the bitter taste of saffron; and safranal, which is responsible for the pungent aroma of saffron (2). In addition to being high-value ingredients added to food, the apocarotenoid compounds of saffron have cosmetic and medical applications (1), including, by way of non-limiting example, the protective function against degenerative diseases of the retina demonstrated for crocetin and crocins (3, 4). Under consideration of such applications and under consideration of the high cost of saffron apocarotenoids, which are currently produced by means of costly extraction processes, there is an industrial interest in the development of methods for the economical biotechnological production thereof by means of expression of the enzymes

involved in the biosynthesis thereof in microbial systems or in plants with high productivity. The biosynthesis of such compounds starts with the symmetrical cleaving, in positions 7,8 and 7'8', of zeaxanthin by means of a dioxygenase (CCD, Carotenoid Cleavage Dioxygenase) (Fig. 1). The two products of the cleavage of zeaxanthin, i.e. 3-OH- β -cyclocitral and crocetin dialdehyde, are dehydrogenated and glycosylated in order to produce, respectively, picrocrocin and crocins. The glucosyltransferases responsible for the glycosylation reaction have been described in saffron and *Gardenia* (5) (6). With regard to the first biosynthetic step, some authors have reported that an enzyme called ZCD cleaves zeaxanthin symmetrically at positions 7,8/7',8', producing crocetin dialdehyde (7), which is an observation contradicted by subsequent works (8). The modelling of the three-dimensional structure of ZCD has shown that ZCD is an incomplete enzyme, devoid of blade VII of the conserved β -propeller structure (9), consistent with the absence of enzymatic activity of ZCD (8). Consequently, it is not currently possible by biotechnological methods to produce the apocarotenoid compounds of saffron, and in particular crocetin and crocins, since the identification of the first enzyme of the biosynthetic pathway, i.e. CCD, which is responsible for the symmetrical cleaving in positions 7,8 and 7',8' of zeaxanthin, is still uncertain.

The object of the present invention is to provide a new carotenoid dioxygenase and methods for the biological production in microorganisms and plants of compounds of high added value derived from saffron.

Summary

Also disclosed is a nucleotide sequence of SEQ ID No. 1 coding for a carotenoid dioxygenase enzyme able to cleave carotenoids symmetrically in positions 7,8 and 7',8' with the production of crocetin dialdehyde.

Also disclosed is an enzyme coded from the nucleotide sequence as described herein, having the amino acid sequence of SEQ ID No. 2.

In one aspect, there is provided a method for the production of crocetin dialdehyde and/or crocetin in microorganisms that accumulate zeaxanthin, comprising a step of expression in said microorganisms of the enzyme coded by the nucleotide sequence of SEQ ID No. 1.

In another aspect, there is provided a method for the production of crocetin in vegetable cells that express zeaxanthin, comprising a step of expression in said cells of the enzyme coded by the nucleotide sequence of SEQ ID No. 1.

Also disclosed is use of an enzyme coded by the nucleotide sequence of SEQ ID No. 1 or of an enzyme having SEQ ID No. 2 to cleave symmetrically the

molecule of zeaxanthin in positions 7,8 and 7,'8' with the production of crocetin dialdehyde and/or crocetin.

Description

The above-mentioned and further objects of the invention, as will become clear from the following description, are achieved by means of a gene coding for a Carotenoid Dioxygenase enzyme, an enzyme coded in this way, and methods for producing crocin precursors as claimed in the corresponding independent claims.

The present invention will be better understood from the following detailed description of some preferred embodiments, provided with reference to the accompanying drawings, in which the corresponding figures show:

Fig. 1 Biosynthetic pathway of apocarotenoids in saffron. Proposed biosynthetic pathway for the biosynthesis of the apocarotenoids. Zeaxanthin is cleaved in positions 7,8 and 7',8' by the activity of a CCD. The cleavage product C₂₀, crocetin dialdehyde, is converted into crocetin from an aldehyde dehydrogenase, then into crocin by the activity of at least two UDPG glycosyltransferases. The product C₁₀, 3-OH- β -cyclocitral, is converted into picrocrocin from a UDPG-glycosyltransferase, and then into safranal.

Fig. 2 Expression and structural characteristics of saffron CCDs. (A) Expression of various CCDs during the development of the stigma of saffron. The graph above, the “heatmap”, shows the kinetics of accumulation of the different apocarotenoids; (B) Phylogenetic relationships between the saffron CCDs (Cs), *Arabidopsis* (At), rice (Os), tomato (Sl), lettuce (Ls), clementine (Cc) and *Synechocystis* (Sy), evaluated by means of the Neighbor-Joining method; (C) Three-dimensional models of the ACO (Apocarotenoid Cleavage Oxygenase) of *Synechocystis* and of the CCD2, ZCD and CCD4a of *Crocus sativus*. The ZCD is missing blade VII of the β -propeller structure (shaded).

Fig. 3 CCD2 expressed in *E. coli* converts zeaxanthin with formation of crocetin dialdehyde. (A) Cells of *E. coli* that accumulate lycopene, β -carotene or zeaxanthin have been transformed with the empty vector pThio1-DAN (C-), or with the same vector in which the CCD2 or the ZCD was cloned. The discolouration of the zeaxanthin exclusively in the cells that express CCD2 should be noted. (B) HPLC-HRMS analyses of the apocarotenoids in cells that express zeaxanthin and CCD2 after incubation at 20°C for 16h. These produce crocetin dialdehyde, of which the chromatographic mass and mobility is identical to that of an authentic standard.

Fig. 4 Affinity assay *in vitro*. (A) Cleaving of zeaxanthin in the *in vitro* assay from extracts of *E. coli* expressing CCD2. (B) Substrates that are not cleaved *in vitro* from CCD2. (C) Substrates that are cleaved *in vitro* from CCD2 and position of the cleaving.

The present invention will now be described in detail with reference to the figures.

RNA isolated from stigmas of saffron was sequenced by means of RNA-Seq 454 Titanium technology as reported in the literature (10) (Fig. 2A). The evolutionary relationships between the gene families CCD1 and CCD4 of saffron, *Arabidopsis*, rice and *Synechocystis*, were studied using the Neighbor-Joining method (11) (Fig. 2B), whilst the CCD proteins were modelled using the web server RaptorX (12) (Fig 2C). The assays in bacteria were obtained by transforming particular strains of *E. coli* engineered to accumulate lycopene, β -carotene, and zeaxanthin, with the expression vectors pThio-DAN1-CCD2, pThio-DAN1-ZCD and pThio-DAN1 (empty vector, used as control) as indicated in the literature (13) (14) (15) (Fig. 3). The transformation of corn endosperm, in *Agrobacterium tumefaciens*-mediated manner, was obtained by cloning the coding sequence of CCD2 in the vector pBI121 (16) and by using protocols already present in the literature (17). The strain of *E. coli* BL21 (p-Gro7) was used to assay the activity *in vitro* of CCD2: the raw protein extracts were prepared and incubated with the appropriate substrates and extracts as described

before (14) (Fig. 4). All the analyses directed to the identification and quantification of the carotenoids and apocarotenoids were performed using a mass spectrometry platform LTQ-Orbitrap coupled with a high-pressure liquid chromatography (HPLC) system and with a diode array detector (DAD) (ThermoFisher Scientific); the chromatographic separation was obtained using a C₃₀ column with inverse phase, whereas for positive ionisation an APCI (Atmospheric Pressure Chemical Ionization) source was selected (18).

The present invention will be better illustrated with reference to the following practical examples, which are provided by way of non-limiting explanation.

Example 1. Identification of the gene *CCD2* expressed in the stigmas of *C. sativus*:

The transcripts *CCD4* and *ZCD* were expressed around anthesis, when the biosynthesis of crocetin, of picrocrocetin and of crocins is essentially complete. Only *CCD2* is expressed in the early stages of development, which actively synthesise apocarotenoids (Fig. 2A) (19). The *CCD2* protein of *Crocus* is a member of a new sub-family of CCDs, similar to, but distant from, the *CCD1*s (Fig. 2B). *CCD2* has all 7 blades forming the β -propeller structure, preserved in all CCDs, whilst *ZCD* is missing blade VII (Fig. 2C).

Example 2. Demonstration that the gene *CCD2*, expressed in the bacteria *E. coli* genetically modified to accumulate zeaxanthin, converts zeaxanthin with production of crocetin dialdehyde and that the gene *ZCD* does not possess such property. The coding sequences of the genes *CCD2* and *ZCD* were cloned in the vector pThio-DAN1, which allows an inducible expression in *E. coli* induced from arabinose. The recombinant proteins were expressed in three strains of *E. coli* that accumulate lycopene, β -carotene and zeaxanthin (Fig. 3A) (13) (14). Whilst no discolouration and no cleavage product were observed in the strains of *E. coli* that accumulate lycopene and β -carotene following the expression of *CCD2* or of *ZCD* (Fig. 3A), the expression of *CCD2* in the strain of *E. coli* that accumulates zeaxanthin lead to clear discolouration of the bacterial cells (Fig. 3A). The HPLC analyses coupled with high-resolution mass spectrometry (LC-HRMS) showed a peak with a m/z ratio of 297.1847, which corresponds to crocetin dialdehyde, which demonstrates the same chromatographic mobility as an authentic standard of such a substance (Fig. 3B). A peak with the chromatographic and the m/z ratio (329.1747) typical of crocetin was also discernible. The *ZCD*, expressed in the same system, did not produce any discolouration or any cleavage product (Fig. 3A-B).

Example 3. Demonstration that the gene *CCD2* expressed in corn endosperm and/or cells of *Arabidopsis thaliana* converts zeaxanthin with production of crocetin. The sequences coding *CCD2* and *ZCD* were cloned in the binary vector pBI121 under the control

of the established promoter 35S. The vectors thus obtained were used to transform corn endosperm and cells of *Arabidopsis thaliana*, which contain zeaxanthin. The endosperm transformed with CCD2 showed discolouration, which is a sign of a cleaving activity of carotenoids, compared with that transformed with a vector containing ZCD or a component of β -glucuronidase. The LC-HRMS analysis carried out on the vegetable cells that express CCD2 did not demonstrate the presence of 3-OH- β -apo-8'-carotenal or of crocetin dialdehyde, but a new peak with a ratio m/z equal to 329.1747, corresponding to crocetin, which demonstrates chromatographic mobility identical to that of an authentic standard of this substance. Such a conversion is caused by the conversion of crocetin dialdehyde to crocetin in the corn endosperm (Table 1) and in the cells of *Arabidopsis thaliana*. No cleavage product of zeaxanthin was instead produced in the corn endosperm expressing ZCD or β -glucuronidase.

Example 4. Development and use of an *in vitro* assay to study the specificity of CCD2 for the various carotenoids. Since the strains of genetically modified *E. coli* can produce only a limited number of different carotenoids, we developed and used an *in vitro* assay, based on semi-purified extracts of *E. coli* that expresses CCD2, in order to study the specificity of this enzyme for the various carotenoids. In such an assay the CCD2 does not cleave β -carotene, violaxanthin, or β -apo-8'-carotenal, but cleaves the zeaxanthin producing 3-OH- β -apo-8'-carotenal, i.e. the product of the single cleave in position 7',8' (Fig. 4). ZCD, instead, does not cleave any of the carotenoids, not even zeaxanthin. 3-OH- β -apo-8'-carotenal, formed from CCD2 on the basis of the zeaxanthin *in vitro*, is then cleaved from the CCD2, with production of crocetin, signalling that the extracts of *E. coli* are able to convert the product of CCD2, crocetin dialdehyde, into crocetin.

Table 1: Relative content of carotenoids and apocarotenoids in corn endosperm transformed with plasmids expressing various enzymes

Molecule	35S: β - glucuronidase	35S:CCD2	35S:ZCD
Lutein	0.81 ± 0.15	$0.48 \pm 0.05^{**}$	0.64 ± 0.09
Zeaxanthin	1.98 ± 0.33	$0.84 \pm 0.18^{**}$	1.73 ± 0.20
β -cryptoxanthin	0.15 ± 0.03	0.17 ± 0.04	0.12 ± 0.04

Crocetin	n.d.	0.008 ± 0.002	n.d.
----------	------	-------------------	------

Peak areas of the ions, standardised by the internal standard, for the more abundant carotenoids and apocarotenoids. The data shown represents the mean $\pm$ SD of 4 biological repetitions. n.d.=not detectable; *=P value 0.05; **=P value 0.01.

References

- (1) Fernandez JA & Pandalai SG (2004) Biology, biotechnology and biomedicine of saffron. *Recent research developments in plant science*. Vol. 2:127-159.
- (2) Caballero-Ortega H, Pereda-Miranda R, & Abdullaev FI (2007) HPLC quantification of major active components from 11 different saffron (*Crocus sativus* L.) sources. *Food Chemistry* 100(3):1126-1131.
- (3) Yamauchi M, Tsuruma K, Imai S, Nakanishi T, Umigai N, Shimazawa M, Hara H: Crocetin prevents retinal degeneration induced by oxidative and endoplasmic reticulum stresses via inhibition of caspase activity. *European journal of pharmacology* 2011 650(1):110-119.
- (4) Qi Y, Chen L, Zhang L, Liu WB, Chen XY, Yang XG: Crocin prevents retinal ischaemia/reperfusion injury-induced apoptosis in retinal ganglion cells through the pi3k/akt signalling pathway. *Experimental eye research* 2013 107(44-51).
- (5) Moraga AR, Nohales PF, Perez JA, & Gomez-Gomez L (2004) Glucosylation of the saffron apocarotenoid crocetin by a glucosyltransferase isolated from *Crocus sativus* stigmas. *Planta* 219(6):955-966.
- (6) Nagatoshi M, *et al.* (2012) UGT75L6 and UGT94E5 mediate sequential glucosylation of crocetin to crocin in *Gardenia jasminoides*. *FEBS Lett* 586(7):1055-1061.
- (7) Bouvier F, Suire C, Mutterer J, & Camara B (2003) Oxidative remodeling of chromoplast carotenoids: Identification of the carotenoid dioxygenase CsCCD and CsZCD genes Involved in *Crocus* secondary metabolite biogenesis. *Plant Cell* 15(1):47-62.
- (8) Rubio A, *et al.* (2008) Cytosolic and plastoglobule-targeted carotenoid dioxygenases from *Crocus sativus* are both involved in beta-ionone release. *J Biol Chem* 283(36):24816-24825.
- (9) Sui X, Kiser PD, Lintig J, & Palczewski K (2013) Structural basis of carotenoid cleavage: from bacteria to mammals. *Arch Biochem Biophys* 539(2):203-213.
- (10) Alagna F, *et al.* (2009) Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *BMC Genomics* 10(1):399.
- (11) Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4):406-425.
- (12) Kallberg M, *et al.* (2012) Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* 7(8):1511-1522.

- (13) Prado-Cabrero A, Scherzinger D, Avalos J, & Al-Babili S (2007) Retinal biosynthesis in fungi: characterization of the carotenoid oxygenase CarX from *Fusarium fujikuroi*. *Eukaryot Cell* 6(4):650-657
- (14) Alder A, Holdermann I, Beyer P, & Al-Babili S (2008) Carotenoid oxygenases involved in plant branching catalyse a highly specific conserved apocarotenoid cleavage reaction. *Biochem J* 416(2):289-296.
- (15) Trautmann D, Beyer P, & Al-Babili S (2013) The ORF slr0091 of *Synechocystis* sp. PCC6803 encodes a high-light induced aldehyde dehydrogenase converting apocarotenals and alkanals. *FEBS J* 280(15):3685-3696.
- (16) Jefferson RA (1987) Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol Biol Reporter* 5:387-405.
- (17) Reyes FC, Sun B, Guo H, Gruis DF, & Otegui MS (2010) *Agrobacterium tumefaciens*-mediated transformation of maize endosperm as a tool to study endosperm cell biology. *Plant Physiol* 153(2):624-631
- (18) Fantini E, Falcone G, Frusciante S, Giliberto L, & Giuliano G (2013) Dissection of Tomato Lycopene Biosynthesis through Virus-Induced Gene Silencing. *Plant Physiol* 163(2):986-998.
- (19) Moraga AR, Rambla JL, Ahrazem O, Granell A, & Gomez-Gomez L (2009) Metabolite and target transcript analyses during *Crocus sativus* stigma development. *Phytochemistry* 70(8):1009-1016.

CLAIMS

1. A method for the production of crocetin dialdehyde and/or crocetin in microorganisms that accumulate zeaxanthin, comprising a step of expression in said microorganisms of the enzyme coded by the nucleotide sequence of SEQ ID No. 1.
2. The method according to Claim 1, wherein said microorganisms are bacterial cells of E.coli.
3. A method for the production of crocetin in vegetable cells that express zeaxanthin, comprising a step of expression in said cells of the enzyme coded by the nucleotide sequence of SEQ ID No. 1.
4. The method according to Claim 3, in which said step of expression is carried out by transforming said cells with the binary vector pBI121, said vector comprising said nucleotide sequence.