Novel carotenoid cleavage dioxygenase catalyzes the first dedicated step in saffron crocin biosynthesis

Sarah Frusciante, Gianfranco Dritto, Mark Bruno, Paola Ferrante, Marco Pietrella, Alfonso Prado-Cabreiro, Angela Rubio-Moraga, Peter Beyer, Lourdes Gomez-Gomez, Salim Al-Babili, and Giovanni Giuliano

*Italian National Agency for New Technologies, Energy, and Sustainable Development, Casaccia Research Centre, 00123 Rome, Italy; †Sapienza, University of Rome, 00185 Rome, Italy; ‡Faculty of Biology, University of Freiburg, D-79104 Freiburg, Germany; §Center for Desert Agriculture, Division of Biological and Environmental Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia; and ††Instituto Botánico, Facultad de Farmacia, Universidad de Castilla–La Mancha, 02071 Albacete, Spain

Edited by Rodney B. Croteau, Washington State University, Pullman, WA, and approved July 3, 2014 (received for review March 16, 2014)

Crocus sativus stigmas are the source of the saffron spice and accumulate the apocarotenoids crocetin, crocins, picrocrocin, and safranal, responsible for its color, taste, and aroma. Through deep transcriptome sequencing, we identified a novel dioxygenase, carotenoid cleavage dioxygenase 2 (CCD2), expressed early during stigma development and closely related to, but distinct from, the CCD1 dioxygenase family. CCD2 is the only identified member of a novel CCD clade, presents the structural features of a bona fide CCD, and is able to cleave zeaxanthin, the presumed precursor of saffron apocarotenoids, both in Escherichia coli and in maize endosperm. The cleavage products, identified through high-resolution mass spectrometry and comigration with authentic standards, are crocetin dialdehyde and crocetin, respectively. In vitro assays show that CCD2 cleaves sequentially the 7,8 and 7′,8′ double bonds adjacent to a 3-OH-β-ionone ring and that the conversion of zeaxanthin to crocetin dialdehyde proceeds via the C30 intermediate 3-OH-β-apo-8′-carotenal. In contrast, zeaxanthin cleavage dioxygenase (ZCD), an enzyme previously claimed to mediate crocetin formation, did not cleave zeaxanthin or 3-OH-β-apo-8′-carotenal in the test systems used. Sequence comparison and structure prediction suggest that ZCD is an N-truncated CCD4 form, lacking one blade of the β-propeller structure conserved in all CCDs. These results constitute strong evidence that CCD2 catalyzes the first dedicated step in crocin biosynthesis. Similar to CCD1, CCD2 has a cyttoplasmic localization, suggesting that it may cleave carotenoids localized in the chromoplast outer envelope.

Significance

Saffron is a triploid, sterile species whose red stigmas constitute the most expensive spice on Earth. The color, the taste, and the aroma of the spice are owed to the crocus-specific apocarotenoid accumulation of crocetin/crocins, picrocrocin, and safranal. Through deep transcriptome analysis, we identified a novel carotenoid cleavage dioxygenase (CCD) whose expression profile parallels the production of crocetin. Using in bacteria, in vitro, and in planta functional assays, we demonstrate that CCD2 is the dioxygenase catalyzing the first dedicated step in saffron crocin biosynthesis starting from the carotenoid zeaxanthin.

Author contributions: S.A.-B. and G.G. designed research; S.F., G.D., and P.F. performed research; P.B., L.G.-G., M.B., M.P., A.P.-C., and A.R.M. contributed new reagents/analytic tools; S.F. and G.D. analyzed data; and G.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Ferrocyanide available through the PNAS open access option.

Data deposition: The sequence of CCD2 reported in this paper has been deposited in the GenBank database (accession no. KJ541749).

1To whom correspondence should be addressed. Email: giovanni.giuliano@enea.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1404629111/-/DCSupplemental.

the plant Crocus sativus L. (Iridaceae) is a perennial, sterile, vegetatively propagated triploid widely cultivated in a temperate belt extending from Spain to Kashmir (1). Albeit its site of domestication is uncertain, the earliest archaeological evidence of its cultivation is provided by Minoan frescoes dated 1,700–1,500 B.C. Its dried red stigmas (Fig. 1A) constitute the saffron spice, which is commonly considered the most expensive spice on Earth, with retail prices ranging between 2,000 and 7,000 €/kg. These high prices are due to the labor associated with its harvesting: because one stigma of saffron weighs about 2 mg, 1 kg of dry saffron requires the manual harvest of stigmas from around 110,000–170,000 flowers (www.europeansaffron.eu) (1).

Saffron stigmas accumulate large amounts (up to 8% on dry weight) of the apocarotenoids crocetin (and its glycosylated forms, crocins), responsible for the red pigmentation of the stigmas; picrocrocin, responsible for their bitter flavor; and safranal, responsible for the pungent aroma of saffron (Fig. 1A) (2). The proposed biosynthetic pathway (3, 4) starts through the symmetric cleavage of zeaxanthin at the 7,8′/7′,8′ positions by a nonheme iron carotenoid cleavage dioxygenase (CCD) (Fig. 1B). The two cleavage products, 3-OH-β-cyclolact and crocin dialdehyde, are dehydrogenated and glycosylated to yield picrocrocin and crocins, respectively. Putative glucosyl transferases responsible for the synthesis of crocins have been characterized in saffron and in Gardenia (5, 6).

Plant CCDs can be classified in five subfamilies according to the cleavage position and/or their substrate preference: CCD1, CCD4, CCD7, CCD8, and nine-cis-epoxy-carotenoid dioxygenases (NCEDs) (7–9). NCEDs solely cleave the 11,12 double bond of 9-cis-epoxy-carotenoids to produce the ABA precursor xanthotoxin. CCD7 and CCD8 act sequentially in the strigolactone pathway, leading to strigolactone precursor carlactone (10). Enzymes of the CCD1 family cleave a wide spectrum of different carotenoids at several different positions (9,10; 9,10,9′,10′; 5,6,5′,6′; or 7,8′,7′) (11, 12). CCD4 enzymes cleave carotenoids at the 9′,10′ or the 7′,8′ positions and determine the level of pigmentation in plant tissues, including Chrysanthemum petals (13), peach flesh (14), potato tubers (15), Citrus peel (16, 17), and Arabidopsis seeds (18).

Structurally, all CCDs are characterized by a rigid, seven-bladed β-propeller structure, at the axis of which a Fe²⁺ atom is located (19). The propeller is covered by a less-conserved dome formed by a series of loops. The reaction is catalyzed by the Fe²⁺ atom via the introduction of oxygen (20).

To date, conflicting data have been reported about the identity of the enzyme catalyzing the cleavage reaction in saffron. A zeaxanthin cleavage dioxygenase (ZCD) was reported to cleave zeaxanthin symmetrically at the 7′,8′/7′,8′ positions, yielding the crocin precursor crocetin dialdehyde (4). However, later work has suggested that ZCD is a truncated form of a plastoglobule-localized CCD4 enzyme, devoid of cleavage activity, and that the database deposited in the GenBank database (accession no. KJ541749).

To whom correspondence should be addressed. Email: giovanni.giuliano@enea.it.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1404629111/-/DCSupplemental.
Approximately 120,000 454 reads from each stage were assembled and their biosynthesis is essentially complete at the R stage (22). Dark red stigmas in closed bud outside the perianth tubes; 0dA, perianth tubes (0.8 mm in length); (around 0.4 mm in length); R, red stigma, closed bud inside the stigma, closed bud inside the perianth tubes (around 0.3 cm in length). Sequencing of six different stigma developmental stages: Y, yellow; C, closed; Z, 3-OH-β-cyclocitrinal; 8; Safranal.

Identification of CCD transcripts expressed during saffron stigma development. CCD1, responsible for the cleavage step leading to crocetin biosynthesis, is a novel type of plant CCD. We report that CCD2 is the enzyme responsible for the cleavage step leading to crocetin biosynthesis from zeaxanthin. In particular, CCD2 expression peaks early, at the O stage (Fig. 2A) coincident with crocetin and crocin accumulation (22), whereas ZCD and CCD4 are expressed late during stigma development.

A phylogenetic analysis of CCD protein sequences from several plants was inferred using the neighbor-joining method using Synecochysist apocarotenoid cleavage oxygenase (ACO) as an outgroup (Fig. 2B). The results suggested that Crocus CCD2 is a member of a clade closely related to, but distinct from, angiosperm CCD1 enzymes. A lettuce enzyme labeled as CCD2 (23) clustered with CCD1 enzymes, whereas an enzyme known to cut zeaxanthin at the 7,8 position, Citrus CCD4b1 (16), clustered with CCD4 enzymes (Fig. 2B). ZCD appeared to be a member of the CCD4 family (Fig. 2B), truncated at the N terminus (Fig. S1).

Because the ZCD cdna was originally isolated by rapid amplification of cDNA ends (RACE) (4, 24) that can lead to the cloning of truncated transcripts, we carried out a 5′-RACE analysis of CCD4 transcripts. Next to a 950-base full-length transcript, whose length is compatible with a full-length CCD4 protein, a series of abundant 5′-truncated transcripts are detectable, the longest of which is compatible with the length of the ZCD protein, which is encoded starting from an internal ATG codon (Fig. S2). It is therefore likely that the original ZCD clone (4) corresponds to a truncated CCD4 transcript. This cannot be either CCD4a or CCD4b (25), which are only 98% identical to ZCD at the nucleotide level. To further address this point, we cloned the 400- to 350-bp RACE products shown in Fig. S2A, containing the internal ATG codon, and sequenced multiple clones. The sequence of eight of the clones corresponds to CCD4a, of one to CCD4b, of five to ZCD (4), and of four to a yet-unidentified CCD4. All of them contain the internal ATG codon.

We modeled the CCD2, ZCD, and CCD4a structures using the RaptorX web server (26) based on the known crystal structure of the Synecochysist ACO (20) as a reference (Fig. 2C). The deduced models show that ZCD is an incomplete enzyme in comparison with the other CCDs predicted structures and ACO. In particular, it lacks blade VII of the β-propeller and part of the dome, whereas CCD2 displays all of the structural features of bona fide CCDs (Fig. 2C).

Saffron CCD2 Expressed in Escherichia coli Cleaves Zeaxanthin to Yield Crocetin Dialdehyde. The CCD2 and ZCD coding sequences were cloned to yield thioredoxin fusion proteins in the pThio-DAN1 vector, allowing arabinose-inducible expression in E. coli (27). The recombinant proteins were expressed in genetically engineered E. coli strains, accumulating lycopene, β-carotene, and zeaxanthin, respectively (Fig. 3A) (28). SDS/PAGE analysis showed that both CCD2 and ZCD fusions were expressed with an apparent molecular mass of 81 and 59 kDa, respectively (Fig. S2B).

No decoloration was observed in E. coli strains accumulating lycopene or β-carotene upon expression of CCD2 or ZCD (Fig. 3A), and no cleavage product was detected in these strains (Fig. S3). In contrast, CCD2 expression in zeaxanthin accumulating E. coli cells led to evident decoloration (Fig. 3A). Analysis by HPLC coupled with high-resolution mass spectrometry (LC-HRMS) revealed a new peak with an m/z of 297,1847 corresponding to the protonated C20-dialdehyde, 8.8′-diapocarotene-8.8′-dial (crocetin dialdehyde) that coeluted with the authentic standard (Fig. 3B). We therefore concluded that CCD2 cleaves zeaxanthin symmetrically at the 7,8,7′,8′ positions to yield crocetin dialdehyde.
ZCD showed no decoloration and no detectable cleavage products in any of the strains (Fig. 3 and S3).

Saffron CCD2 Expressed in Maize Endosperm Cleaves Zeaxanthin to Yield Crocetin. To investigate the cleavage activity of CCD2 and ZCD in planta, we used Agrobacterium-mediated transient expression (29) in yellow maize endosperm, which is known to accumulate several xanthophylls, including zeaxanthin. The CCD2 and ZCD coding sequences were cloned into a binary plasmid or vector under the control of 35S promoter. A vector containing the intron-bearing β-glucuronidase reporter gene (p35S: GUS_INT:NOS) (30) was used to optimize the transformation protocol (Table S2) and as a control for transformation efficiency. Fig. 4 shows pictures of maize kernels transformed with the three constructs. Kernels transformed with CCD2 show decoloration, compared with those transformed with the control plasmid or ZCD. Analysis of the CCD2-expressing samples by quantitative LC-HRMS (Fig. 4B) showed neither the cleavage intermediate 3-OH-β-apo-8′-carotenal (β-citrin) nor the final product crocetin dialdehyde (Fig. 4B). However, we identified a new peak with an m/z of 329.1747 expected for crocetin that was chromatographically indistinguishable from an authentic crocetin standard. Thus, contrary to E. coli, maize endosperm most likely possesses an endogenous aldehyde dehydrogenase, allowing this crocetin dialdehyde oxidation step. This product was not detectable in endosperm overexpressing ZCD or GUS_INT (Fig. 4B).

LC-HRMS of carotenoids of transformed maize kernels revealed significant decreases in the content of both zeaxanthin and lutein, but not in that of β-cryptoxanthin, indicating that also lutein may be a CCD2 substrate (Table 1).

In Vitro Substrate Specificity of Saffron CCD2. Because only a limited number of carotenoids can be produced in genetically engineered E. coli, we used an in vitro assay to explore the substrate specificities and regional cleavage specificities of CCD2 and ZCD (Fig. 5). In the in vitro assay, CCD2 did not convert β-carotene, violaxanthin, β-apo-8′-carotenal, or β-cryptoxanthin (Fig. S4), but it cleaved zeaxanthin yielding a C30 apocarotenoid identified on the basis of its m/z and its chromatographic identity with the authentic standard, as 3-OH-β-apo-8′-carotenal (β-citrin) (Fig. S5 A and B), i.e., the product of a single cleavage at the 7,8′ position. ZCD did not convert any of the substrates tested, including zeaxanthin (Fig. S5A).

We also tested whether the product 3-OH-β-apo-8′-carotenal, formed by CCD2 from zeaxanthin in vitro, can act as well as a substrate. Indeed, the formation of crocetin dialdehyde and crocetin was observed (Fig. S5C). This suggests that 3-OH-β-apo-8′-carotenol is a substrate of CCD2 and that the conversion of zeaxanthin to crocetin dialdehyde likely occurs in two sequential steps. Furthermore, E. coli seems to contain an aldehyde dehydrogenase activity that is not active in vivo, but partially activated in vitro.

CCD2 cleaved also lutein (Fig. S6 A and B), yielding a C30 apocarotenoid with a chromatographic mobility different from that of 3-OH-β-apo-8′-carotenol. Despite the unavailability of an authentic standard, this compound could be identified unambiguously as 3-OH-ε-apo-8′-carotenol on the basis of its m/z of 415.2981, indicative of the loss of a water molecule (432.3028 + [H+] − [H2O]). Molecules that have an OH group at an allylic position, such as the 3 position of an ε-ionone ring, readily eliminate one molecule of water upon ionization (31). The above results suggest that the CCD2 cleavage site is always at the 7,8′ position adjacent to the 3-OH-β-ionone ring (Fig. 5). CCD2 cleaved also 3-OH-β-apo-10′-carotenol (C27) and 3-OH-β-apo-12′-carotenol (C25) (Fig. 5 and Fig. S6 C and D), yielding a C17′ and a C15′-dialdehyde, respectively. This indicates that CCD2 is regiospecific, always targeting the C7-C8 double bond and tolerating variations in the length of the apocarotenoid polyene moiety.

Fig. 2. Expression and structural characteristics of carotenoid cleavage dioxygenases from saffron stigma. Transcript levels of saffron CCDs in different stigma developmental stages, based on 454 RNA-Seq data; −2dA, 2 d preanthesis; 0dA, day of anthesis; +2dA, 2 d postanthesis (A). Data expressed as reads per kilobase per million (RPKM). The graph above the heat map indicates the kinetics of accumulation of the different apocarotenoids. Phylogenetic relationships of CCDs from saffron (C), Arabidopsis (At), rice (Os), tomato (Sl), lettuce (Ls), clementine (Cc), and Synechocystis (Sy) inferred using the neighbor-joining method; GCD1, CAC79592.1; GCD4a, ACDE62476.1; GCD4b, ACDE62477.1; GCD2, CAD33262.1; ACOCD1, AT3G63520; ACOCD4, AT4G19170; OsCD1, Os12g0640600; OsCCD4a, Os02g0704000; OsCCD4b, Os12g0435200; SICCD1a, Soly01g087250.2; SICCD4a, Soly08g07480.2; LsCCD2, BAE72095.1; CcCCD4b1, Cicle10028113m; SyACO, P74334 (B). Topology diagrams of Synechocystis apocarotenoid cleavage oxygenase (ACO) and Crocus sativus CCD2, ZCD, CCD4a (C). Secondary structural elements consisting of α-helices and β-sheets are colored in pink and yellow, respectively. The seven blades are labeled from I to VII for ACO and is the same for the other topology diagrams. The ferrous catalytic iron is colored in green. All structural elements located outside the seven blades form part of the dome. The gray shaded structural elements in ZCD are lacking; please note the alternative N terminus. Most of the dome is lacking in this protein, together with most of blade VII. CCD4a topology diagram is shown for comparison.
To assess the affinity of CCD2 for its different substrates, we measured the percentage conversion rates of these substrates in the in vitro assay (Table S3). Although the data are only semiquantitative, due to the differential solubility of the different substrates, 3-OH-β-apo-8′-carotenal showed the highest (52.7%) conversion rate, followed by 3-OH-β-apo-12′- and 3-OH-β-apo-10′-carotenal (18.5% and 12.5%, respectively). Zeaxanthin and lutein showed the lowest (4.8% and 1.7%, respectively) conversion rates among the cleaved substrates.

**Subcellular Localization of Saffron CCD2.** Based on ChloroP analysis, CsCCD2 lacks a recognizable plastid transit peptide (Fig. S1). Because as many as 12% of chloroplast-localized proteins present this feature (32), we studied the localization of a C-terminal fusion of CCD2 to green fluorescent protein (CCD2:GFP) in Nicotiana benthamiana-agroinfiltrated leaves. The results (Fig. S7) suggest that CCD2:GFP is a cytoplasmic protein.

**Discussion**

Using deep transcriptome analysis of developing saffron stigmas, we have identified a novel CCD enzyme, CCD2, expressed during early stigma development, consistent with the time course of crocetin formation. Analysis of the amino acid sequences of several CCDs belonging to saffron, *Arabidopsis*, lettuce, *Citrus*, rice, and cyanobacteria indicates that saffron CCD2 represents a novel branch close to, but distinct from the CCD1 family. CCD1 enzymes are known to cleave carotenoids, linear and cyclic, at several bonds (9,10; 9,10,9,10′; 5,6,5′,6′; or 7,8,7′,8′) (33).

On the basis of the evidence presented, we suggest that CCD2 is the enzyme that catalyzes the zeaxanthin cleavage step in crocetin biosynthesis. The previously described ZCD enzyme (4) appears to be an N-truncated form of a CCD4 enzyme, encoded by a 5′-truncated *CCD4* transcript distinct from both *CCD4a* and *CCD4b* (25). This truncated enzyme was inactive in all of our in vivo and in vitro assays.

In contrast to ZCD, CCD2 displays all of the structural features of a bona fide CCD. It is highly expressed at the orange stage of stigma development, when crocetin accumulation is maximal, and when expressed in *E. coli*, it is able to convert zeaxanthin to crocetin dialdehyde via two sequential cleavage reactions at the 7,8 and 7′,8′ positions. In vivo expression in maize kernels and in vitro assays confirm this activity and provide evidence for the subsequent conversion of crocetin dialdehyde to crocetin, probably through the action of nonspecific maize aldehyde dehydrogenases.

Like the related CCD1 enzymes (21, 34), CCD2 lacks a recognizable plastid transit peptide and is localized to the cytoplasm.
Carotenoids are synthesized in plastids and are found in several plastid compartments, including the outer envelope, which is particularly rich in xanthophylls (35). Therefore, a likely hypothesis is that CCD2 transiently associates with the outer envelope of saffron stigma chromoplasts and cleaves the xanthophylls localized in it.

The first CCD 3D structure was obtained from ACO, a cyanobacterial enzyme synthesizing the C_{20} apocarotenoid retinal (20). The protein structure revealed that the enzyme contains a Fe^{2+} ion in the active site, coordinated by four conserved histidine residues, an arrangement common to all CCDs. The iron in ACO is encased by a rigid, seven-bladed β-propeller structure, overarched by a dome of six large loops (Fig. 2C). The β-propeller portion of the structure is present in all CCDs characterized to date, from bacteria to animals to plants (19). To understand the differences between the here-identified CCD2 and ZCD, we modeled the tertiary structure using ACO as template (Fig. 2C). This revealed that ZCD lacks blade VII of the propeller [known to participate in the coordination of the central iron atom (19)] and part of the dome. 5′-RACE experiments (Fig. S24), suggest that ZCD is encoded by a truncated CCD4 transcript, leading to a nonfunctional protein. All assays aimed at uncovering a cleavage activity of ZCD, in bacterio, in planta, and in vitro, were met with negative results. This is consistent with the observation made by Rubio et al. (21) with the CCD4a-211 truncated enzyme, which is almost identical to ZCD.

Our results confirm the pathway proposed for saffron apocarotenoid biosynthesis (3, 4) in two important aspects: zeaxanthin is a substrate for the cleavage reaction and the cleavage occurs at the 7,8 and 7′,8′ positions. The in vitro assays indicate that the cleavage reaction occurs in two subsequent steps: a first cleavage generates 3-OH-β-apo-8′-carotenal, which is then reclaved by the same enzyme to yield crocetin dialdehyde. We envisage two possible mechanisms through which this double cleavage can occur: (i) a sliding mechanism, in which the carotenoid molecule is bound in the hydrophobic tunnel in a position that brings the catalytic iron close to the 7,8 double bond, and then, after the first cleavage has occurred, it slides to bring the iron close to the 7′,8′ double bond for the second cleavage; or (ii) a flipping mechanism, in which after the first cleavage the apocarotenoid exits the tunnel and reenters it in the opposite orientation, to be cleaved at the symmetric position. We favor the second mechanism, in view of the fact that 3-OH-β-apo-8′-carotenal is accumulated in free form in the in vitro reaction.

The combined data obtained in bacterio, in planta, and in vitro give a rather precise idea of the steric requirements of CCD2 for its substrate: CCD2 cleaves zeaxanthin, lutein, and all tested 3-OH-β-apocarotenoids at the 7,8 position, but it does not cleave β-carotene and lycopene, indicating an absolute requirement for 3-OH-β-ring at the proximal end of the molecule. The distal end of the molecule can be a 3′-OH-β- or ε-ring or an aldehyde moiety because zeaxanthin, lutein, and 3-OH-β-apocarotenals of varying lengths are accepted substrates. However, the in planta and in vitro data suggest that some constraints exist also for the distal end, because β-cryptoxanthin, which has an unsubstituted β-ring at the distal end, is not cleaved by CCD2. We measured the percentage conversion of the various substrates in the in vitro assay. The results indicate that zeaxanthin and 3-OH-β-apo-8′-carotenal are the preferred substrates, respectively, for the first and second cleavage reaction. The very high conversion of 3-OH-β-apo-8′-carotenol provides an explanation for the fact that this intermediate is not accumulated in bacterio or in planta.

Although the principal objective of this study, i.e., the identification of the enzyme catalyzing the initial cleavage step in the saffron apocarotenoid pathway and the characterization of its activity, has been met, the enzymatic steps downstream of the cleavage step still await complete elucidation. Several aldehyde dehydrogenases and glucosyl transferases have been identified in our transcriptome data and hold promise for a complete reconstruction of the saffron apocarotenoid pathway.

Materials and Methods

454 Titanium RNA-Seq sequencing of saffron stigma DNAs was performed according to published methods (36) and will be reported elsewhere. The sequence of CCD2 has been submitted to GenBank under accession number KJ541749. Evolutionary relationships were inferred using the neighbor-joining method (37), and evolutionary analyses were conducted in MEGA8 (38). Heat maps were created using Genesis (39). CCD models were drawn with the RaptorX web server (26). Chloroplast transit peptides were deduced using ChloroP (40). 5′-RACE was performed using a commercial kit (Life Technologies; catalog number 18374-058). In bacterio assays were performed using E. coli strains accumulating lycopene, β-carotene, zeaxanthin (28, 41), and CCD2 or ZCD expressed in the pTHIO-DAN1 expression vector (27). For in vitro assays, the expression vectors were transformed into E. coli BL21 (pGro7) (Takara); crude lysates were prepared, incubated with appropriate substrates, and extracted as described (41). Carotenoid/apocarotenoid analysis was performed on an LTQ-Orbitrap mass spectrometry system coupled to an Accela U-HPLC system equipped with a photodiode array detector (Thermofisher Scientific) using positive mode atmospheric pressure chemical ionization and a C18 reverse-phase column (31). Ion peak areas were normalized to the internal standard (α-tocopherol acetate). For transient transformation of maize kernels, CCD2 was cloned in the pBl121 vector (42) and transformed using a published method (29). For subcellular localization, CCD2 was fused C-terminally to enhanced green fluorescent protein (eGFP) (43) using Gibson assembly (44) and agroinfiltrated in Nicotiana benthamianna leaves as described (45). After 48 h, leaves were analyzed by confocal laser-scanning microscopy. Green and red fluorescence were used to detect eGFP and chlorophyll signals, respectively. A detailed description of all materials and methods used is provided in SI Materials and Methods.

Table 1. Normalized ion peak areas

<table>
<thead>
<tr>
<th>Compound</th>
<th>Control</th>
<th>ZCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lutein</td>
<td>0.81 ± 0.15</td>
<td>0.48 ± 0.05</td>
</tr>
<tr>
<td>Zeaxanthin</td>
<td>1.98 ± 0.33</td>
<td>0.84 ± 0.18</td>
</tr>
<tr>
<td>β-Cryptoxanthin</td>
<td>0.15 ± 0.03</td>
<td>0.17 ± 0.04</td>
</tr>
<tr>
<td>Crocetin</td>
<td>n.d.</td>
<td>0.008 ± 0.002</td>
</tr>
</tbody>
</table>

Ion peak areas, normalized for the internal standard, for the main kernel carotenoids and apocarotenoids. Data are the average ± SD of four biological replicates. n.d., not detectable; *P value, 0.01.
ACKNOWLEDGMENTS. We thank Hanseog Ernst for providing the synthetic substrates, Nara Lico for a gift of N. benthamiana plants and for help in agroinfiltration experiments; Elena Romano and Emanuela Viaggii at the Centre of Advanced Microscopy “Patrizia Albertano” for the confocal images; Gaetano Perrotta, Paolo Facella, and Fabrizio Carbone for 454 sequencing; and Alessia Fiore for critical reading of the manuscript. This work was supported by the Italian Ministry of Research (Project “Integrated Knowledge for the Sustainability and Innovation of Italian Agri-Food”), German Research Foundation (Deutsche Forschungsgemeinschaft) Grant AL 892/1-4, the European Union [The development of tools and effective strategies for the optimisation of useful secondary metabolite production in plants, Seventh Framework Programme (FP7) Contract 244348], and the Iberoamerican Network for the Study of Carotenoids as Food Ingredients (112RTO445). S.F. was supported by short-term fellowships of the PlantEngine (FA1006) and Saffronomics (FA1101) European Cooperation in Science and Technology actions.

The evolutionary relationships of carotenoid cleavage dioxygenase 1 (CCD1) and CCD4 family genes from saffron, *Arabidopsis*, rice, and *Synecochystis* were inferred using the neighbor-joining method (3) with a bootstrap value of 500. Evolutionary analyses were conducted in MEGA5 (4). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Heat maps were created using Genesis (5). Models for CCD2, zeaxanthin cleavage dioxygenase (ZCD), and CCD4a were deduced using the crystal structure of *Synecochystis* apocarotenoid cleavage oxygenase. Molecular graphics and analyses were performed with the RaptorX web server (6). Chloroplast transit peptides were deduced using ChloroP (7).

### 5′-RACE.

5′-RACE was performed using a commercial kit (Life Technologies; catalog number 18374-058) according to the manufacturer’s instructions. cDNA was obtained from 1 μg of saffron total RNA at stage 0 day with primers ZCD rev 319 (CCTTGAAAGAAGAGATGG). Tail-d cDNA was amplified using 5′-RACE abridged anchor primer supplied in the kit and primer ZCD rev 285 (TCTTGGCTGACGTCTGGA). Two microliters of the PCR were loaded on 1.5% (wt/vol) agarose gel (1× TBE) stained with GelRed Nucleic Acid Gel Stain (Biotium; catalog number 41003) using as a marker 1 Kb Plus DNA Ladder (Life Technologies; catalog number 10787018).

#### Cloning of ZCD and CCD2.

A plasmid containing the ZCD coding sequence (8) was kindly provided by Prof. B. Camara (Strasbourg, France), whereas CCD2 was isolated from saffron stigma RNA at the “orange” stage using the SMART PCR cDNA synthesis kit (Clontech; catalog number 634902) with SuperScriptII reverse transcriptase (Life Technologies; catalog number 18064-014). This novel CCD2 is similar to (97% amino acid identity) to a *Crocus sativus* CCD2 previously deposited in GenBank (EU523661), which was not present in our transcriptome samples. ZCD and CCD2 CDS were then amplified using the Advantage 2 Polymerase Mix (Clontech; catalog number 639201) and cloned in pGem-Teasy vector system I (Promega; catalog number A1360) with the following oligos: ZCD-for, GGTTAGCAAGGTGGACCCAACCA, and ZCD-rev, GTCCTAGACTGGCTGACGACGCTC. Overnight cultures of these were inoculated to an OD 0.7, and induced with 0.2% (wt/vol) arabinose grown for 16 h at 20 °C, pelleted, and resuspended in gel loading buffer according to their ODs. The samples were boiled for 10 min and 5 μL of the supernatants were loaded on 10% (wt/vol) SDS/PAGE gels using a Bio-Rad device (Bio-Rad Mini Protein). Protein bands were visualized using Colloidal Coomassie Blue.

#### SDS/PAGE.

One-milliliter cultures of *Escherichia coli* cells were grown for 16 h at 20 °C, pelleted, and resuspended in loading buffer according to their ODs. The samples were boiled for 10 min and 5 μL of the supernatants were loaded on 10% SDS/PAGE gels using a Bio-Rad device (Bio-Rad Mini Protein). Protein bands were visualized using Colloidal Coomassie Blue.


*E. coli* strains accumulating carotenoid substrates (lycopene, β-carotene, zeaxanthin) (10, 11) were transformed with pTHIO-ZCD/CCD2 or the control plasmid pTHIO-DAN1. Overnight cultures of these were inoculated into 50 mL of LB medium containing half-strength of antibiotics (25 μg·L⁻¹ kanamycin, 50 μg·L⁻¹ ampicillin), grown at 37 °C to an OD₆₀₀ of 0.7, and induced with 0.2% (wt/vol) arabinose for 6 h at 28 °C or 20 °C for 16 h. Cells were pelleted and extracted in 10 mL of acetone and dried, and the extracts were redissolved in 100 μL of chloroform before being subjected to HPLC analyses.

#### In *vitro* assay.

Carotenoids and apocarotenoids were purchased from Sigma-Aldrich, CaroteNature, and provided by BASF. pTHIO vectors were transformed into *E. coli* BL21 cells harboring pGro7, a plasmid encoding groES-groEL-chaperone system under the control of an arabinose-inducible promoter (Takara). Cells were grown at 28 °C to an OD₆₀₀ of 0.7, and induced with 2% (wt/vol) arabinose at 20 °C for 16 h (11). Crude lysates were prepared, resuspending pellet from 50-mL cultures in 1 mL of LEW buffer (50 mM NaH₂PO₄, 300 mM NaCl, 1 mM DTT, 0.1% Triton X-100, 1 mg/mL lysozyme) followed by incubation on ice for 30 min. Samples were sonicated on ice and centrifuged (30 min; 12,000 g; 4 °C). The in vitro assays were carried out as described (11) using a substrate concentration of 40 μM. Assays were incubated at 28 °C and extracted as described (11).

#### HPLC-Atmospheric Pressure Chemical Ionization–MS Analyses of Carotenoids and Apocarotenoids.

Twenty microliters of extracts were analyzed with a Discovery LTQ-Orbitrap mass spectrometry system using atmospheric pressure chemical ionization (APCI) (positive mode), coupled to an Accela U-HPLC system equipped with a photodiode array detector (ThermoFisher Scientific). LC separations were performed using a C₁₈ reverse-phase column (100 × 3.0 mm; YMC Europe). The mobile phases used were methanol (A), water/methanol 20/80 (vol/vol), containing 0.2% ammonium acetate (B), and tert-methyl butyl ether (C). The separation was developed using 95%/A/5% B for 1.3 min, followed by 80%/A/5%/B/15%/C for 2.0 min and by a subsequent 9.2-min linear gradient to 30%/A/5%/B/65%/C. UV-VIS detection was continuous from 220 to 700 nm. All solvents used...
Agrobacterium are both involved in beta-ionone release. 

CCD2 was fused C-terminally to enhanced green fluorescent protein (eGFP) using Gibson assembly (16) to create pBI121:CCD2:eGFP. The sequence of CCD2 and eGFP coding sequences was amplified with Q5 High-Fidelity DNA Polymerase (New England BioLabs, Beverly, MA) using the following primers: CCD2-F (5′-GGATCCATATGGAGCCAACTAAAC-3′) and CCD2-R (5′-GGATCCTCTAGAGCTCGAGAGCTC-3′) for CCD2 and eGFP-F (5′-GGATCCATATGGAGCCAGCGATGCTC-3′) and eGFP-R (5′-GGATCCTCTAGAGCTCGAGCGCGCC-3′) for eGFP. The amplified products were digested with BamHI and XhoI and ligated into the BamHI and XhoI-digested pBI121 vector (Clontech, Mountain View, CA) to generate pBI121:CCD2:eGFP. The sequence of pBI121:CCD2:eGFP was confirmed by DNA sequencing.

Protein Localization.

To verify the localization of CCD2 and eGFP in maize kernels, we performed transient expression in maize kernels using Agrobacterium tumefaciens-mediated transformation. Maize kernels were surface-sterilized and placed in Whatman 3MM paper soaked with 5 mL of infiltration medium (18). Each construct was agroinfiltrated with C58C1 containing either PB121:eGFP or PB121:CCD2:eGFP at least four independent clones were analyzed by confocal laser-scanning microscopy after 48 h of infiltration. Red fluorescent signals were detected with an inverted IX81 microscope (Olympus) and the GFP fluorescent signals were detected with a confocal laser-scanning microscope (Leica). Images were processed using the IMARIS software.

Plant Physiol 405, 10.1104/pp.120.310546.
Fig. S1. ClustalW alignment of carotenoid cleavage dioxygenases (CCD) families 1, 2, and 4 in saffron (Cs, *C. sativus*), *Arabidopsis* (At, *A. thaliana*), and rice (Os, *O. sativa*) from Fig. 2B. The red shading indicates conserved residues, and the gray shading indicates the chloroplast transit peptide, based on ChloroP prediction.
Fig. S2. (A) 5′-RACE on RNA from 0dA stage stigmas, indicating the presence of several CCD4 transcripts: The full-length product covers the full-length CCD4 CDS, whereas several of the truncated products cover the ZCD CDS (solid arrows below the CCD4 gene). The arrowhead on the CCD4 gene indicates the position of the ZCD rev283 primer, used for the final RACE amplification. The solid arrows below the CCD4 gene indicate the extent of the CDS 5′ regions, from the primer until the ATG. The sizes of the transcripts have been subtracted for the length of the primers (56 bp). Sixteen truncated products of 400 bp were cloned, sequenced, and found to be homologous to CCD4A (eight clones), ZCD (four clones), CCD4B (one clone), and other CCD4 (three clones). (B) SDS/PAGE analysis of total protein extracts of E. coli cells expressing thioredoxin (C−), or thioredoxin fused with CCD2 or ZCD. Extracts were obtained from noninduced (−) cells or after 16-h induction at 20 °C with arabinose.
Fig. S3. (A) LC-APCI-MS analysis of lycopene-accumulating *E. coli* cells. Lycopene-accumulating *E. coli* cells expressing CCD2 or ZCD were induced for 16 h at 20 °C with arabinose [0.2% (wt/vol)], extracted with acetone, and the extracts were run on HPLC-coupled high-resolution mass spectrometry (LC-HRMS) system. The accurate masses of lycopene (first panel), a possible C14 dialdehyde produced by 9-10′-10′ cleavage of lycopene (second panel), and crocetin dialdehyde (third panel) were extracted. No cleavage product was detected in the CCD2 or ZCD-overexpressing cells or in control cells expressing thioredoxin alone. (B) LC-APCI-MS analysis of β-carotene–accumulating *E. coli* cells. β-Carotene–accumulating *E. coli* cells expressing CCD2 or ZCD were induced for 16 h at 20 °C with arabinose [0.2% (wt/vol)], extracted with acetone, and the extracts were run on LC-HRMS system. The accurate masses of β-carotene (first panel), a possible C14 dialdehyde produced by 9-10′-10′ cleavage of β-carotene (second panel), and crocetin dialdehyde (third panel) were extracted. No cleavage product was detected in the CCD2 or ZCD-overexpressing cells or in control cells expressing thioredoxin alone.
Fig. S4. Lack of detectable in vitro cleavage activity on other (apo)carotenoids. β-carotene (A); violaxanthin (B); β-apo-8′-carotenal (C); β-cryptoxanthin (D). HPLC–photodiode array detection (HPLC-PDA) analysis shows chromatographic profiles recorded in a range of 350- to 550-nm analysis of in vitro protein of CCD2 incubated with several substrates. No conversion was observed in crude lysates from cells expressing CCD2 or from control cells.
Fig. S5. (A) In vitro assay using zeaxanthin as a substrate. HPLC-PDA analysis shows chromatographic profiles recorded in a range of 350–550 nm. The crude lysate of CCD2-expressing E. coli cells (CCD2) converted zeaxanthin (I) into 3-OH-β-apo-8′-carotenal (II). The UV-Vis spectrum of the obtained product is shown in the Inset; no conversion was observed in the lysate of ZCD-overexpressing or of control cells. The ZCD assay was also repeated in the conditions described by Bouvier et al. (8) with negative results. (B) Confirmation of the identity of the cleavage product by LC-HRMS; the extracts of the incubation mix were run on HPLC-coupled high-resolution mass spectrometry (LC-HRMS) system alongside authentic standards. The ion extraction of the accurate mass of 3-OH-β-apo-8′-carotenal confirms the identification. (C) In vitro assay using 3-OH-β-apo-8′-carotenal as a substrate. HPLC-PDA analysis shows chromatographic profiles recorded in a range of 350–550 nm. The crude lysate of CCD2-expressing E. coli cells (CCD2) converted 3-OH-β-apo-8′-carotenal (II) into crocetin dialdehyde (III) and crocetin (IV). UV-Vis spectra of obtained products are shown in the Insets. (D) Confirmation of the identity of the cleavage product by LC-HRMS; the extracts of the incubation mix were analyzed on by LC-HRMS alongside authentic standards. The ion extraction of the accurate masses of crocetin dialdehyde (III) confirms the identification. No conversion was observed in the lysate of thioredoxin-overexpressing cells (control). (E) Confirmation of the identity of the cleavage product by LC-HRMS; the extracts of the incubation mix were analyzed by LC-HRMS alongside authentic standards. The ion extraction of the accurate masses of crocetin (IV) confirms the identification. No conversion was observed in the lysate of thioredoxin-overexpressing cells (control).
**Fig. S6.** (A) In vitro assay using lutein as a substrate. HPLC-PDA analysis shows chromatographic profiles recorded in a range of 350–550 nm. The crude lysate of CCD2-expressing *E. coli* cells (CCD2) converted lutein (I) into 3-OH-α-apo-8-carotenal (II); UV-Vis spectra of obtained products is depicted in the Inset. (B) Confirmation of the identity of the cleavage product by HPLC-coupled high-resolution mass spectrometry (LC-HRMS); the extracts of the incubation mix were run on LC-HRMS system alongside authentic standards. The ion extraction of the accurate masses of lutein and its cleavage product 3-OH-α-apo-8'-carotenal (M + H+ - H2O) confirm the cleavage at the 7,8' position. No conversion was observed in the lysate of thioredoxin-overexpressing cells (control). (C) In vitro assay using 3-OH-β-apo-10'-carotenal (5 h, 28 °C) as substrate. The crude lysate of thioredoxin-CCD2-expressing *E. coli* cells (CCD2) converted 3-OH-β-apo-10'-carotenal (III) into C17 dialdehyde (IV). The online absorbance spectra of the products are shown. mAU, milli-absorbance units; UV-Vis spectra of obtained products are depicted in the Insets. (D) In vitro assay using 3-OH-β-apo-12'-carotenal (16 h, 28 °C) as substrate. The crude lysate of thioredoxin-CCD2-expressing *E. coli* cells (CCD2) converted 3-OH-β-apo-12'-carotenal (V) into C15 dialdehyde (VI). The online absorbance spectra of the products are shown. mAU, milli-absorbance units. No conversion was observed in the control corresponding to crude lysate of THX-overexpressing cells (control).
Table S1. Carotenoid cleavage dioxygenases identified in the saffron stigma transcriptome

<table>
<thead>
<tr>
<th>Classification</th>
<th>Protein length, aa</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCD1</td>
<td>546</td>
<td>Constitutive</td>
</tr>
<tr>
<td>CCD2</td>
<td>562</td>
<td>Early</td>
</tr>
<tr>
<td>CCD4a</td>
<td>579</td>
<td>Late</td>
</tr>
<tr>
<td>CCD4b</td>
<td>574</td>
<td>Late</td>
</tr>
<tr>
<td>CCD4c</td>
<td>618</td>
<td>Late</td>
</tr>
<tr>
<td>CCD7</td>
<td>627</td>
<td>Constitutive (very low)</td>
</tr>
<tr>
<td>ZCD</td>
<td>369</td>
<td>Late</td>
</tr>
</tbody>
</table>

Table S2. Optimization of agroinfiltration conditions for maize kernels

<table>
<thead>
<tr>
<th>A. tumefaciens strain</th>
<th>OD</th>
<th>Cocultivation, d</th>
<th>GUS activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>A281</td>
<td>0.2</td>
<td>2</td>
<td>+</td>
</tr>
<tr>
<td>A281</td>
<td>0.2</td>
<td>3</td>
<td>+</td>
</tr>
<tr>
<td>A281</td>
<td>0.6</td>
<td>2</td>
<td>+++</td>
</tr>
<tr>
<td>A281</td>
<td>0.6</td>
<td>3</td>
<td>++</td>
</tr>
<tr>
<td>A281</td>
<td>1.0</td>
<td>2</td>
<td>++</td>
</tr>
<tr>
<td>A281</td>
<td>1.0</td>
<td>3</td>
<td>++</td>
</tr>
<tr>
<td>C58C1</td>
<td>0.2</td>
<td>2</td>
<td>++</td>
</tr>
<tr>
<td>C58C1</td>
<td>0.2</td>
<td>3</td>
<td>++</td>
</tr>
<tr>
<td>C58C1</td>
<td>0.6</td>
<td>2</td>
<td>+++</td>
</tr>
<tr>
<td>C58C1</td>
<td>0.6</td>
<td>3</td>
<td>++</td>
</tr>
<tr>
<td>C58C1</td>
<td>1.0</td>
<td>2</td>
<td>++</td>
</tr>
<tr>
<td>C58C1</td>
<td>1.0</td>
<td>3</td>
<td>++</td>
</tr>
</tbody>
</table>

Maize kernels were agroinfiltrated using different Agrobacterium strains containing the p35S:GUS_INT:Nos plasmid at different cell densities and cocultivation days, and GUS activity was evaluated visually.

Fig. S7. Confocal microscopic detection of GFP fluorescence in N. benthamiana leaves control and agroinfiltrated with GFP or CCD2:GFP constructs. For each construct, red (chlorophylls), green (GFP), and merge (overlap of chlorophylls and GFP signals) are shown. Scale bars: 10 μm. For details, see Materials and Methods.
Table S3. Percentage conversion of different CCD2 substrates in in vitro assays performed for 16 h at 28 °C

<table>
<thead>
<tr>
<th>Substrate</th>
<th>% conversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-OH-(\beta)-apo-8′-carotenal</td>
<td>52.7</td>
</tr>
<tr>
<td>3-OH-(\beta)-apo-12′-carotenal</td>
<td>18.5</td>
</tr>
<tr>
<td>3-OH-(\beta)-apo-10′-carotenal</td>
<td>12.5</td>
</tr>
<tr>
<td>Zeaxanthin</td>
<td>4.8</td>
</tr>
<tr>
<td>Lutein</td>
<td>1.7</td>
</tr>
<tr>
<td>(\beta)-Carotene</td>
<td>0.0</td>
</tr>
<tr>
<td>(\beta)-Cryptoxanthin</td>
<td>0.0</td>
</tr>
<tr>
<td>Lycopene</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Data are the average of two independent assays.