A GDSL lipase-like from *Ipomoea batatas* catalyzes efficient production of 3,5-diCQA when expressed in *Pichia pastoris*

Sissi Miguel\(^1,6\), Guillaume Legrand\(^2,6\), Léonor Duriot\(^1\), Marianne Delporte\(^2\), Barbara Menin\(^3\), Cindy Michel\(^1\), Alexandre Oly\(^3\), Gabrielle Chataigné\(^2\), Aleksander Salwinski\(^1\), Joakim Bygdell\(^4\), Dominique Vercaigne\(^2\), Gunnar Wingsle\(^5\), Jean Louis Hilbert\(^2,7\), Frédéric Bourgaud\(^1,7\), Alain Hehn\(^6,7\) & David Gagneul\(^2,7\)

The synthesis of 3,5-dicaffeoylquinic acid (3,5-DiCQA) has attracted the interest of many researchers for more than 30 years. Recently, enzymes belonging to the BAHD acyltransferase family were shown to mediate its synthesis, albeit with notably low efficiency. In this study, a new enzyme belonging to the GDSL lipase-like family was identified and proven to be able to transform chlorogenic acid (5-O-caffeoylquinic acid, 5-CQA, CGA) in 3,5-DiCQA with a conversion rate of more than 60%. The enzyme has been produced in different expression systems but has only been shown to be active when transiently synthesized in *Nicotiana benthamiana* or stably expressed in *Pichia pastoris*. The synthesis of the molecule could be performed in vitro but also by a bioconversion approach beginning from pure 5-CQA or from green coffee bean extract, thereby paving the road for producing it on an industrial scale.
Plants can be seen as factories that produce a large variety of specialized molecules displaying interesting biological properties. Among them, phenolic compounds are reported for their human health benefits and especially for their antioxidant and anti-inflammatory properties. A well-known example is chlorogenic acid (5-O-caffeoylquinic acid, 5-CQA, CGA) formed by the esterification of caffeic acid with quinic acid. This molecule can be further combined with a second caffeic acid to produce dicaffeoyl quinic acids like 3,4-dicaffeoylquinic acid or its isomer 3,5-dicaffeoylquinic acid (3,5-DiCQA). 5-CQA and 3,5-DiCQA were detected in many higher plants, and among them, the main sources are coffee, potato and vegetables. According to epidemiological studies, their daily intake is between 100 and 224 mg for 5-CQA and between 0.2 and 7.7 mg for 3,5-DiCQA. The increasing complexity of these caffeate derivatives provides additional properties to the molecules, such as a higher solubility or stability, but also new biological activities. Many reports highlight the health benefits of 5-CQA and 3,5-DiCQA. For instance, 5-CQA prevents low-density lipid oxidation, whereas 3,5-DiCQA has antihypertensive and anti-inflammatory activities. Due to their neuroprotective activity largely described in the literature, 5-CQA and 3,5-DiCQA are considered for the treatment of Alzheimer’s disease. Additionally, 3,5-DiCQA and other chlorogenic acids have been associated to anti-virus-cell fusion properties against numerous human viruses like human immunodeficiency virus, influenza A, herpes simplex 1 and 2, Coxsackie B3, adenovirus and respiratory syncytial viruses. In parallel to these anti-viral activities, 3,5-DiCQA has also been assigned as a potent anti-inflammatory compound against lipopolysaccharide-induced acute lung injury through the inhibition of human neutrophil elastase activity, making it a potential new therapeutic candidate against acute viral pneumonia diseases. Finally, 3,5-DiCQA has also been identified as a hypotensive molecule, inhibiting angiotensin I-converting enzyme (ACE). This ACE-inhibiting activity could help protect the lung cells infected by SARS-CoV-2 since the virus is using angiotensin II-converting enzyme (ACE2), a protein homolog to ACE, as a cell entry point, which, in return, generates a disruption in the ACE/ACE2 balance.

The synthesis of 5-CQA is well documented, and four pathways have been described in the literature (Fig. 1). In the first pathway, p-coumaroyl-CoA is esterified with quinic acid by a hydroxycinnamoyl-CoA: shikimate/quinate transferase (HCT), converting enzyme (ACE). This ACE-inhibiting activity could be used as a hypotensive molecule, inhibiting angiotensin I-converting enzyme (ACE2). This ACE2-inhibiting activity could make it a potential new therapeutic candidate against respiratory syncytial viruses.

Nutritional and functional characterization of IbHCT. Abscisic acid (ABA) and 15-cis-9-epoxy-11,13-epoxy-10-deoxy-12-oxo-phytodienoic acid (15c-9oxo-PDA) are both signals for responses to severe stress conditions, such as cold or drought. The production of these molecules is tightly regulated by the expression of specific genes. In Arabidopsis thaliana, the expression of the ABA biosynthesis gene Os3234.7 is induced by cold and drought. This gene encodes a protein involved in the synthesis of ABA and 15c-9oxo-PDA. To perform functional characterization of this enzyme, it was expressed in E. coli BL21 (DE3) strains and, after purification, incubated in the presence of caffeoyl-CoA and quinic acid at 30°C and pH 7.0 for 1 h, as described in Comino et al. (Fig. 2a). In the presence of IbHCT, a new product (P1), absent in control reactions, was detected (Fig. 2a.i). P1 was identified as 5-CQA on the basis of UV spectra, mass spectra (MS) and retention time in high-performance liquid chromatography (HPLC).
comparison to an authentic standard (Fig. 2a.ii). P1 gave a $[\text{M-H}]^-$ ion at a mass-to-charge ratio ($m/z$) of 353 and exhibited a maximum absorption peak at 326 nm (Fig. 2a.iii). These results showed that IbHCT has the ability to synthesize 5-CQA as HCTs from other plant species\textsuperscript{2–5}. An additional peak (not present in the initial reaction mixtures) corresponding to caffeic acid also appeared in the positive and control reactions. This molecule may arise from caffeoyl-CoA or 5-CQA degradation, as reported by Kojima and Kondo\textsuperscript{34} and Lallemand et al.\textsuperscript{32}. In addition to these results, incubation of recombinant IbHCT with 5-CQA and caffeoyl-CoA at pH 6.0 led to the production of two new compounds (P2 and P3; Fig. 2b.i) absent in the negative control after 1 h of incubation at 30 °C. P2 was identified as 3,5-DiCQA based on UV spectra (maximum absorption at 326 nm) and MS spectra ($[\text{M-H}]^-$ ion at a $m/z$ of 515) in comparison with an authentic standard (Fig. 2b.ii and iii). P3 eluted before 3,5-DiCQA and was identified as 3,4-DiCQA by mass spectrometry (not shown). P3 may result from the isomerization of 3,5-DiCQA (Supplementary Fig. 11), as previously described in other studies\textsuperscript{32}. These results demonstrated that IbHCT is able to catalyze both 5-CQA and 3,5-DiCQA synthesis in vitro. Regarding IbHQT, only a very slight activity towards the synthesis of p-coumaroylquinic acid was monitored. No production of DiCQA was detected under our experimental conditions.

HCTs are often described to be involved in the synthesis of various hydroxycinnamoyl esters. The specificity of the enzyme in vitro conditions with cinnamoyl-, p-coumaroyl-, caffeoyl- or feruloyl-CoA as acyl donors in combination with quinate or shikimate as acceptors was assessed. As expected, in the presence of the enzyme cinnamoyl-quinate (P4 in Supplementary Fig. 2 A), p-coumaroyl-quinate (P5 in Supplementary Fig. 2B), caffeoyl-shikimate (P6 in Supplementary Fig. 3A), p-coumaroyl-shikimate

---

**Fig. 1 Simplified biosynthetic pathway of 3,5-DiCQA.** IbHCT and IbICS (in green) are two enzymes of *Ipomoea batatas* characterized in this study. PAL L-phenylalanine ammonia-lyase, C4H cinnamate 4-hydroxylase, C3H 4-coumarate 3-hydroxylase, 4CL 4-coumaroyl-CoA ligase, HCT/HQT 4-hydroxycinnamoyl CoA: shikimate/quinate hydroxycinnamoyl transferase, C3’H 4-coumaroyl shikimate/quinate 3’-hydroxylase, CSE caffeoyl shikimate esterase, IbICS isochorogogenate synthase from *I. batatas*, IbHCT HCT from *I. batatas*. 

---
(P7, P8 and P9 in Supplementary Fig. 3B) and caffeoyl-shikimate (P10 in Supplementary Fig. 3C) were detected in the reaction mixtures containing the different combinations of substrates. However, derivatives with feruloyl entities were not detected in the presence of quinate or shikimate.

This study led us to three main conclusions. First, IbHCT has a broad substrate activity as orthologous enzymes from coffee and tobacco. With the exception of the production of 3,5-DiCQA and its isomer, 3,4-DiCQA, when the enzyme was incubated with caffeoyl-CoA and quinate, we could not identify any additional peak that could correspond to dicaffeoyl shikimic acid or any other dihydroxycinnamoyl acids. Second, in a preliminary experiment, the content of 3,5-DiCQA was higher in roots grown in hydroponic conditions without nitrogen (2 g of 3,5-DiCQA / 100 g dry weight) than in medium supplemented with nitrogen (1 g of 3,5-DiCQA / 100 g dry weight) (Supplementary Fig. 4). The results obtained regarding the low efficiency of IbHCT are insufficient to explain the high concentration of 3,5-DiCQA in...
I. batatas roots grown under nitrogen depletion conditions. Third, the molecular mass of the enzyme (50 kDa) does not match the value of the chlorogenic acid:cholerogetane caffeoyl transferase described previously by Villegas et al. Therefore, another strategy was initiated in an attempt to identify additional enzymes involved in the synthesis of 3,5-DiCQA.

**Identification and functional characterization of IbICS**

cDNA library construction. Little genetic information concerning I. batatas was available in public databases, which led us to generate a RNA-seq library. Total RNA was extracted from the roots of plants deprived of nitrogen. As mentioned above, in roots of plants grown under nitrogen-deprived conditions, the amounts of 3,5-DiCQA doubled compared to the control conditions. RNA was extracted from roots of plants to construct a 454-based de novo RNA-seq library (ID: PRJNA647243). This library was normalized to ensure representation of most of the genes expressed in these tissues. Clustering and assembly of these singletons led to the identification of 65264 contigs. These contigs were annotated and functionally assigned using Blast2GO. A Gene Ontology analysis realized on the library showed 8454 different annotations (Supplementary Fig. 5A), among which 37% were assigned to catalytic activities (Supplementary Fig. 5B). To assess the coverage of our library, we searched and confirmed the presence of IbHCT along with 13 additional sequences of putative acyltransferases belonging to the BAHD family. These enzymes were distributed in almost all BAHD acyltransferase clades (Supplementary Fig. 6).

Preparation of IbICS enriched protein fractions and sequence analysis. Based on the work performed by Villegas and collaborators, slices of I. batatas tubers were incubated in the dark for 48 h at 25 °C to induce 3,5-DiCQA accumulation and promote an increase in target enzyme activity. Proteins from an acetone extract of slices of roots at 25 °C to induce 3,5-DiCQA accumulation and promote an increase in target enzyme activity. Proteins from an acetone extract of slices of roots harvested 48 h at 25 °C were separated by SDS-PAGE, and the proteins were collected in 8 different acrylamide gel fragments (~1 cm). The proteins extracted from each acrylamide piece were tested, and 3,5-DiCQA synthase activity could be highlighted in fragments #4 and #5, which correspond to proteins with an apparent molecular weight of approximately 30 kDa. The proteins present in these two bands, as well as the proteins included in the surrounding fragments (#3 and #6), were digested with trypsin and further analyzed using a UHPLC MS/MS approach. This analysis provided several hundred peptide sequences (Table 1, and Supplementary Table 1) that have been further identified and assigned to 25 fragment #3) up to 64 (fragment #5) homologous proteins present in public databases.

**Table 1 Analyses of the peptides obtained by trypsin digestion of the proteins collected by SDS-PAGE.**

| Acrylamide slice ID displaying an activity | Number of different peptides | Number of related proteins |
|------------------------------------------|-----------------------------|---------------------------|
| 3                                        | 261                         | 25                        |
| 4                                        | 230                         | 27                        |
| 5                                        | 437                         | 64                        |
| 6                                        | 779                         | 52                        |

Slices of SDS-PAGE exhibiting 3,5-DiCQA synthesis activity were analyzed by LC-MS/MS. Numbers of corresponding proteins identified in public databases are indicated.

Identification of a full-length cDNA encoding IbICS. The peptides identified by UHPLC MS/MS were used in a tBLASTn search of the de novo normalized RNA-seq library obtained from I. batatas roots to subsequently identify the corresponding coding sequences. A total of 24 proteins were identified in the four selected gel samples (Supplementary Table 1), and among these proteins, a sequence matching with a predicted Ipomoea nil GDSL esterase/lipase At1g28590-like (XM_019313071.1) was detected in fragments #3, #4 and #5. This result led us to identify IbICS, a 1158 bp coding sequence (CDS) (Genbank ID: MT291823) (Supplementary Fig. 7) encoding a putative 385-amino-acid-long protein (Supplementary Fig. 7 and 3) with a calculated molecular mass of 41.93 kDa and pI of 4.92. In silico analysis performed with the SignalP and TargetP programs revealed a predicted 25-amino acid long N-terminal signal that might target IbICS into the secretory pathway. Removing this sequence leads to a protein of 39.54 kDa with an unchanged pI. A GDSL motif was detected between amino acids 34 and 37 of the deduced IbICS amino acid sequence (Fig. 3). A sequence alignment with three additional characterized GDSL lipase-like enzymes demonstrated a notable high-sequence identity. Within GDSL, IbICS belongs to the subclass of SGNH hydrolases that are characterized by the presence of four conserved amino acids in four conserved blocks (I, II, III and V). A phylogenetic analysis showed that these enzymes clustered with GDSL identified in Rapisia serpenina (AAW88230), S. lycopersicon (CBV37053.1) and Alopecurus myosuroides (CAG27610). All these enzymes were shown to be involved in specialized metabolism (Fig. 4).

Functional characterization of IbICS. To determine the function of IbICS, the CDS was cloned into two different inducible expression systems based on E. coli and Saccharomyces cerevisiae. To analyze the effective expression of the protein in both systems, we added a polyhistidine tag at the C-terminus of the recombinant protein. The production of the protein was efficient in both systems (Fig. 5a, b). To further characterize the function, some enzymatic assays were performed in vitro in the presence of 5-CQA under the same conditions as previously described with the protein purified from the I. batatas tissues. Unfortunately, none of the proteins displayed any 3,5-DiCQA synthase activity under these conditions. We extended our investigation to a Nicotiana benthamiana plant expression system. The infiltrated leaves were harvested 96 h post infiltration, and the crude extract was used to perform incubations in the presence of 5-CQA at 28 °C for 22 h. UHPLC-MS analysis showed the appearance of a new product (P11) that has been identified as 3,5-DiCQA (molecular weight at 516) by mass spectrometry (mass in negative mode at 515) (Supplementary Fig. 8). This product was not detected when the crude extract was prepared from control leaves, which led us to assume that the enzyme is active in this system.

Kinetic parameters and substrate specificity of the recombinant protein. The protein expressed in N. benthamiana leaves was further purified using a His-Trap Ni column. The presence and identity of the recombinant protein were assessed by immunoblot analysis using an anti-His antibody (Fig. 5c). The activity of the purified recombinant protein was measured in vitro using 5-CQA as a substrate. The products of the reaction analyzed by UHPLC-MS confirmed the production of 3,5-DiCQA (Fig. 6a) with an m/z of 515 comparable to a commercial standard (Fig. 6c). No product was detected when 5-CQA was incubated in the presence of a protein extract of N. benthamiana infiltrated with an empty plasmid (Fig. 6b). The optimal pH and temperature were set at pH 6.3 and 39.9 °C, respectively (Supplementary Fig. 9). The Km was determined at 3.5 mM (Supplementary Fig. 9). The activity of IbICS with various substrates including acyl donors (various hydroxyxycinnamate...
derivatives) and other acyl acceptors (5-CQA and quinic acid) was tested, but no activity was observed under our experimental conditions (Supplementary Fig. 10).

Development of a system amenable to produce 3,5-DiCQA at industrial level. To establish a suitable process for mass production of 3,5-DiCQA, we switched from the plant system to \textit{P. pastoris} as an alternative expression factory. The IbICS CDS was cloned into the pICZ\textsubscript{\alpha} integrative plasmid that contains the \textit{S. cerevisiae} alpha secretor factor sequence favouring the excretion of the protein into the culture medium. The recombinant \textit{P. pastoris} strain was cultured under standard conditions, and protein expression was induced with methanol after 96 h. These conditions led to the production of the protein detectable in the culture medium (Fig. 5d). In a second step, different amounts of pure 5-CQA were added to the medium and the production of 3,5-DiCQA was monitored. The highest bioconversion rate of 5-CQA (61.3\%) was obtained with 7.5 mM 5-CQA (Fig. 7a), leading to the synthesis of 2.3 mM of DiCQA (instead of 3.75 mM expected for 100\% efficiency). Using a higher concentration of 5-CQA (over 10 mM) led to an inhibitory effect, and a lower concentration decreased the yields of 3,5-DiCQA produced.

Fig. 3 Alignment of the deduced amino acid sequence of IbICS with biochemically characterized GDSL lipase-like proteins. The IbICS peptide sequence was deduced from cDNA and aligned with sinapine esterase from \textit{Brassica napus} (AAX59709), acetylajmaline esterase from \textit{Rauvolfia serpentina} (AY762990) and chlorogenate: glutarate caffeoyltransferase from \textit{Solanum lycopersicum} (FR667689). The conserved GXSXXDXG motif is illustrated in bold. The predicted N-terminal leader sequences of all enzymes are shown in italics. Black box: conserved blocks in the SGNH-hydrolysase family (I, II, III and V). Amino acid residues forming the catalytic triad in the consensus sequences of blocks I and V are marked by black triangles. Sub-sequences identified by mass spectrometric sequencing after trypsin digestion are marked in red or in blue.
A 3,5-DiCQA-specific synthesis could also be highlighted when adding a green coffee extract containing 5-CQA (Fig. 7b). The bioconversion allowed a 7.5-fold increase of 3,5-DiCQA in comparison to native extract with a transformation efficiency estimated at 36%. The difference in efficiency in comparison with pure 5-CQA described above might be related to the complex composition of the green coffee bean extract where compounds structurally close to 5-CQA or other unrelated molecules might exert an inhibitory activity against IbICS.

**Discussion**

Given the interest of 3,5-DiCQA, many scientists have been searching for more than 30 years for the gene(s) responsible for its biosynthesis. In this report, our phytochemical analyses show that a decrease in nitrogen availability increased the concentrations of 5-CQA acid and 3,5-DiCQA in *I. batatas* plant root tissues. The amount of 3,5-DiCQA can reach almost 2 g for 100 g of dried roots. Such results have already been described by Galieni and collaborators, who have evidenced in lettuce that the accumulation of phenolic acids was significantly enhanced, especially 5-CQA and 3,5-DiCQA. Consequently, we carried out a combined proteomic and transcriptomic approach in this tissue to identify and functionally characterize two enzymes involved in the final step of the synthesis of 3,5-DiCQA from 5-CQA: IbHCT and IbICS.

![Phylogenetic tree of GDSL lipase/esterase protein sequences available in public databases](https://doi.org/10.1038/s42003-020-01387-1)
3,5-DiCQA was found to be optimal at pH 6, while the synthesis of chlorogenic acid was optimal at pH 7 (not shown). This difference could be explained by the fact that the activity of the enzyme may change depending on the cell compartment. It is therefore likely that the activity of 3,5-DiCQA synthesis takes place in an acidic compartment of the cell, such as the vacuole. We also observed that the reaction leading to the synthesis of 3,5-DiCQA was notably slow. Indeed, the product became detectable after a 1 h incubation time. These data are consistent with the results described by Lallemand et al. for coffee HCT. Finally, small amounts of 3,5-DiCQA were produced in vitro, which did not allow quantification. According to several studies, HCTs perform reversible reactions. It is therefore possible that the low reaction yields are related to an equilibrium between the precursor and the product. However, the low catalytic performance of HCTs for the synthesis of 3,5-DiCQA makes it difficult to explain the accumulation of high concentrations of this molecule in plants like sweet potato. Some authors suggest that 5-CQA synthesized in the cytoplasm is further transferred into vacuoles, where it is stored as 3,5-DiCQA. This side activity has been referred to as “moonlight activity” by Moglia and collaborators.

Using a proteomic/transcriptomic approach, a completely new player in the synthesis of these DiCQAs was isolated. IbICS is a gene encoding an enzyme belonging to a subgroup of the GDSL family called SGNH hydrolase. The name of this enzyme derives from the four strictly conserved Ser-Gly-Asn-His residues in four conserved blocks: I, II, III and V. It has been reported that this enzyme is generally involved in the regulation of development, morphogenesis and the synthesis of specialized metabolites and compounds involved in defence mechanisms. Our experiments showed that this enzyme could be the main actor in the synthesis of 3,5-DiCQA in sweet potato. Whereas the experiments performed with IbHCT led to the production of trace amounts of 3,5-DiCQA, we were able to convert nearly 80% of the 5-CQA that was added to the reaction mix containing IbICS in in vitro experiments. An enzyme using chlorogenate as an acyl donor (a chlorogenate-dependent caffeoyl transferase) was previously described in tomato roots, but no molecular evidence was provided. In this case, the enzyme transfers the caffeoyl group of

Fig. 5 Expression analysis of IbICS in different heterologous expression systems. Western blot analysis of proteins produced and purified from a E. coli, b S. cerevisiae, c N. benthamiana and d P. pastoris. For each sample, 1 corresponds to the purification performed on cells/tissues transformed with an empty vector. 2 corresponds to the purification performed on cells/tissues transformed with a plasmid containing the IbICS CDS. P: pellet of S. cerevisiae culture after protein extraction; S: supernatant. Parts of gels separated by a space were grouped together and lined up to facilitate the visualization of results.
5-CQA to ethanol to produce ethyl caffeate. This enzyme is inhibited by PMSF, which is a feature common to characterized members of the GDSL family and suggests that this enzyme could also belong to the GDSL family\textsuperscript{31,40,41}. This finding indicates an evolutionarily driven metabolic diversification that was already highlighted for\textsuperscript{32} SICGT\textsuperscript{42}. BAHD and SCPL acyltransferases are currently considered the two main families of acyltransferases involved in phenylpropanoid metabolism\textsuperscript{40}. In the future, it might be relevant to add the GDSL family if additional examples of such activity are evidenced.

The expression of IbICS has been successfully realized in four heterologous expression systems. Under standard conditions, the expression in \textit{S. cerevisiae} and \textit{E. coli} was effective for both systems but no activity in 3,5-DiCQA synthase was detected. This finding could be related to a poor solubilization of the proteins or to post-translational modification. Regarding the second possibility, an in silico analysis using the online GlycoEP tool (http://crdd.osdd.net/raghava/glycoep/submit.html) revealed 3 putative N-glycosylation sites (N136, N319 and N334) and 7 putative O-glycosylation sites (T51, T69, T234, T320, T347, T355 and T356). The third expression system based on agroinfiltration of \textit{N. benthamiana} leaves led to the production of a functional enzyme able to convert 5-CQA into 3,5-DiCQA. This result is consistent with other studies. For example, Tan et al. could not express a functional \textit{Brassica napus} GDSL esterase (BnGLIP) in bacteria or yeast but succeeded in \textit{N. benthamiana} leaves. In this system, these researchers obtained up to 50 µg of recombinant protein from 1 g of leaves\textsuperscript{40}. Similarly, expression of SICGT, a GDSL lipase from tomato, failed to be active in \textit{E. coli} and yeast but was successfully produced and active in tobacco plants\textsuperscript{42}. The plant expression system has also been used for the expression of CpLIP2, a GDSL lipase isolated from \textit{Carica papaya}\textsuperscript{41}. 3,5-DiCQA has been identified as a potential anti-viral and anti-inflammatory candidate molecule but it is scarcely available at industrial scale considering the low concentrations found in plants\textsuperscript{43}. Hence, producing this molecule in high amounts is an interesting challenge. In our hands, the best method was the \textit{P. pastoris} expression platform. The main advantages of this
system are that the protein is excreted in the culture medium and the IbICS produced is active. Other GDSL lipases were described to be active and are secreted out of the plant tissue. For example, JNP1, a GDSL lipase/esterase isolated from Jacaranda mimosifolia, a tropical tree, could be highlighted in the floral nectar 44. GLIP1, another enzyme belonging to this same family, has been detected in the secretome of Arabidopsis thaliana. In this case, the enzyme is produced in response to salicylic acid treatments 45.

The proposed Pichia-based bioconversion system is rather efficient since 61.3% of the 5-CQA could be converted, leading to a final concentration of 1.2 g/L 3,5-DiCQA when starting from 2.6 g/L 5-CQA. This system could be further improved by using a fed batch approach consisting in adding substrate during the process without reaching substrate toxic concentration, which seems to be a limiting factor. Another possible process improvement might be to replace the pure substrate molecule and use a plant extract harbouring high concentration of 5-CQA. This last strategy was tested in our experiments but the conversion rate of the 5-CQA present in a green coffee extract could only reach 36%. The result might be related to the presence of numerous substrates in the extract that can display an inhibitory effect on the GDSL enzyme. Finally, IbICS is soluble and active in aqueous solution. However, in our complex culture medium, with an improper pH and temperature, there might be a decline in the enzymatic activity, despite preservative buffers. The production yield might, for example, be increased through enzyme immobilization approaches that would preserve it from degradation. Such approaches were investigated in several cases and were described to provide interesting benefits, such as convenient handling, easy and efficient recovery of the product and reuse of the same enzyme batch in several reactions of conversion 6,47.

This finding was reported for a lipase isolated from Candida rugosa and immobilized on magnetic hollow mesoporous silica microspheres. In this case, the enzyme was used with fatty acids and triglycerides in a solvent-free system. The authors obtained a conversion rate of 90%, and the enzyme could be reused in 50 successive cycles.

Other potential players involved in 3,5-DiCQA biosynthesis. Our results also raise new questions. The investigation of the biochemical properties of IbICS revealed some discrepancies with the pioneering work of Villegas et al. 31. First, during the purification process, ICS activity was detected in protein fractions precipitated with 40% ammonium sulphate, while in a previous report, the enzyme precipitated between 40 and 70% saturation. Second, IbICS has a calculated molecular weight of 39.54 kDa and a pI of 4.92, whereas the chlorogenic acid:chlorogenate transferase isolated by Villegas et al. had a deduced molecular weight of 3.5 mM and an optimal pH of 6.5, while Villegas determined K_m of 0.87 mM and an optimal pH of 5.0, with the enzyme losing 50% of its activity at pH 6.8. Taken together, these differences between the two studies suggest that additional enzymes might be involved in the outstanding efficient synthesis of 3,5-DiCQA in sweet potato roots. In our study, we overcome a new step thanks to the identification of IbICS. However, given the complexity of specialized metabolism, it is possible that we have not reached the end of the story. Further experiments are needed to test new hypotheses, such as the deletion of IbICS in sweet potato, through genome editing.

Methods

Plant material. I. batatas were grown in hydroponic conditions (Supplementary Fig. 4). Young cuttings of sweet potato roots were soaked in water for 2 weeks to induce root generation. During the following 3 weeks, the plants were cultivated either in nitrogen-enhanced medium (15:10:30, N:P:K, PlantProd) or in nitrogen-free medium (0:15:40, PlantProd) with electrophoretically adjusted each day to 1.0–1.2 and a pH of 5.6–5.8.

For enzyme purification, sweet potato tubers were purchased at a local supermarket. Discs (1 cm diameter) were prepared from tuber slices (5 mm thick) using a cork borer, rinsed with distilled water and immediately placed on moist filter paper in closed plastic boxes. Discs were incubated at 25°C in the dark for 48 h and subsequently used as a source for identifying the enzyme.

N. benthamiana seeds were sown on soil and cultivated in culture rooms under a 16 h/8 h day/night photoperiod with artificial light (70 µmol m^-2 s^-1) at 26°C and 70% humidity. Three-week-old plants were transplanted into individual pots and were used after 3–4 weeks for agroinfiltration experiments, as described in Munakata et al. 48.

Cloning, expression in E. coli and purification of IbHCT. To amplify the IbHCT coding sequence (ID: BA14794), we used IbHCT_Fw: 5′-CGGTGTGATC-GAACGTGCAAGGCAC-3′ and IbHCT_Rev: 5′-CCGTGAGCA-GATATTATCA-GACACTTGAAG-3′ primers. First-strand cDNAs were synthesized from RNA extracted from I. batatas roots cultivated in nitrogen-free medium and using the SuperScript™ III One-Step RT-PCR System with Platinum® Taq High Fidelity DNA Polymerase (Invitrogen®, Thermo Fisher Scientific). The PCR products were ligated into the pcR8 plasmid (Invitrogen®, Thermo Fisher Scientific) and further transferred between the Ndel and Xhol restriction sites in the pET28b expression vector (Addgene, https://www.addgene.org/). Expression in E. coli and purification of the proteins were performed as described by Comino et al. 67.

Enzymatic in vitro assays of IbHCT. Synthesis of 5-CQA, cinamoyl-shikimate, p-coumaroyl-shikimate and caffeoyl-shikimate from cinamyl-, p-coumaryl- and caffeoyl-CoA, respectively with 5-CQA or shikimic acid. IbICS was added to a volume of 400 µL with 0.6 M of ester-CoA (p-coumaroyl-, cinamoyl- and caffeoyl-CoA), 0.6 mM of quinic acid (or shikimic acid) and 15 mg/mL of purified enzyme in 0.1 M NaPi at pH 7. The reactions were incubated at 30°C during 1 h. The reactions were stopped by adding of 10 µL of mix acetone/HCl 2 M and the solution was mixed and centrifuged at 10,000 x g for 10 min at room temperature. Products were then purified by ethyl acetate extraction for HPLC analyses.

Synthesis of 3,5-DiCQA from caffeoyl-CoA and 5-CQA was performed in a volume of 200 µL with 0.1 M of caffeoyl-CoA, 0.1 M of 5-CQA and 15 mg/mL of purified enzyme in 0.1 M NaPi at pH 6. The reaction was incubated at room temperature during 15 h, stopped by adding of 5 µL of mix acetone/HCl 2 M, mixed and centrifuged at 10,000 x g for 10 min at room temperature. Product was directly analyzed in HPLC and MS.

Construction of an RNA-seq library. Total messenger RNA (mRNA) was extracted from I. batatas roots cultivated in nitrogen-free medium using the Plant RNeasy® kit (Qiagen). A de novo normalized RNA-seq library was constructed using GS FLX+ chemistry by Eurofin MWG Operon (Ebersberg, Germany). Deep sequencing of 600,000 ESTs was performed. Clustering and assembly of the reads was performed by the supplier.

Preparation of the IbICS-enriched protein fraction and monitoring of ICS activity. Sweet potato tuber discs (200 g) were chopped with a Waring Blender in the presence of precooled aceton (−20°C) (4 times 30 s runs with 30 s cooling-off periods). The homogenate was filtered through filter paper (Whatman), and the residue was washed with cold acetone (−20°C) until the flow-through became transparent. After drying at room temperature, the aceton extract was loaded on a Sephadex G-50 column equilibrated with buffer A and eluted with a step gradient of NaCl (from 0.15 to 1 M) prepared in 50 mM Tris-HCl, pH 7.0. The protein solution was applied to a DEAE Sepharose anion exchange column equilibrated with buffer A and eluted with a step gradient of NaCl (from 0.15 to 1 M) prepared in 50 mM Tris-HCl, pH 7.0. After centrifugation to remove insoluble residues (4000 g, 5 min, 4°C), the extract was loaded on a Sephadex G-50 column equilibrated with buffer A and the proteins eluted with buffer B (0.4 mM l−1). One-millilitre fractions were collected. The fractions containing enzyme activity were pooled and concentrated using a Vivaship 6 centrifugal concentrator (100,000 MWCO, Sartorius). The concentrated solution was applied to a DEAE Sephadex A-25 column equilibrated with buffer A and the proteins eluted with buffer B (0.4 mM l−1). One-millilitre fractions were collected. The fractions containing enzyme activity were pooled and concentrated using a Vivaship 6 centrifugal concentrator (100,000 MWCO, Sartorius). SDS-PAGE analysis was conducted as described elsewhere. After electro-elution, one lane of the gel was stained with Coomassie blue, while another lane was washed twice with 2% Triton (v/v) to remove SDS, which might interfere with the enzyme activity. To localize enzyme activity, the two lanes were identically and separately cut into 8 slices (between 20 and 75 kDa, 1 cm height). The slices of the lane prewashed with Triton were ground, buffer A was added to solubilize the protein for 10 min at 4°C, and the mixture was measured at 540 nm in a spectrophotometer as described elsewhere.

ARTICLE COMMUNICATIONS BIOLOGY | https://doi.org/10.1038/s42003-020-01387-1 | www.nature.com/commsbio
For the standard assay of IbICS activity, enzyme solutions were incubated at 30°C for 15 min in 35 mM acetate buffer, pH 5.0, containing 1 mM 5-CQA and 10 mM 20 mM sodium phosphate, 0.5 M NaCl, pH 7.4) using a Polytron® PT2000E (Kinematica). The crude plant extract was clarified by centrifugation at 14,000 x g for 10 min at 4°C. The supernatant was further filtered through a 0.2 µm filter and loaded onto 1 mL HisTrap HP columns (GE Healthcare). Due to elution with 500 mM imidazole, one-milliliter fractions were collected and individually analyzed by SDS-PAGE. The fractions harbouring the target protein were pooled, dialysed with 5 volumes of phosphate buffer and concentrated using centrifugal 10 kDa molecular mass cutoff filters (Amicon® Ultra 0.5 mL Centrifugal Filters — PMNI 10 kDa, MILLIPORE). Purified proteins were stored at −20°C.

The standard enzymatic in vitro assay was performed in a volume of 50 µL with 0.2–0.4 µg of purified enzyme. For evaluation of the optimal pH, the reaction mixture was incubated for 30 min at 25°C in a buffer (0.1 M acetic acid, 0.1 M Tris, 20 mM MES) at pH values ranging from 4.0 to 9.0 with a single 5-CQA concentration of 10 mM. To evaluate the best substrate concentration, the 5-CQA concentration was modulated from 0 to 30 mM. The reactions were stopped by adding 150 µL absolute ethanol, and the solution was mixed and centrifuged at 14,000 x g for 10 min at 4°C. The supernatant was analyzed by UHPLC-MS. Reactions were run in triplicate. The m/z value was determined in triplicate by fitting Michaelis–Menten curves using SigmaPlot Software.

Expression of IbICS in E. coli and S. cerevisiae. Expression in E. coli was realized in the BL21 Rosetta strain, and the culture was performed as for CcHCT described in Comino et al.3 Expression in S. cerevisiae was performed in the WAT11 yeast strain26, as described for cytochrome P450s37.

Expression and in vivo bioconversion of 5-CQA to 3,5-DiCQA by recombinant IbICS produced by P. pastoris. The P. pastoris X-33 strain (Invitrogen™, Thermo Fisher Scientific) was transformed as specified by the supplier with pPICZαA-IbICS or empty pPICZαA used as a negative control. The culture of recombinant P. pastoris was performed as described by the supplier. Briefly, a colony was picked into 25 mL of culture medium including glycerol at 28°C at 250 rpm for 16–18 h. The yeast was centrifuged at 5000 x g for 10 min at 4°C. The yeast pellets were suspended in 250 mL of culture medium buffered at pH 6.0 and containing 0.5% methanol as an inducer. The induction was maintained for 3 days and supplied with methanol every 24 h. Three days after the onset of induction, 10 mM 5-CQA (final concentration) was added directly to the culture medium, and 3 days later, the supernatant of Pichia culture was sampled and analyzed by UHPLC-MS.

Determination of protein concentrations and immunoblot analysis. Protein concentrations of the extracts were determined using the Qubit® Fluorometer (InvitrogenTM, Thermo Fisher Scientific) with the Qubit® protein assays kit (Invitrogen™, Thermo Fisher Scientific) according to the supplier’s instructions. For immunodetection, 25 µL of protein mix was analyzed by SDS-PAGE. The separated proteins were blotted on a polyvinylidene fluoride membrane (Membrane PVDF 0.45 µm Amersham™ Hybond® P, GE Healthcare). His-tagged IbICS was detected using a rabbit primary polyclonal anti-6xHistidine Epitope Tag (Cat#R1181, Acracid) at a 1:5000 dilution in PBS (137 mM NaCl, 2.7 mM KCl, 8.1 mM Na2HPO4, 1.5 mM KH2PO4, pH 7.4) and a secondary anti-rabbit antibody conjugated to alkaline phosphatase (Cat#A0418 – 1ML, Anti-Rabbit IgG (whole molecule)—Alkaline Phosphatase antibody produced in goat, Sigma-Aldrich) diluted 1:5000 in PBS. Revelation was performed with NBT/BCIP (Promega, USA) as substrates.

Chromatography analyses. Quantification of 5-CQA and 3,5-DiCQA was performed by:

a. UHPLC Shimadzu Prominance XR system (Shimadzu, Kyoto, Japan) with a PDA detector (applied range: 190–800 nm, detection at 300 nm) using a LiChrospher® 100 RP-18 Gold column (250 mm x 4.6 mm, 5 µm, Merck, Darmstadt, Germany), maintained at 30°C. The mobile phase was composed of water, pH 2.6 adjusted using orthophosphoric acid (A) and pure methanol (B), delivered at 0.7 mL/min with the gradient of B phase as follows: 25%–5% (0–45 min); 55%–70% (45–50 min); 70% (50–55 min), 70–23% (55–60 min), 23% (60–65 min).

b. UHPLC Shimadzu Nexis X2 system (Shimadzu, Kyoto, Japan) with a PDA detector (applied range: 220–370 nm, detection at 300 nm) coupled to a mass spectrometer LCMS2020 (electrospray ionization in a negative ion mode, m/z
The 454-based de novo RNA-seq library has the ID: 12

The contents of 3,5-DiCQA and 5-CQA were determined using the standards of 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The concentrations of other hydroxycinnamate derivatives were expressed as 3,5-DiCQA equivalents.

Data availability

The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

Received: 26 April 2020; Accepted: 9 October 2020; Published online: 13 November 2020

References

1. Biesalski, H. K. Polyphenols and inflammation: basic interactions. *Carr. Optin. Clin. Nutr.* 10, 724–728 (2007).
2. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* 278, 95–103 (2003).
3. Comino, C. et al. Isolation and functional characterization of a cDNA coding a hydroxycinnamoynyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. BMC Plant Biol. 7, 1–14 (2007).
4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The contents of 3,5-DiCQA and 5-CQA were determined using the standards of 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The concentrations of other hydroxycinnamate derivatives were expressed as 3,5-DiCQA equivalents.

Data availability

The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

Received: 26 April 2020; Accepted: 9 October 2020; Published online: 13 November 2020

References

1. Biesalski, H. K. Polyphenols and inflammation: basic interactions. *Carr. Optin. Clin. Nutr.* 10, 724–728 (2007).
2. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* 278, 95–103 (2003).
3. Comino, C. et al. Isolation and functional characterization of a cDNA coding a hydroxycinnamoynyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. BMC Plant Biol. 7, 1–14 (2007).
4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The contents of 3,5-DiCQA and 5-CQA were determined using the standards of 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The concentrations of other hydroxycinnamate derivatives were expressed as 3,5-DiCQA equivalents.

Data availability

The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

Received: 26 April 2020; Accepted: 9 October 2020; Published online: 13 November 2020

References

1. Biesalski, H. K. Polyphenols and inflammation: basic interactions. *Carr. Optin. Clin. Nutr.* 10, 724–728 (2007).
2. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* 278, 95–103 (2003).
3. Comino, C. et al. Isolation and functional characterization of a cDNA coding a hydroxycinnamoynyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. BMC Plant Biol. 7, 1–14 (2007).
4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The contents of 3,5-DiCQA and 5-CQA were determined using the standards of 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The concentrations of other hydroxycinnamate derivatives were expressed as 3,5-DiCQA equivalents.

Data availability

The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.

Received: 26 April 2020; Accepted: 9 October 2020; Published online: 13 November 2020

References

1. Biesalski, H. K. Polyphenols and inflammation: basic interactions. *Carr. Optin. Clin. Nutr.* 10, 724–728 (2007).
2. Hoffmann, L., Maury, S., Martz, F., Geoffroy, P. & Legrand, M. Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J. Biol. Chem.* 278, 95–103 (2003).
3. Comino, C. et al. Isolation and functional characterization of a cDNA coding a hydroxycinnamoynyltransferase involved in phenylpropanoid biosynthesis in *Cynara cardunculus* L. BMC Plant Biol. 7, 1–14 (2007).
4. Niggeweg, R., Michael, A. J. & Martin, C. Engineering plants with increased 3,5-DiCQA and 5-CQA, respectively (both 100 mg/L in DMSO/water, pH 3.0). The data that support the findings of this study have been deposited in Genbank (https://www.ncbi.nlm.nih.gov). The 454-based de novo RNA-seq library has the ID: PRJNA647243. The iBCS CDS has the ID: MT291823.
Acknowledgements

The authors would like to thank Clément Charles for his technical support. This work was financially supported by the “Bioprolor2” project (Région Grand-Est) (to A.H.) and by the “Impact Biomolecules” project of the “Lorraine Université d’Excellence” (Investissements d’avenir-ANR) (to A.H.). Experiments performed at Charles-Violette Research Institute were supported by an ALIBIOTECH program (2015-2020) obtained from the CPER/FEDER Région Hauts-de-France. Marianne Delporte and Guillaume Legrand were supported by doctoral fellowships from the doctoral school 104 SMRE.

Author contributions

M.S. performed the genetic molecular cloning, the bioinformatic analysis and the enzymatic assays in the 4 expression systems for IbICS. L.D., C.G., V.D., J.L.H. and G.D. purified the protein from the roots of I. batatas. B.J. and W.G. performed the MS protein analyses. L.D. cloned, expressed the IbHCT and realized the functional characterization. M.B. constructed the IbICS cDNA library. O.A., S.A. and M.S. performed the enzymatic characterization of IbICS. M.S. realized the experiments with green coffee extracts and S.A. performed the UHPLC-MS investigations of IbICS. M.S., F.B., H.A., and G.D. wrote the article. H.J.L., B.F., H.A. and G.D. supervised the project and are the principal investigators.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s42003-020-01387-1.

Correspondence and requests for materials should be addressed to A.H. or D.G.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2020