Synthetic biology strategies for the production of furanocoumarins

Joana L. Rodrigues[1*], Lígia R. Rodrigues[1]

[1]Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
[*]Corresponding author. Email: joanarodrigues@ceb.uminho.pt

Background

Furanocoumarins are polyphenolic compounds produced in plants as secondary metabolites and with several therapeutic activities. Furanocoumarins such as xanthotoxin and bergapten are known to be used in several severe and difficult to treat cases of vitiligo and psoriasis. However, they are also the first line of treatment for diseases such as mycosis fungoides, the most common type of cutaneous T-cell lymphoma. Moreover, several clinical trials are ongoing to assess their potential to treat other disorders such as Crohn's disease or graft *versus* host disease [1]. Furanocoumarins are obtained from plant extraction with low yields and purities, since they accumulate in low amounts in plants, in an expensive and environmentally unfriendly process. Therefore, more sustainable and profitable methods need to be developed to obtain these interesting compounds in higher amounts and purities. In this work, we explore the use of microorganisms to produce these compounds using synthetic biology approaches.

Methods

Our previous know-how, acquired during the study of curcuminoids heterologous pathway in *Escherichia coli* [2-7], was used to design the furanocoumarins biosynthetic pathway. This pathway has the first two steps in common with the curcuminoids pathway. Herein, as a proof-of-concept we designed a biosynthetic pathway to produce psoralen, a known furanocoumarin and the first one from furanocoumarins pathway. The biosynthetic pathway involves six enzymes: tyrosine ammonia lyase (TAL), 4-coumarate-CoA ligase (4CL), coumaroyl-CoA 2'-hydroxylase (C2'H), prenyltransferase (PT), marmesin synthase (MS) and psoralen synthase (PS) (**Figure 1**). The genes were cloned in pET28 containing glutathione S-transferase (GST) tag and Duet plasmids from Novagen that are suitable for the expression of pathways containing several enzymes. In a first approach, each gene of the pathway was cloned in a single plasmid to evaluate the activity of each enzyme separately. *In vivo* and *in vitro* assays were performed according to available literature [8-10]. Heterologous protein expression was evaluated through sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE).

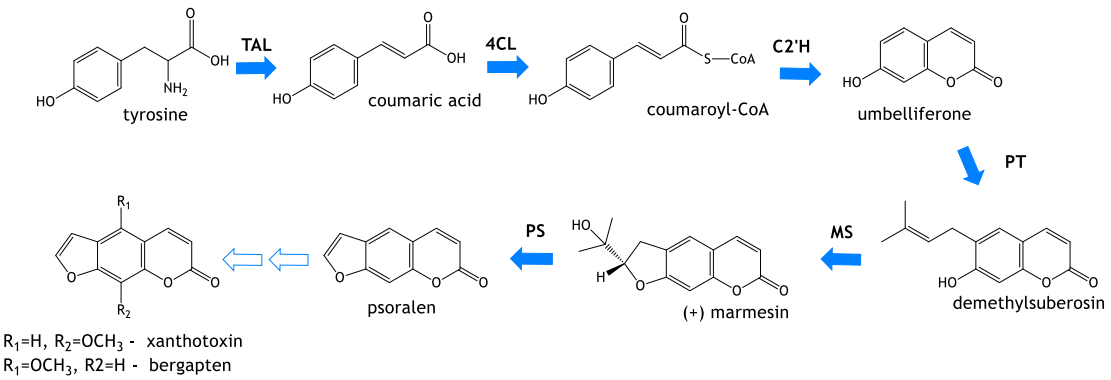

Figure 1: Biosynthetic pathway for furanocoumarins production in *Escherichia coli*. 4CL, 4-coumarate-CoA ligase; C2'H, coumaroyl-CoA 2'-hydroxylase; MS, marmesin synthase; PS, psoralen synthase; PT, prenyltransferase; TAL, tyrosine ammonia lyase.

Results

C2'H was successfully expressed (**Figure 2**) and active although mostly present in the insoluble phase. Therefore, in a first approach, two of the enzymes used in curcuminoids biosynthetic pathway (TAL from *Rhodotorula glutinis* and 4CL from *Arabidopsis thaliana*) were combined with C2'H from *Ipomoea batatas* to produce the umbelliferone coumarin *in vivo*. This combination of enzymes allowed to obtain 127 µM of umbelliferone in 18 h, while the directed supplementation of coumaric acid led to the production of 240 µM. The supplementation of caffeic acid and ferulic acid also originated escoletin and scopoletin coumarins, respectively, but with lower titers.

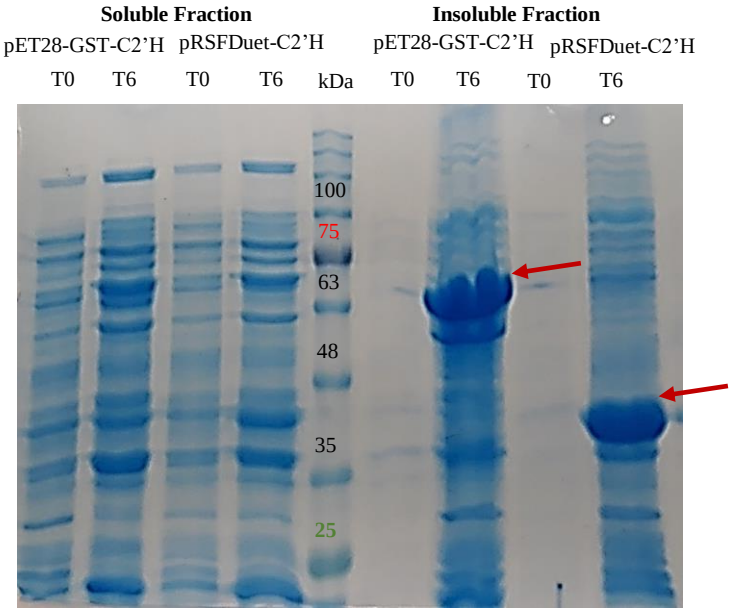

Figure 2: Protein gel showing the expression of coumaroyl-CoA 2'-hydroxylase (C2'H) from *Ipomoea batatas* in pET28-GST (glutathione S-transferase) and pRSDduet-1. T0 – before protein induction, T6 – 6 h after protein induction. Expected sizes: C2'H – 40.4 kDa, GST+C2'H – 66.1 kDa.

Next, codon-optimized PT from *Pastinaca sativa*, MS from *Ficus carica* and codon optimized PS from *Ammi majus* were evaluated *in vitro*. PT and MS were not visible in protein gels. Regarding PS, although the heterologous protein expression was observed (**Figure 3** - left), the protein size was not the expected. In addition, none of the three enzymes demonstrated any *in vitro* activity. Therefore, PT, MS and PS proteins were engineered to remove the N-terminal amino acids representing potential transit peptides related to plastid localization that can affect protein solubility. However, truncated enzymes were also not visible in the protein gels and did not present any activity. Regarding PS, that is a cytochrome P450 enzyme (CYP450), was also engineered to replace the potential transit peptide (corresponding to 48 aa) for other N-terminal sequences known to improve CYP450s solubility. However, the expression in soluble phase did not improve neither the *in vitro* activity. In **Figure 3** (on the right) it is shown an example of the protein gel after PS N-terminal replacement. The PS N-terminal was replaced by Banes sequence that corresponds to N-terminus of bovine steroid hydroxylase P450. The protein was highly expressed in the insoluble phase.

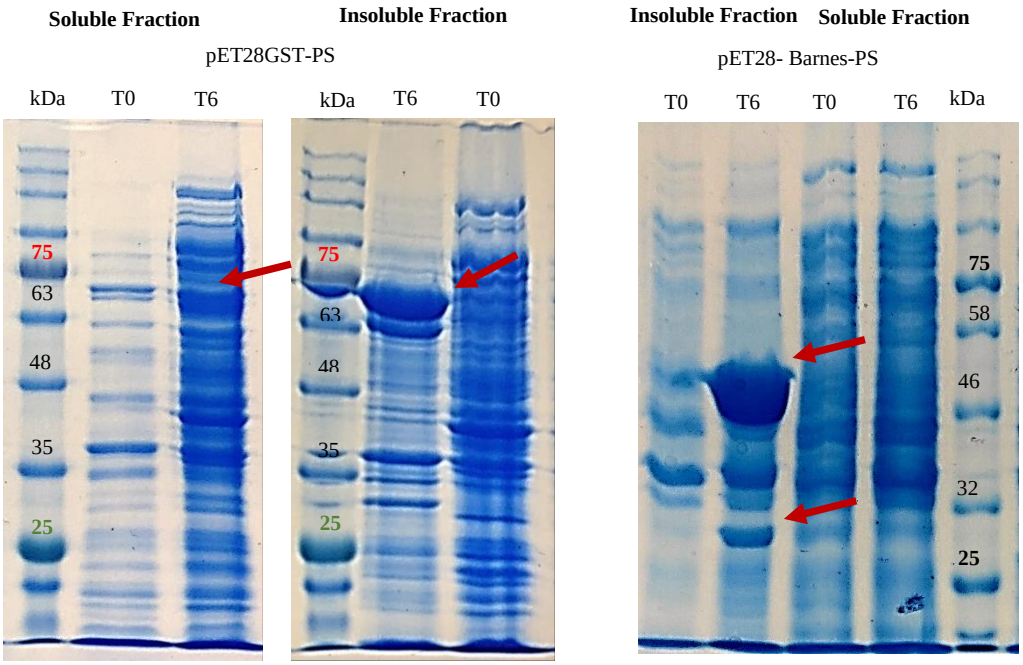

Figure 3: Protein gel showing the expression of psoralen synthase (PS) from *Ammi majus* in pET28-containing GST (glutathione S-transferase) (gels on the left) or truncated PS with Barnes sequence (MALLLAVF) upstream expressed in pET28 without GST (gel on the right). T0 – before protein induction, T6 – 6 h after protein induction. Expected sizes: GST+PS - 82 kDa, Barnes+PS – 52 kDa. PS appears to have a smaller size than expected when expressed with GST. When expressed with Barnes sequence, PS size appears correct but a smaller unidentified band is also observed.

Currently, the co-expression of chaperones (namely, the GroESL system) to facilitate furanocoumarins folding is being evaluated. Moreover, as PS is a CYP450 enzyme and

E. coli does not contain an endogenous monooxygenase system, a CYP450 reductase will be also co-expressed. In addition, PT from *Petroselinum crispum*, a PT that demonstrated higher substrate affinity and catalytic activity in *E. coli*, and XimD, a protein with MS activity, recently identified in *Streptomyces xiamenensis* bacteria [11] are also being used for the construction of the biosynthetic pathway.

Conclusion

The construction of the furanocoumarins biosynthetic pathway in microorganisms is still in an early stage. The first steps of the pathway related to coumarins production have been successfully constructed in microorganisms. However, so far, the enzymes specific of the furanocoumarins' pathway (PT, MS and PS) demonstrated low solubility and activity when expressed in *E. coli*. The expression of these enzymes in a eukaryotic microbe such as *Saccharomyces cerevisiae* should be considered. In addition, the use of enzymes identified in microorganisms, such as XimD or aromatic PTs, that are able to perform the same reactions as plant enzymes but are more soluble, might be an excellent alternative to assemble the whole biosynthetic pathway in microorganisms.

Funding

This study was supported by the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UID/BIO/04469/2020 unit.

References

[1] Rodrigues JL & Rodrigues LR. Biosynthesis and Heterologous Production of Furanocoumarins: Perspectives and Current Challenges. *Nat Prod Rep* 38, 869-879 (2020).

[2] Rodrigues JL *et al.* Heterologous production of curcuminoids. *Microbiol Mol Biol Rev* 79, 39–60 (2015).

[3] Rodrigues JL *et al.* Hydroxycinnamic acids and curcumin production in engineered *Escherichia coli* using heat shock promoters. *Biochem Eng J* 125, 41–49 (2017).

[4] Rodrigues JL, Gomes D & Rodrigues LR. A combinatorial approach to optimize the production of curcuminoids from tyrosine in *Escherichia coli*. *Front Bioeng Biotechnol* 8, 2020.00059 (2020).

[5] Couto MR, Rodrigues JL & Rodrigues LR. Optimization of fermentation conditions for the production of curcumin by engineered *Escherichia coli*. *J R Soc Interface* 14, 20170470 (2017).

[6] Rodrigues JL *et al*. Heterologous production of caffeic acid from tyrosine in *Escherichia coli*. *Enzyme Microb Technol* 71, 36–44 (2015).

[7] Rodrigues JL *et al*. Production of curcuminoids from tyrosine by a metabolically engineered *Escherichia coli* using caffeic acid as an intermediate. *Biotechnol J* 10, 599–609 (2015).

[8] Karamat F *et al.* A coumarin-specific prenyltransferase catalyzes the crucial biosynthetic reaction for furanocoumarin formation in parsley. *Plant J* 77, 627-638 (2014).

[9] Larbat R *et al.* Molecular cloning and functional characterization of psoralen synthase, the first committed monooxygenase of furanocoumarin biosynthesis. *J Biol Chem* 282, 542-554 (2007).
[10] Villard C *et al*. A new P450 involved in the furanocoumarin pathway underlies a recent case of convergent evolution. *New Phytol* 231, 1923-1939 (2021).
[11] Bu XL *et al*. Constructing microbial hosts for the production of benzoheterocyclic derivatives. *ACS Syn Bio* 9, 2282–2290 (2020).