BOUNTY HUNTERS: PROSPECTING POTENTIAL NOVEL PCA DECARBOXYLASES FOR MICROBIAL MUCONIC ACID PRODUCTION

Richard D. Cruz[1,2], Matheus A. dos Santos[2,3], Bruna L. Merlin[2], Filipe S. R. S. Benvenuto[2], Cláudio B. C. da Silva[2], Gabriel L. Chaves[2], Gabriela F. Persinoti[2], Leticia M. Z. Murakami[2], Cintia R. Sargo[2*]

[1]*Genetic and molecular biology graduate program, Unicamp / Biology Institute, Campinas, SP, Brazil.*

[2]*Brazilian Center for Research in Energy and Materials (CNPEM) / Brazilian Biorenewables National Laboratory (LNBR), Campinas, SP, Brazil.*

[3]*Federal University of Latin American Integration (UNILA), Foz do Iguaçu, PR, Brazil.*

**Corresponding author email address: cintia.sargo@lnbr.cnpem.br.*

ABSTRACT

Muconic acid (MA) is a bioprivileged molecule that can be converted into various biopolymers essential to modern society. Engineered microorganisms offer an environmentally sustainable alternative for producing MA from renewable feedstocks. However, the efficiency of this process is hindered by the low activity of protocatechuate (PCA) decarboxylase, leading to the accumulation of large amount of PCA that, subsequently, reduces MA yield and productivity. In this study, Sequence Similarity Network analysis was employed to identify potential PCA decarboxylases from assembled metagenomes and Uniprot databases. Four candidates were selected to be evaluated in vivo. Each candidate PCA decarboxylase was co-expressed with the other two enzymes of the heterologous MA pathway and tested in *E. coli* W strain for the ability to complete the pathway and produce muconic acid. Three proteins evaluated, PCA_U1, PCA_U2 and PCA_M1, showed PCA decarboxylase activity, allowing the conversion of PCA into catechol. The introduction of these enzymes, which are still unexplored in the literature, into the heterologous pathway enabled the production of 38 - 62 mg L^{-1} of MA. Expression assays showed that protein PCA_U1 was overexpressed in their soluble forms which will enable future purification steps and biochemical characterization.

Keywords: PCA decarboxylase. Muconic acid. Sequence similarity networks. Synthetic Biology.

1 INTRODUCTION

Currently, the chemical industry is considered one of the main causes of anthropogenic greenhouse gas emissions, due to its heavy dependence on fossil resources. Thus, the development of an efficient biotechnology industry, based on renewable resources, is essential to reduce environmental impacts and contribute to a more sustainable development, meeting the global demand for bioproducts. Among these bioproducts, there is a growing interest in Muconic Acid (MA), an unsaturated dicarboxylic acid, which can be functionalized into various biopolymers that are indispensable to modern society, such as PET and nylon-6,6. The use of microorganisms as platforms that convert renewable raw materials into MA has been considered a promising alternative for the development of more sustainable processes[1] .

In silico studies conducted by our research group indicated that the metabolic pathway derived from the natural intermediate 3-dehydroshikimate (DHS-to-MA pathway) presents the highest maximum theoretical yield of MA. This pathway is composed of three heterologous reactions: (1) conversion of DHS into protocatechuate (PCA); (2) conversion of PCA to catechol and (3) conversion of catechol to MA, catalyzed by DHS dehydrogenase (*asbF* gene), PCA decarboxylase (*aroY* gene) and catechol 1,2-dioxygenase (*catA* gene), respectively.

According to literature, the low activity of PCA decarboxylase is considered an important bottleneck for DHS-to-MA pathway, leading to the accumulation of the intermediate PCA and limiting MA yields and productivity from renewable feedstocks such as sugars[2,3]. In addition to being relevant to produce MA, the action of PCA decarboxylase enzymes plays a central role in the biosynthesis of catechol, an industrially relevant aromatic chemical precursor with myriad applications, such as synthetic fragrances (vanillin, eugenol), tanning agents, biomaterials and therapeutic agents[4]. Despite the biotechnological importance, few PCA decarboxylase variants have been investigated to date, highlighting an open gap for discovery of more efficient alternatives. In this study, Sequence Similarity Network (SSN) analysis was employed to expand the repertoire of these enzymes and potentially find variants with improved enzymatic activity. Four iso-functional enzyme sequences were selected and tested using an in vivo approach. PCA_M1, from metagenome and PCA_U1 and PCA_U2 from Uniprot database, successfully enabled muconic acid production indicating that it was possible to identify three active PCA decarboxylase enzymes.

2 MATERIAL & METHODS

SSN analysis was performed for prospecting new potential PCA decarboxylases. LNBR/CNPEM assembled metagenomes and Uniprot database were evaluated. Bacterial PCA decarboxylases from *Klebsiella pneumoniae* (*aroY_Kp*) and *Enterobacter cloacae* (*aroY_Ec*), widely described in the literature, were used as a seed protein in the SSN analysis. SSN analysis was conducted using a Hidden Markov Model profile (Finn et al., 2011) to identify candidates with conserved domains, setting an E-value threshold of 1e-10. The resulted SSN was analyzed in the Cytoscape v3.9 (https://cytoscape.org/) and identity threshold of 60% was adjusted to discretize the iso-functional clusters.

Selected sequences with codon-optimization for *E. coli* expression were synthesized by GeneScript and then subcloned into the pET28a(+) vector with a histidine-tag at the C and N-terminus. PCA decarboxylase from *Enterobacter cloacae* was also synthesized and used as control in the *in vivo* analyses

The genes of the heterologous MA pathway, *asbF* from *Bacillus thuringiensis*, variants of *aroY* and *catA* from *Corynebacterium glutamicum*, were subcloned into a high copy number plasmid (pRSM6 ePathbrick vector), resulting in 5 plasmids. All genes were assembled in a pseudo-operon configuration. *E. coli* W was selected as chassis for muconic acid production. Small-scale cultivations were performed in a 96-well plate setup in the Growth Profiler 960 (EnzyScreen, Netherlands). Growth was monitored for 72 hours at 37°C and 220 rpm, in triplicate. MR modified medium with supplementation of 20 g/L of glycerol was used for muconic acid production. The heterologous pathway was induced by 1 mM IPTG after the culture reached an OD of 0.6. Glycerol, acetic acid, muconic acid, catechol, protocatechuate (PCA) and 3-dehydroshikimate (3-DHS) were quantified by HPLC.

Expression assays were conducted to verify the solubility of the proteins that showed biological activity in the in vivo test. Variants of *aroY* genes were cloned into pET28a plasmid. *E. coli* BL21(DE3) cells were transformed with these plasmids and were cultivated in small scale (50 mL, LB medium, 30ºC, 220rpm). The solubility test indicated the candidates to be investigated in future purification and biochemical characterization steps.

3 RESULTS & DISCUSSION

Four potential PCA decarboxylases were selected from SSN analysis for gene synthesis and further investigation due to their proximate sequence identity (ranging from 57% to 62%), with slight differences in sequence length and molecular weight, compared to *aroY* from *Klebsiella pneumoniae* (*aroY_Kp*), as detailed in **Table 1**. Two of them are from Uniprot database and the other two from Sugarcane Bagasse Soil metagenome (LNBR consortia)

***Table 1**: Potential protocatechuate (PCA) decarboxylases selected from Sequence Similarity Network (SSN) analysis to be evaluated in vivo. PCA decarboxylase from Klebsiella pneumoniae (aroY_Kp) and Enterobacter cloacae (aroY_Ec) were applied as seed protein.*

Protein Code	Database	Probable Organism	Identity *aroY_Kp* (%)	Length (aa)	MW (kDa)
PCA_U1	Uniprot	*Klebsiella aerogenes*	62	498	54.21
PCA_U2	Uniprot	*Treponema* sp.	59	505	55.01
PCA_M1	LNBR consortia	*Bryobacteraceae* sp.	57	489	52.89
PCA_M2	LNBR consortia	*Thermoleophilia* sp.	59	494	53.60

As far as we know, there is a lack of suitable and conclusive in vitro PCA decarboxylase enzyme activity assay. Therefore, an in vivo approach was applied. Each of the PCA decarboxylase candidates was co-expressed in plasmids containing the DHS dehydratase (*asbF* gene) from *Bacillus thuringiensis* and catechol 1,2-dioxygenase (*catA* gene) from *Corynebacterium glutamicum.* PCA decarboxylase from *Enterobacter cloacae*, widely described in the literature, was employed as control.

In vivo assay of PCA decarboxylase candidates showed that three (PCA_U1, PCA_U2 and PCA_M1) among the four prospected proteins demonstrated capabilities to convert PCA into catechol when inserted into the muconic acid pathway (**Figure 1A**). Strains containing these prospected genes were able to produce 38 - 62 mg L^{-1} of MA, while control strain produced 114 mg L^{-1} of MA in 72 h small-scale cultivations.

The results achieved here support literature information regarding the PCA decarboxylase bottleneck, evidenced by the PCA accumulation by MA-producing strains, particularly for strains harboring PCA_U2 and PCA_M1, in which the production of PCA was more than 3 times higher than the production of muconic acid.

It also stands out that a high level of 3-DHS accumulation was observed (Figure 1A), indicating that the DHS dehydrogenase enzyme also needs optimization steps

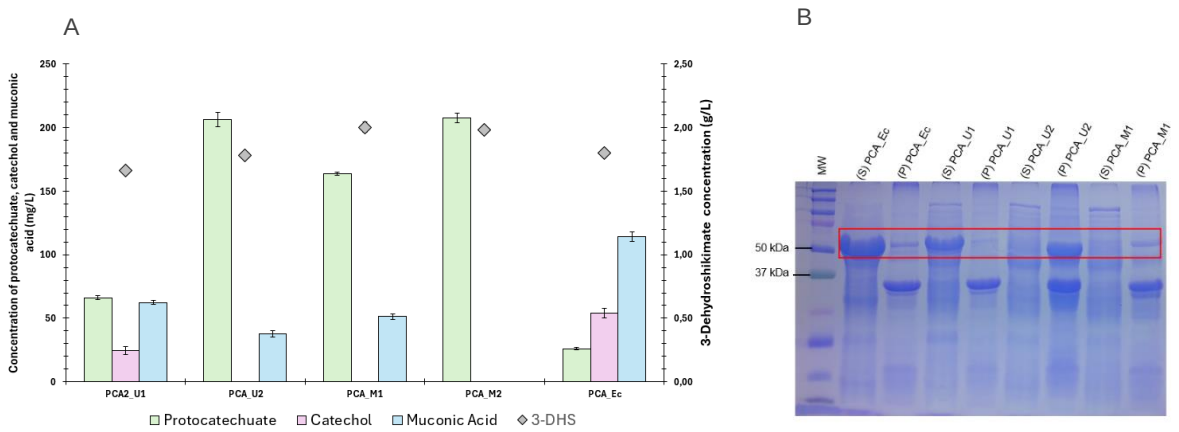

Figure 1. A) Production of Muconic Acid and intermediates in 72 h small-scale cultivations by *E. coli* strains harboring each potential PCA decarboxylase prospected in SSN analysis co-expressed with DHS dehydratase from *Bacillus thuringiensis* and catechol 1,2-dioxygenase from *Corynebacterium glutamicum.* PCA decarboxylase from *Enterobacter cloacae* was employed as control. Error bars represent standard deviations based on triplicate experiments. B) SDS-PAGE gel to evaluate expression and solubility of PCA decarbolyases propected. (S): supernatant of lysis; (P): cell pellet after lysis.

As presented in Figure 1B, the three prospected enzymes with biological activity were heterologously expressed by E. coli BL21(DE3), but only PCA_Ec (control) and PCA_U1 were soluble. Purified enzymes and/or in vitro assay using cell free extracts will be employed for the biochemical characterization of the new PCA decarboxylases prospected.

4 CONCLUSION

In conclusion, this study gives new ways to explore promising PCA decarboxylase candidates and offers insights to get over bottlenecks in the muconic acid pathway. Four PCA decarboxylase candidates prospected in Sequence Similarity Networks analysis were evaluated in vivo. Three proteins, PCA_U1, PCA_U2 and PCA_M1, showed PCA decarboxylase activity, allowing the conversion of PCA into catechol. The introduction of these enzymes into the muconic acid heterologous pathway enabled the production of 38 - 62 mg L^{-1} of MA. While further investigations are required to elucidate enzyme kinetics and optimize production yields, the findings give an important step towards developing more sustainable biotechnological solutions to a necessaire transition about dependence on fossil fuels resources at chemical industry.

The results observed here reinforce that bioinformatics tools, such as Sequence Similarity Networks, are efficient in exploring the protein sequence space and discovering new candidate biosynthetic enzymes.

5 REFERENCES

1. CHOI, S., LEE H. N., PARK, E., LEE, S. J., KIM, E. S. 2020. Biomolecules.10 (9).1-14.
2. WEILAND, F., BARTON, N., KOHLSTEDT, M., BECKER, J., WITTMANN, C. 2023. Metab Eng. 75. 153-169.
3. BRÜCKNER, C., OREB, M., KUNZE, G., BOLES, E., TRIPP, J. 2018. FEMS Yeast Res.18 (2).
4. ZEUG, M., MARKOVIC, N., IANCU, C. V., TRIPP, J., OREB, M, CHOE, J. Y. 2021. Sci Rep. 11 (1).

ACKNOWLEDGEMENTS

This research was supported by grants from Fundação de Amparo à Pesquisa do Estado de São Paulo (grant #2020/16187-6). The ph.D scholarship is provided by CAPES. Here are also the thanks of all the collaborators of the project at LNBR/CNPEM.