Search for new halogenases in native strains of *Streptomyces*

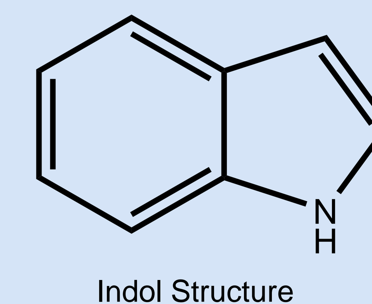
Florencia Cavallaro[1]; Sofía Zeballos[2]; Valentina Croce[2]; María Inés Siri[2]; Victoria Giorgi[1]; Agustina Vila[1]; César Iglesias[1,2]

1. Laboratorio de Biocatálisis y Biotransformaciones, DQO-DEPBIO, Facultad de Química, Universidad de la República, Uruguay

2. Laboratorio de Microbiología Molecular, Área Microbiología, DEPBIO, Facultad de Química, Universidad de la República, Uruguay

fcavallaro@fq.edu.uy

UNIVERSIDAD DE LA REPÚBLICA URUGUAY



Introduction

Many halogenated compounds are used as intermediates in the synthesis of compounds of interest to the chemical, pharmaceutical industries; however, chemical halogenation involves the use of high-cost reagents that are often harmful to the environment (1). A large number of halogenated products produced by various organisms of all kingdoms have been found in nature (2), where these reactions are catalyzed by halogenases which can be classified according to their catalytic mechanism into: a) heme, vanadium and flavin-dependent halogenases, b) non-heme iron halogenases or c) S-adenosyl-L-methionine (SAM) halogenases that react through a nucleophilic pathway (3). This work focuses on flavin-dependent halogenases.

Indol Structure

In this work we present the analysis by genomic mining in native strains for the search for new halogenases applicable to the functionalization of indole compounds of interest. From a collection of 235 strains of *Streptomyces* isolated from soil in Uruguay, 12 strains were selected for their biosynthetic potential of secondary metabolites from which their complete sequenced genomes were obtained. With data from the literature, a phylogenetic tree was created that allowed the classification of the different halogenases documented in groups of similar regioselectivity. Reference sequences were selected for each specific group, and homology searches were performed in the available genomes using pBLAST. Thus, candidate genes were identified that were analyzed by sequence alignment, detecting conserved motifs and essential amino acids for the mechanism of action of FADH-dependent halogenases. These genes were cloned into pET-28a(+) vectors for expression in *E. coli* BL21(DE3). The analysis of the reaction crudes by HPLC suggests that these biocatalysts are capable of halogenating the tryptophan indole core.

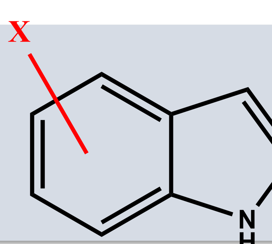
Methodology

1. Searching for sequences in the database

HHMER and pBLAST software were used and sequences with a % identity > 30% were selected in reference to reported halogenases that met amino acid requirements at the active site for halogenation. A phylogenetic tree was constructed using the Geneious® Software to classify the different documented halogenases into groups of similar regioselectivity.

2. Genomic mining for the search for new halogenases in *Streptomyces* genomes

Local database with sequenced genomes. Analysis with pBLAST coupled to Geneious Software. → Selection of ORFs (Open Reading Frames) % similarity of 30 to 75% with respect to halogenases already reported → Analysis of preserved motifs Regioselectivity prediction



3. Expression of selected enzymes

Flask Scale Culture Medium Evaluation: AIM vs. LB → Separation of the soluble (supernatant) and insoluble (pellet) phases by centrifugation. → Analysis of the expression of proteins of interest using SDS-PAGE

4. Study of biotransformation

With the recombinant strains of *E. coli* BL21(DE3) that adequately expressed the enzymes of interest, tryptophan biotransformation was performed for 24 hours at 25 °C using cell lysate, 500mM phosphate buffer, pH 7.4, 1mM tryptophan substrate. HPLC-DAD reaction crudes were analyzed to evaluate the formation of new products compared to the corresponding reaction controls (substrate and biocatalyst control).

Results

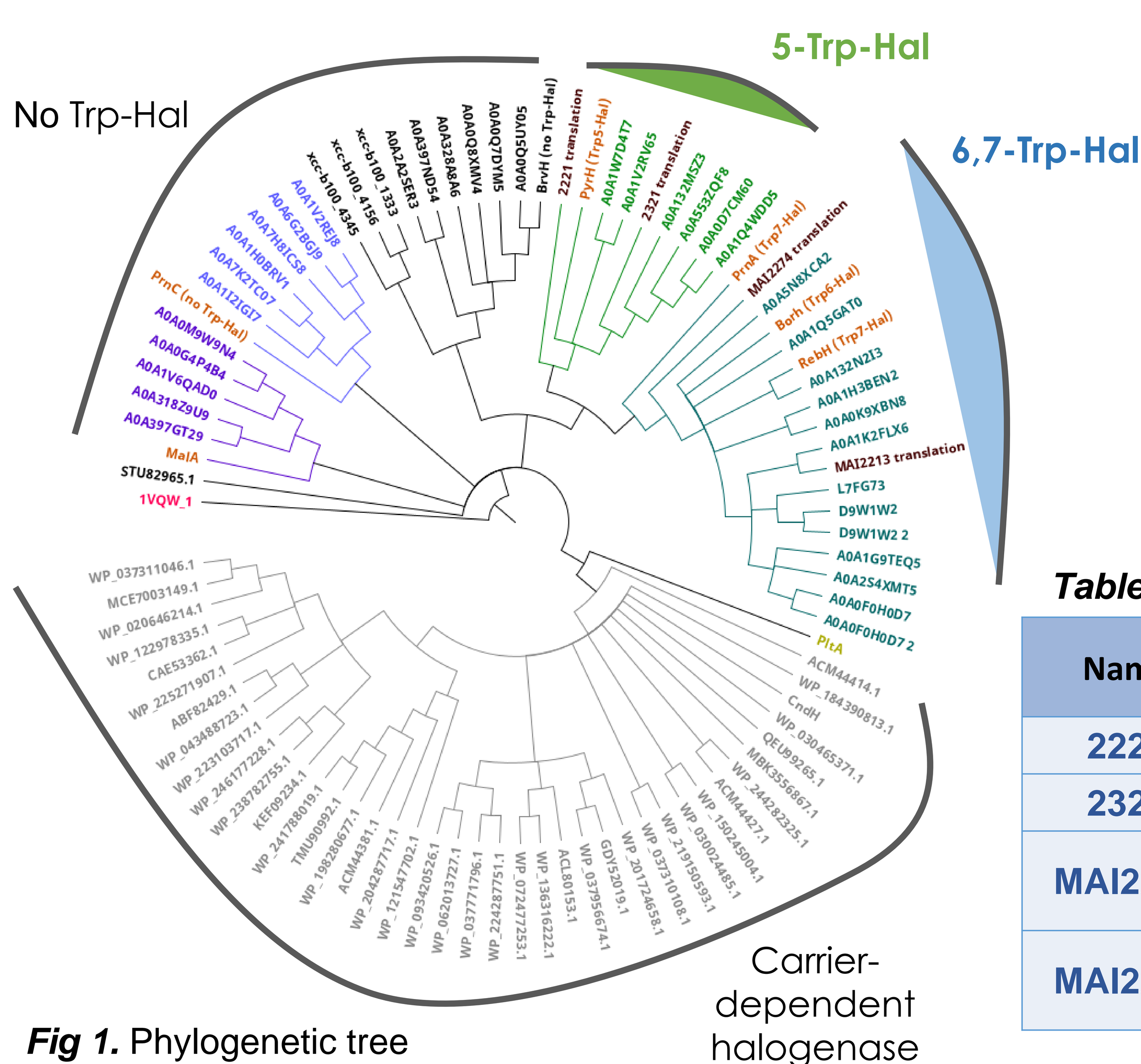


Fig 1. Phylogenetic tree

Table 1. Percentage of Identity Between Reported Enzymes and Identified Enzymes in *Streptomyces*

	2221 (possible Trp-5-Hal)	2321 (possible Trp-5Hal)	MAI2213 (possible Trp6,7-Hal)	MAI2274 (possible Trp6,7-Hal)
Borh (Trp6-Hal)	39.4	40.5	71.9	64.2
PrnA (Trp7-Hal)	38.6	37.0	53.7	54.8
RebH	38.9	37.7	67.1	63.3

Table 2. Analysis of preserved motifs

Name	Motif 1	Motif 2	Motif 3	Regioselectivity predicted
2221	ATFSTV	HPFER-QATQ	HYHGFEAYS	Trp5-Hal
2321	ATFSTV	HPFER-QRAQ	YYHGFESYS	Trp5-hal
MAI2213	ATVPN	HPFGL	YYGNFEAEFRNF WTNGSYY	Trp6-Hal
MAI2274	ATVPN	HPFGL	YYGNFEAEFRNF WTNGSYY	Trp6-Hal

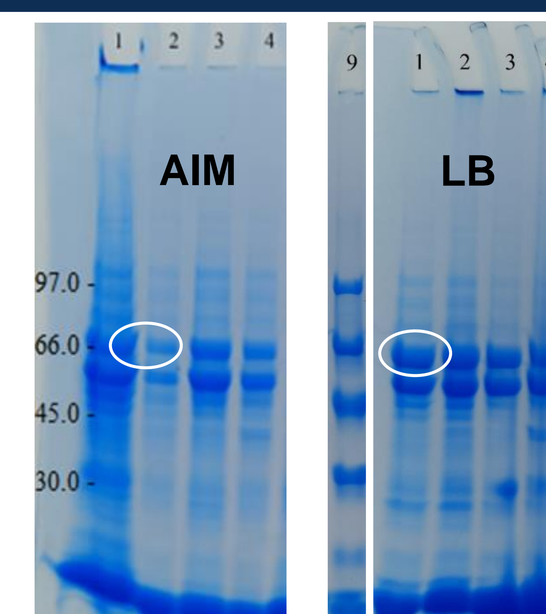


Fig.3 SDS-PAGE comparing expression in AIM and LB media. 1 and 2: soluble and insoluble phases after induction at 20°C, 24 h; 3 and 4: soluble and insoluble phases after induction at 28°C, 24 h.

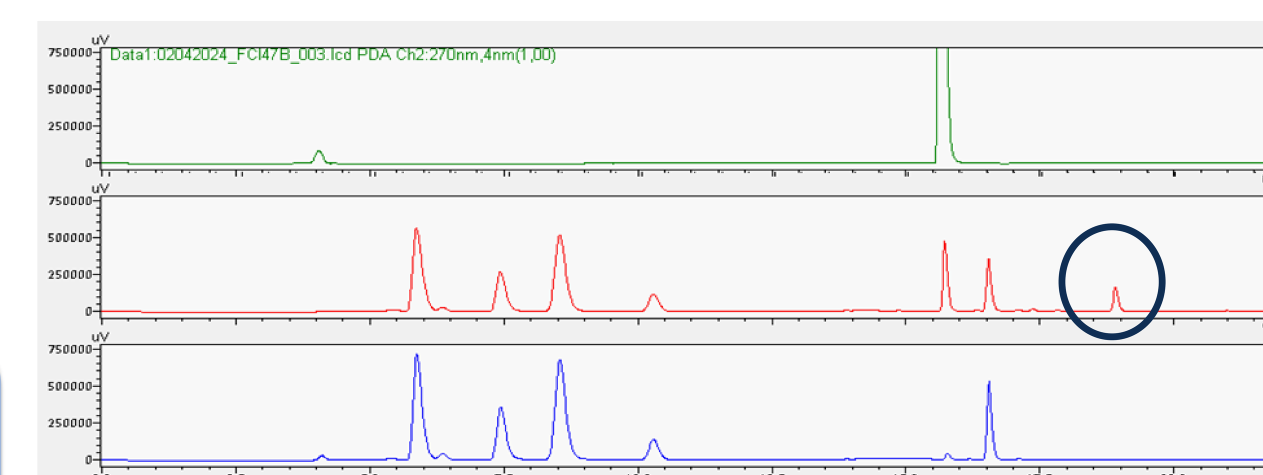


Fig.4 Chromatogram of tryptophan biotransformation using MAI2274 (red) and the corresponding controls: substrate control (green), cell control (blue).

4 strains of *Streptomyces* were identified that harbor enzymes with possible halogenase activity and that present a low to moderate similarity to reported FADH-dependent halogenase (Table 1).
The analysis of the sequences suggests that genes MAI2213 and MAI2274 belong to the Trp7-Hal group, while genes 2221 and 2321 are associated with the Trp6-Hal group (Figure 1 and Table 2).

A better expression of the enzymes of interest was observed when growing the recombinant biocatalysts in AIM medium (Figure 3).
HPLC analysis showed the formation of a new product in the biotransformation of tryptophan (Figure 4).

Perspectives

- Characterize the obtained product
- Optimize reaction conditions and evaluate range to other indole substrates
- Develop mutants of these enzymes to expand the reach to bulkier substrates

References

(1) Frese, M., *et al.*. (2014). *ChemCatChem*, *6*(5), 1270–1276.
(2) Crowe, C., *et al.* (2021). *Chemical Society Reviews*, *50*(17), 9443–9481.
(3) Büchler, J., *et al.* (2019). *Catalysts*, 9, 1030.

Acknowledgements