UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI BIOLOGIA

Corso di Laurea in Biologia

ELABORATO DI LAUREA

A COMPUTATIONAL APPROACH TO PFAS BIOREMEDIATION: COMPARATIVE ANALYSES OF DEHALOGENASES AND LACCASES ENZYMES – AN IGEM PROJECT

Tutor: Prof.ssa Brini Marisa
Dipartimento di Biologia

Co-tutor: Prof. Filippini Francesco
Dipartimento di Biologia

Laureanda: Sabrina Salmaso
Matricola 2042748

ANNO ACCADEMICO 2023/2024

page

INDEX

1. Abstract

Per- and Poly-fluoroalkyl substances (PFAS) have been an amazing industrial invention used for a wide range of daily products. Recent health-related discoveries have raised concerns that these molecules are persistent in the environment because they are not degraded naturally, so remediation solutions are needed. In this thesis, two possible enzyme candidates for PFAS biodegradation have been considered, namely dehalogenases and laccases, which are studied by bioinformatic approach to investigate whether they can be of help to PFAS degradation. Therefore, a comparative analysis of these enzymes was performed in microorganisms capable to live and grow in high PFAS concentrations.

2. Introduction

2.1 The PFAS problem

PFAS is an acronym for per- (with one fully fluorinated carbon chain) and poly- (partially) fluoroalkyl substances. PFAS are a class of synthetic chemicals comprising more than 9,000 substances with wide industrial and commercial use since the 1950s. They can be classified based on the number of carbon atoms present in their carbon chain: when there are 7 or more, they are called “long-chain”, while those with less are “short-chain” (Berhanu et al., 2023; Marchetto et al., 2021; Wackett, 2021).

Due to their persistence in nature, PFAS have been referred to as “forever chemicals”. This is because the F-C bond is the strongest and shorter bond known. These substances have started to be considered a problem since the 2000s. Thousands of tons of PFAS are produced annually and used to realize an enormous number of objects we use daily, thanks to their hydrophobic and lipophobic properties. On the other side, their continuous production has inevitably led to environmental contamination by these substances, mainly due to industrial wastewater. The discovery of bioaccumulation in living beings and some correlation between high PFAS concentrations and health problems raised some concerns. Contaminated food and water, are the main sources for human uptake of PFAS, which then accumulate in blood and organs after long exposure (Meegoda et al., 2022; Zhang et al., 2022).

Rules on PFAS production and prevention are continuously updated worldwide. For example, some types of PFAS were banned and other substitutes were produced with GenX technology. However, these new short-chain molecules

are even more dangerous to the environment (Giglioli et al., 2023; Zhang et al., 2022).

Veneto Region, in north-eastern Italy, where Università degli Studi di Padova is found, shows one of the worst situations related to PFAS contamination: in May 2013 it was discovered that due to wastewater, a large area around an industrial plant in Vicenza province was contaminated with PFAS (Figure 1), resulting in spillover to groundwater and drinking water, leading local population to a warning exposure (Giglioli et al., 2023).

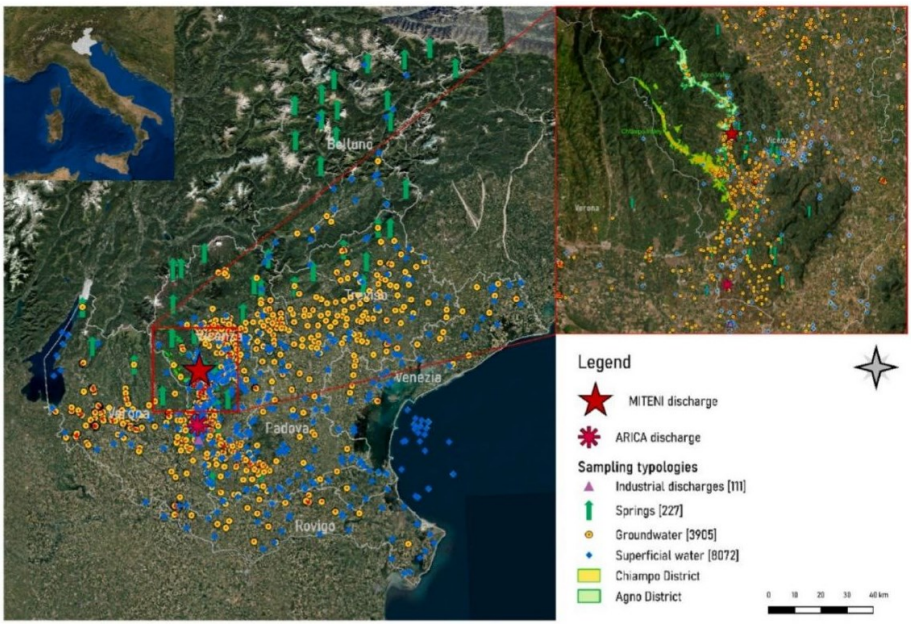

Figure 1. The location of monitoring points inside the Veneto Region (Giglioli et al., 2023).

2.2 State of the art

After the rise of the PFAS problem, many physical and chemical treatments were considered, that are still used. One of the main treatments until now has been using granular activated carbon (GAC) filters and their incineration after use, which makes PFAS spread again into the environment through the air. Others include nanofiltration (NF), exchange resin (IXR) and reverse osmosis (RO). The problem with these methods is that there is no degradation, so they remain in the environment or are stored until a definitive solution is found. On the other hand, a biological treatment would be preferred for many reasons, such as lower overall costs, less disruption to soil and water environments and possibly complete degradation (Harris et al., 2022; Meegoda et al., 2022; Zhang et al., 2022).

The main problem is that the biodegradation solution still needs many studies as there is no current solution. Studies performed so far mainly focused on the two most common PFAS types, PFOA (Perfluorooctanoic Acid) and PFOS

(Perfluorooctane Sulfonate) (Figure 2). Both molecules have an 8-carbon chain, and PFOA are listed as Substances of Very High Concern, with many negative effects already known. The problem is that defluorination on fully fluorinated compounds has been quite complicated rather than on partially fluorinated ones (Marchetto et al., 2021; Zhang et al., 2022).

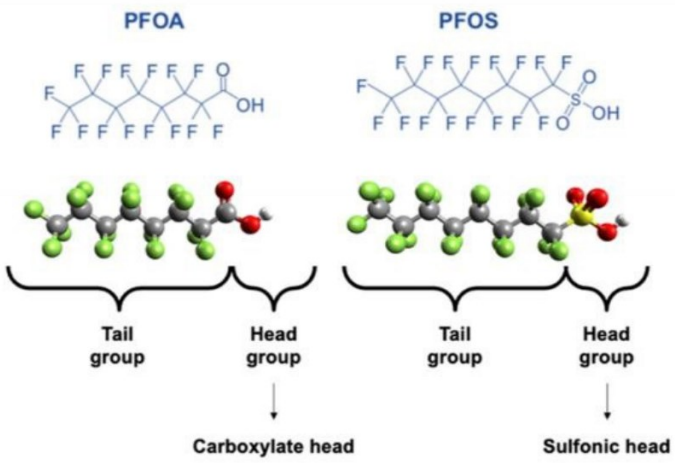

Figure 2. PFOA and PFOS chemical structures (Meegoda et al., 2022).

For a biodegradation approach, three ways can be followed:

1. Bacterial degradation, which is seemingly more effective under aerobic conditions, but also occurs under anaerobic ones. Particularly, the *Pseudomonas* sp. genus proved able to degrade PFOS, even though by an as yet unknown mechanism.
2. Fungal degradation, which is of major importance in natural environments but still needs thorough studies related to PFAS.
3. Phyto-microbial degradation, which is an important *in situ* remediation of increasing interest. Plants can uptake PFAS, and microbial collaboration can lead to some degradation. However, information is lacking, and more studies are required. Also, their toxicity in plants is still unclear and the major risk is to transmit PFAS through the food chain.
4. Phyco-remediation. In particular, the growth and respiration of *Synechocystis* sp. PCC6803 has been studied. This cyanobacterium could be a perfect candidate because it is already used in bioremediation projects and its complete genome sequence is available, allowing it to perform computational analyses. At low PFOA and PFOS concentrations, Synechocystis shows no growth alteration, while at higher an incremented level of photosynthesis and respiration are observed. Indeed, when investigating individual molecules, PFOS is found to be internalised in the cells without degradation, while PFOA is partially degraded (Marchetto et al., 2021).

In general, problems related to this approach are:

- Not all intermediate metabolites in the degradation pathways are known and need more studies.
- Complete defluorination still seems impossible.
- Short-chain PFAS obtained could be more dangerous than actual ones.
- There is a lack of characterization and identification of enzymes that can cleave the C-F bond. For now, fluoroacetate dehalogenases (FAcDs) that have been isolated from different kinds of bacteria, have been shown to process a hydrolytic defluorination and are now molecular systems for defluorination studies (Berhanu et al., 2023; Marchetto et al., 2021; Zhang et al., 2022).

In nature, defluorination doesn't occur often. However, polyfluorinated compounds can be degraded, although this process is more challenging compared to dechlorination (Table 1) (Wackett, 2021).

Table 1. Differences between dechlorination and defluorination.

DECHLORINATION	DEFLUORINATION
Polychlorinated compounds are common in nature and dechlorination can occur by multiple enzymes that have evolved for millions of years	Only a few monofluorinated compounds and no polyfluorinated ones are found in nature. Also, fluorine is rarer, resulting in a much less spread evolution of defluorinases
There are no negative effects from high levels of chloride, meaning that biodegradation can occur without many problems	Fluorine inside the cell has many negative effects on the metabolism. However it has been observed that in environments where fluoride concentration is >1mM, antifluoride protections have evolved
Polychlorinated compounds can be used as the final electron acceptor in cell metabolism instead of traditional ones because the redox potential is adequate to generate a membrane gradient	Polyfluorinated compounds have a negative redox potential, so they can't be used as final electron acceptors

2.3 The MUTANS project

MUTANS is a multidisciplinary students' project affiliated with Università degli Studi di Padova. One of the main objectives of the MUTANS team is to carry out a problem selected by the students and to find a solution to it by using a synthetic biology approach. The team then presents the results to iGEM, an international synthetic biology competition.

In 2023, a second MUTANS team was formed following the initial team's founding in 2021. The new team has decided to address the significant local PFAS problem and to work on a potential way to obtain a degradation on some of the primary compounds. The project is called *surPFAS* and covers different aspects.

Briefly, the system consists of genetically modified *Escherichia coli* that express dehalogenase and laccase enzymes on their surface. PFAS are accumulated in a GAC filter which is then regenerated. The PFAS are subsequently introduced into a bioreactor where bacteria are present, and finally PFAS-biodegradation can occur in a controlled environment (Figure 3).

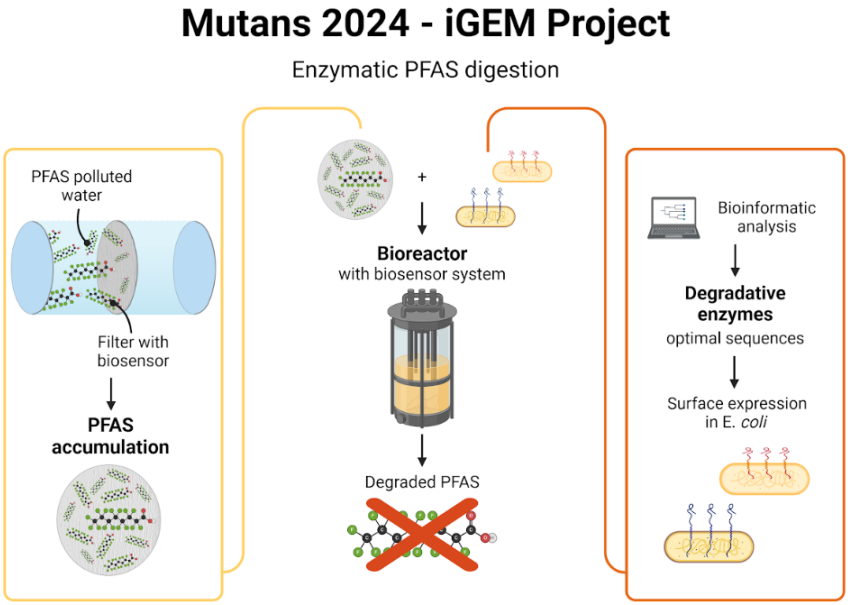

Figure 3. Graphical abstract of the MUTANS project

3. Aim of the study

In this thesis, one part of the project is presented. Specifically, the first part of the project, consisting of *in silico* screening of enzyme candidates for PFAS biodegradation to be used in the subsequent engineering part. It was decided to choose two dehalogenases and two laccases to be expressed. Our research group decided to use bioinformatic tools to study some enzymes from organisms that proved useful in previous bioremediation studies. We decided to focus our studies on:

- Fluoroacetate dehalogenase (FAc-DEX) from *Burkholderia* sp. (Vascon, 2019).
- Dehalogenases (DeHa 1, 2, 3, 4, 5) from *Delftia acidovorans* (Steel, 2021).
- Dehalogenases and laccases from *Synechocystis* sp. PCC6803 (Marchetto et al., 2021).
- Dehalogenases and laccases from *E. coli.*
- Fungal laccases, specifically from *Pleurotus ostreatus* (Luo et al., 2015) and *Pycnoporus* sp (Luo et al., 2018).

3.1. What are dehalogenases

Dehalogenases are a class of enzymes which catalyse the removal of halogen atoms (e.g. Fluorine (F), Chlorine (Cl), Bromine (Br), Iodine (I)) from organic molecules. There are many kinds of dehalogenases, based on the type of halogen and the specific substrate they act. For example, haloacid dehalogenases are a subset of dehalogenases that act specifically on organic acids containing halogen atoms, catalysing the removal of the halogen atom typically through hydrolysis. Again, fluoroacetate dehalogenases are a category of haloacid dehalogenases. Specifically, fluoroacetate dehalogenases can catalyse the break of the strong C-F bond and remove fluoride atoms from halogenated molecules through hydrolysis. Naturally, they detoxify fluoroacetate, which is toxic to many animals, by hydrolysing the carbon-fluoride bond, resulting in fluoride and glycolate. The catalytic mechanism is not yet fully understood, but an aspartate residue appears to be crucial for a nucleophilic attack on the carbon of fluoroacetate, releasing the fluoride and forming an enzyme-ester tetrahedral intermediate. Secondly, a histidine residue activates the hydrolysis and regenerates the nucleophile (Figure 4).

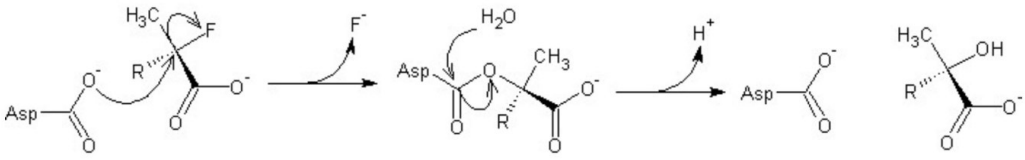

Figure 4. Mechanism of the nucleophilic attack, followed by the tetrahedral intermediate and final regeneration by water (Scott & Miao, 2024)

Due to their properties, it has been suggested that fluoroacetate dehalogenases could be used with substrates other than their natural ones and have been considered one of the main enzymes that could achieve complete PFAS biodegradation. Also, it has been seen that the enzyme research could be extended in all dehalogenase classes. Much research has been done, and some kinds of fluorinated compounds have been proven to be defluorinated, but we are still far from achieving complete biodegradation, suggesting that further research needs to be done (Scott & Miao, 2024).

3.2. What are laccases

Laccases are metalloproteins belonging to the multi-copper oxidases. In fact, they typically contain four copper atoms, distributed in three sites: Type-1 (T1),

Type-2 (T2) and Type-3 (T3), which is the one with two copper atoms. Laccases naturally have a wide range of substrates, as they're specifically important to decomposers for lignin degradation and other organic compounds. In the mechanism, all three copper sites participate in the reaction. They catalyse the oxidation of substrates: T1 obtains an electron from the substrate and transfers it to the trinuclear cluster T2/T3, where an O_2 serves as an electron donor and is reduced to water. This regenerates the enzyme's active site, which can start another cycle (Figure 5). Laccases still raise some questions, as they seem to vary a lot between different organisms under different aspects, such as the length sequence, the three-dimensional structure, and the number of copper sites. Generally, fungal laccases are more studied and widely used for biodegradation, while bacterial laccases are still less known and very different from fungal ones but are still useful in many industrial and environmental applications.

Laccases can help achieve PFAS biodegradation, as they seem to break the internal bonds of these molecules, resulting in shorter chains. Both PFOA and PFOS degradation have been observed as partially fluorinated compounds, starting from their initial state as fully fluorinated (Martin et al., 2024; Scott & Miao, 2024).

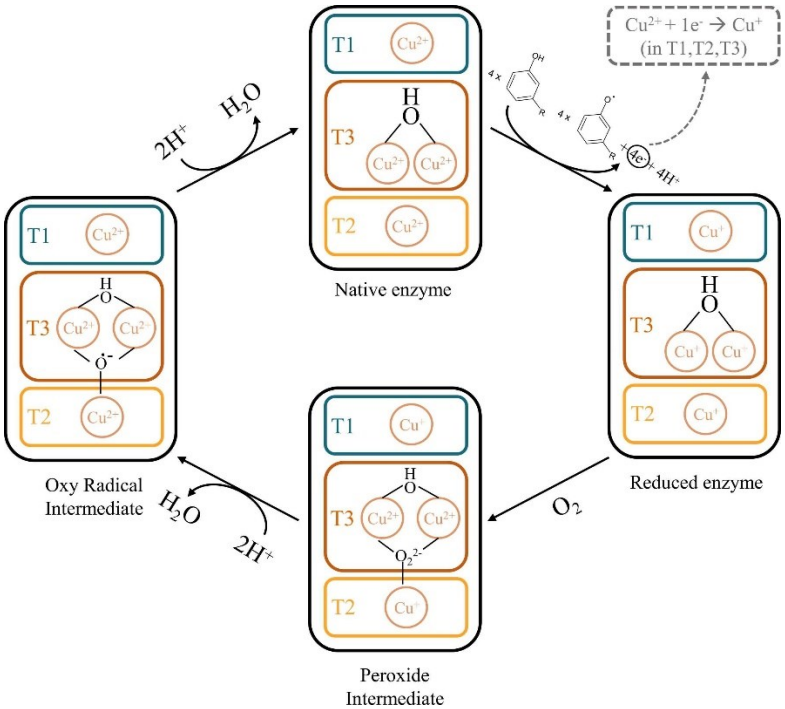

Figure 5. Mechanism of laccases function (Martin et al., 2024)

Shorter chains can be a double-edged sword: shorter intermediates might be even more harmful than the actual versions of PFAS, but also, shorter PFAS enter better in dehalogenase's catalytic sites. For this reason, the idea of a combined system utilising both laccases and dehalogenases could be revolutionary for the

research field that has been seeking a solution for PFAS degradation for many years.

4. Materials and Methods

4.1. BLAST analyses

Search for orthologous enzymes was performed using BLAST (Basic Local Alignment Search Tool) at the NCBI server (https://blast.ncbi.nlm.nih.gov/Blast.cgi). BLAST is the most widely used local alignment tool, and it allows to compare either nucleotide or protein sequences to the ones stored in databases, providing users with levels of similarities between them, assigning scores to sequences as an output. At the NCBI graphical interface, the user can change several search parameters, including database selection, introducing eventual taxonomic restrictions and changing some algorithm parameters such as e.g. the E-value, the substitution matrix etc.

In our analyses, we did not change algorithm parameters, and only taxonomic restrictions were applied when needed. In the BLASTp runs, taxonomic restriction with taxid identifiers was used to narrow the research and look for orthologous proteins only in, respectively, the *Pseudomonas* genus (taxid:286), *E. coli* (taxid:562) and *Synechocystis* PCC6803 (taxid:1148).

4.2. Docking simulations

A docking simulation is a bioinformatic technique that helps to predict possible different orientations of a ligand (in our case, different PFAS molecules) when it binds to an enzyme binding site (in this work, dehalogenases and laccases, referred to as targets). The software assigns scores to the possible combinations by calculating the binding's free energy (ΔG) and giving us a final result, rating them from the best to the worst.

For our simulations, we used Autodock Vina (https://vina.scripps.edu/) in UCSF Chimera (https://www.rbvi.ucsf.edu/chimerax/) and CB-dock2 (https://cadd.labshare.cn/cb-dock2/index.php). Since there are many variables, there are also many formulae to assign scores, but the two mentioned software's results are completely comparable. Then, three-dimensional structures of both protein and ligand are needed in PDB format to be used in the software.

4.3. InterProScan

InterProScan (https://www.ebi.ac.uk/interpro/) is a bioinformatics tool used for the functional annotation of protein sequences. It defines features as domains,

protein families, conserved sites, and more. Analyses are made on information from numerous databases to obtain better and clearer results.

5. Results and Discussion

5.1. Dehalogenase candidates

With our research, we wanted to find two dehalogenase candidates for the subsequent expression in *E. coli*. For the first candidate, it was decided to use one of the initial enzymes taken into consideration, among the ones from *Burkholderia*, *D. acidovorans* or *E. coli*. Docking simulations were performed to choose the best candidate. For the second enzyme, we looked for orthologues of the initial enzymes that also obtained high scores in docking simulations.

In searches for the second dehalogenase candidate, the five dehalogenases of *D. acidovorans* were used as protein sequence queries (Table 2). In reality, DeHa3 didn't show any activity, but we wanted to try testing them anyway to verify previous studies (Steel, 2021).

Table 2. Lists of fluoroacetate dehalogenases from *D. acidovorans*

DeHa1	PZP66635.1
DeHa2	WP_011137954.1
DeHa3	WP_198821738.1
DeHa4	WP_012206086.1
DeHa5	WP_099734849.1

In the first research, *E. coli* K12 and BL21 were searched, but the obtained results suggested amplifying the research to E. coli was a better option. Opposite to that, for the *Pseudomonas* genus, it was better to narrow the research to only some species, which we still had to choose. Finally, threshold filters were added to find outputs with identity higher than 25% and coverage higher than 70%.

The BLAST in *Synechocystis* gave results only when DeHa4 was used as a query, while the others had no significant results. Instead, two orthologous of DeHa4 were found, respectively WP_010872272.1 e WP_010872152.1 (Figure 6). On the other hand, *E. coli* provided many results while using all different *D. acidovorans*' dehalogenases (Appendix 1).

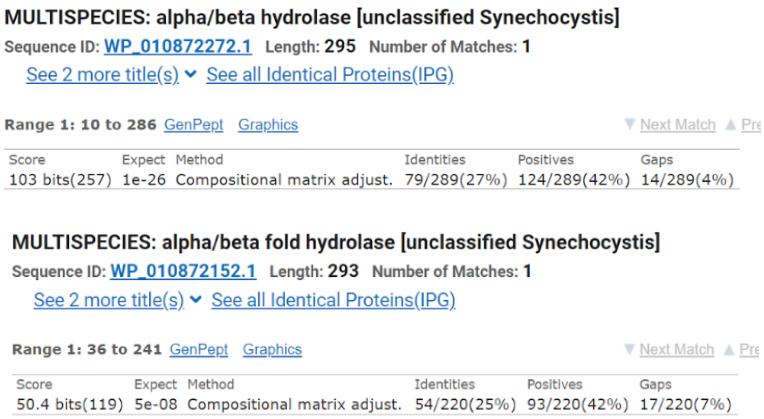

Figure 6. Synechocystis' orthologous to DeHa4 from D. acidovorans

Finally, we started another research with some enzymes found in the supplementary material from one of our initial research papers (Marchetto et al., 2021). Here, possible enzyme candidates in *Synechocystis* are found, named alpha/beta fold hydrolases. There is a list of nineteen enzyme candidates with a similarity of 25-39% to *Burkholderia* sp. Fac-dex, and another fifty-three candidates with a similarity of 22-40% (Appendix 2). A multiFASTA with all 72 enzymes was created and used as query sequences for BLAST alignment to the five dehalogenases from *D. acidovorans*, the ones from *E. coli*, and the FaC-DEX from *Burkholderia*. For *D. acidovorans*, also here DeHa4 was the only one that obtained some results. Specifically, there were 11 results, which were then filtered to maintain only the ones from *Synechocystis* PCC6803 (Appendix 3). This led to 9 results, among which the two best hits from the BLAST were WP_010872272.1 and WP_010872160.1. *Burkholderia*'s FaC-DEX gave 9 results (Appendix 4).

To find the first enzyme for expression, docking simulations were done on the five dehalogenases from *D. acidovorans*. DeHa2 obtained the best score, indicating that it could be the best one for the degradation. Also, it confirmed team USAFA results, to which also seemed the best (Steel, 2021). To find the second enzyme, other simulations were done on all the orthologous found until now. It resulted that our best hit from *D. acidovorans* DeHa4 orthologous research in *Synechocystis* obtained the highest score (WP_010872272.1, in the table, referred to as UPI00000C10BF) (Table 3 and Figure 7).

Table 3. Some of the docking results on different PFAS types. FA - Fluoroacetic Acid; TFA - Trifluoroacetic Acid; PFPA - Perfluoropropionic Acid; PFBA - Perfluorobutanoic Acid; PFBS - Perfluorobutane Sulfonate; PFHxA - Perfluorohexanoic Acid; PFHxS - Perfluorohexane Sulfonate; PFOA - Perfluorooctanoic Acid; PFOS - Perfluorooctane Sulfonate.

DHA	FA	TFA	PFPA	PFBA	PFBS	PFHxA	PFHxS	PFOA	PFOS
PZP66635	-3.4	-4.5	-5.3	-6.0	-7.2	-7.3	-5.8	-7.6	-7.6
WP_011137954	-3.6	-4.9	-4.5	-6.3	-6.3	-6.3	-6.5	-6.6	-6.9
EIQ7173218	-3.6	-4.0	-4.9	-5.5	-6.2	-6.8	-6.5	-7.8	-7.8
HAX5005898	-3.5	-5.1	-6.1	-6.8	X	-6.0	-6.1	-6.4	-6.4
HCN0551385	-3.5	-5.1	-5.8	X	X	-6.0	-6.2	-6.6	-6.2
HCP7681605	-3.5	-5.1	-5.8	-6.9	-5.4	-6.0	-6.3	-6.6	-6.5
LinB	-3.2	-4.4	-5.9	-6.6	-6.5	-7.7	-7.6	-7.7	-5.2
UPI00000C10BF	-3.6	-4.5	-4.9	-6.4	-7.2	-7.2	-6.7	-8.3	-8.3
UPI00000D34C4	-3.7	X	X	X	X	-6.3	X	-7.9	-8.4
FAC-DEX	-4.2	-4.9	X	X	X	X	X	X	X

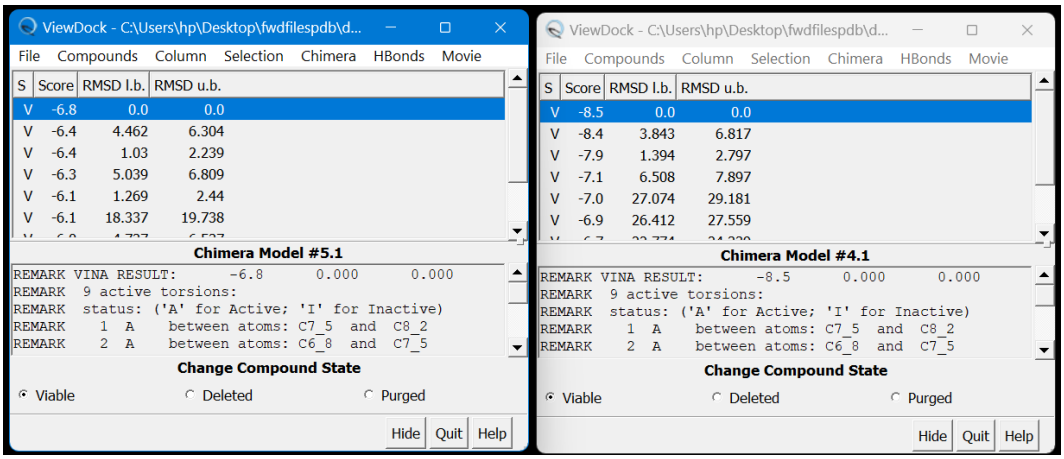

Figure 7. Docking analysis of *Synechocystis* dehalogenase and PFOS

5.2. Laccase candidates

We then started looking for the two laccase candidates. Here we had fewer starting points than for dehalogenases. But from some laccases from the literature, we tried to look for orthologues with BLAST and make docking analyses here too.

To find our laccase candidates, we started with laccase WP_194014464.1, named Multicopper oxidase domain-containing protein (Marchetto et al., 2021). In the supplementary material, this laccase from *Synechocystis* is present, which has 24% similarity to *Pleurotus ostreatus* laccase POX1. We used BLASTp to look for other homologues in *Synechocystis*, using WP_194014464.1 as the query sequence and a filter for coverage higher than 50%. We found four results (Figure 8).

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
multicopper oxidase family protein [Synechocystis salina]	Synechocystis salina	994	994	100%	0.0	100.00%	487	WP_194014464.1
multicopper oxidase family protein [Synechocystis sp.]	Synechocystis sp.	500	500	99%	1e-174	54.77%	492	MEB3229074.1
multicopper oxidase domain-containing protein [Synechocystis sp. PCC 7509]	Synechocystis sp. PCC 7509	165	165	84%	3e-45	29.19%	498	WP_009631957.1
multicopper oxidase family protein [Synechocystis sp. PCC 7509]	Synechocystis sp. PCC 7509	157	157	86%	3e-42	27.15%	500	WP_009631681.1
multicopper oxidase domain-containing protein [Synechocystis sp.]	Synechocystis sp.	85.1	85.1	55%	3e-18	26.06%	301	MEB3160845.1

Figure 8. Laccases homologues in *Synechocystis* sp.

Also, another laccase from *E. coli* was used here as a query in another BLAST research, while using the taxid identifier for *Synechocystis*. The enzyme EKK0468176.1 was used in a previous iGEM project, but we also discovered that it was an incomplete sequence, and we found the entire one, P36649. Eight results were found, many of them equal to the ones of the first BLAST research (Figure 9).

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
multicopper oxidase domain-containing protein [Synechocystis sp. PCC 7509]	Synechocystis sp. PCC 7509	142	142	53%	1e-37	36.21%	498	WP_009631957.1
multicopper oxidase family protein [Synechocystis sp. PCC 7509]	Synechocystis sp. PCC 7509	137	137	82%	1e-35	29.00%	500	WP_009631681.1
multicopper oxidase family protein [Synechocystis sp.]	Synechocystis sp.	114	114	68%	2e-27	29.78%	492	MEB3229074.1
multicopper oxidase family protein [Synechocystis salina]	Synechocystis salina	97.8	97.8	68%	7e-22	28.75%	487	WP_194014464.1
multicopper oxidase domain-containing protein [Synechocystis salina LEGE 06155]	Synechocystis salina LEGE 06155	79.0	79.0	41%	1e-16	30.11%	234	MBE9175477.1
multicopper oxidase domain-containing protein [Synechocystis sp.]	Synechocystis sp.	59.3	59.3	21%	2e-09	34.69%	543	MEB3227411.1
multicopper oxidase domain-containing protein [Synechocystis sp. PCC 7509]	Synechocystis sp. PCC 7509	49.3	49.3	28%	3e-06	26.87%	333	WP_009632146.1
multicopper oxidase domain-containing protein [unclassified Synechocystis]	unclassified Synechocystis	48.9	48.9	22%	3e-06	27.27%	356	WP_190595578.1

Figure 9. Laccases in *Synechocystis* sp. orthologs to *E. coli* laccase

Finally, another BLAST research was conducted, using some fungal laccases and the taxid identifier for *Synechocystis*. The first organism chosen is *Pleurotus ostreatus*. We used its lac1 (Q12729) and lac2 (Q12739) as queries in BLASTp. Both laccases gave the same two results, with slightly different score numbers (Figure 10).

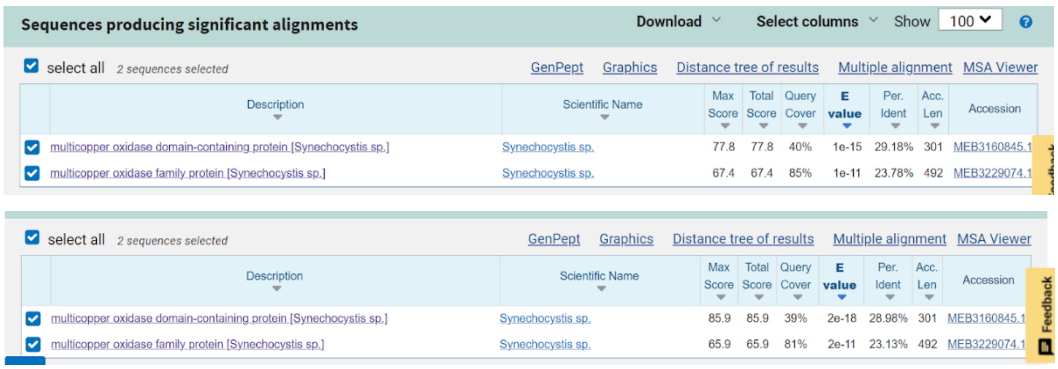

Figure 10. On top, the results from lac1; on bottom, the results from lac2

Then, we used a laccase (O59896) from *Pycnoporus* sp. as a query. We found five results in *Synechocystis*, among which also the two laccases resulted from the previous blast (Figure 11).

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
multicopper oxidase domain-containing protein [Synechocystis sp.]	Synechocystis sp.	89.4	89.4	41%	2e-19	31.35%	301	MEB3160845.1
multicopper oxidase family protein [Synechocystis salina]	Synechocystis salina	72.0	72.0	83%	3e-13	23.53%	487	WP_194014464.1
multicopper oxidase family protein [Synechocystis sp.]	Synechocystis sp.	66.6	66.6	88%	2e-11	24.54%	492	MEB3229074.1
multicopper oxidase domain-containing protein [Synechocystis salina LEGE 06155]	Synechocystis salina LEGE 06155	60.8	60.8	20%	2e-10	34.51%	234	MBE9175477.1
multicopper oxidase domain-containing protein [Synechocystis salina LEGE 06155]	Synechocystis salina LEGE 06155	35.8	35.8	18%	0.033	30.21%	193	MBE9175478.1

Figure 11. Results from Pycnoporus sp.

Among all results, MEB3229074.1 is the *Synechocystis* laccase with the highest score and similarity to the three fungal laccases. The same laccase also resulted in the two previous BLASTs, both the one with *E. coli* and the one with *Synechocystis*.

We used Interproscan to understand the presence of the three copper domains in the laccases resulting from the BLAST research. From the main page, we just inserted the FASTA of the protein sequences and the result is given in a few seconds. This passage was added as it seemed from the aminoacidic sequence that one of the copper sites was missing in *Synechocystis* laccase (MEB3229074.1). In InterProScan, T1, T2 and T3 sites are referred to as CuRO_1, CuRO_2 and CuRO_3 domains. Under the "Residues" section, *E. coli* laccase (P36649) presents all three domains, while the *Synechocystis* one only has domains CuRO_1 and CuRO_3, missing CuRO_2.

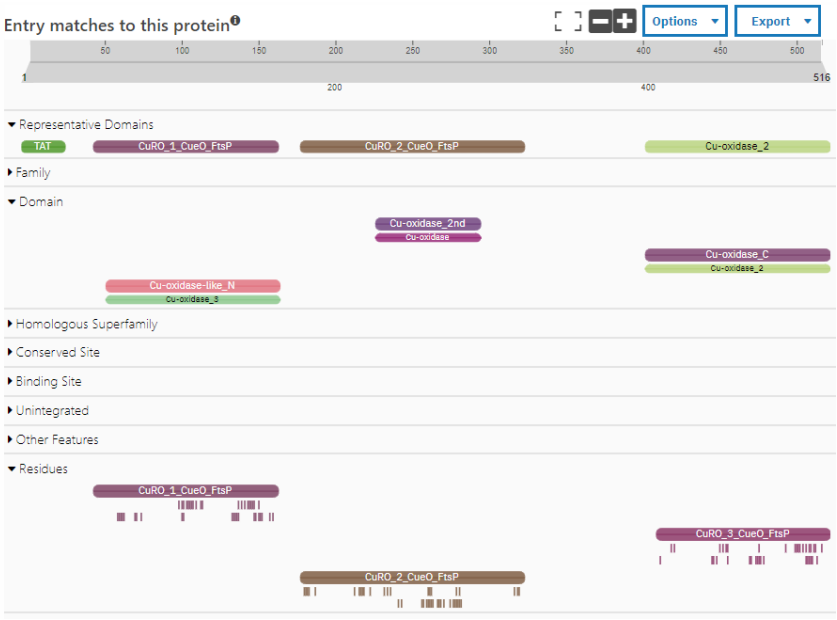

Figure 12. *E. coli* laccase

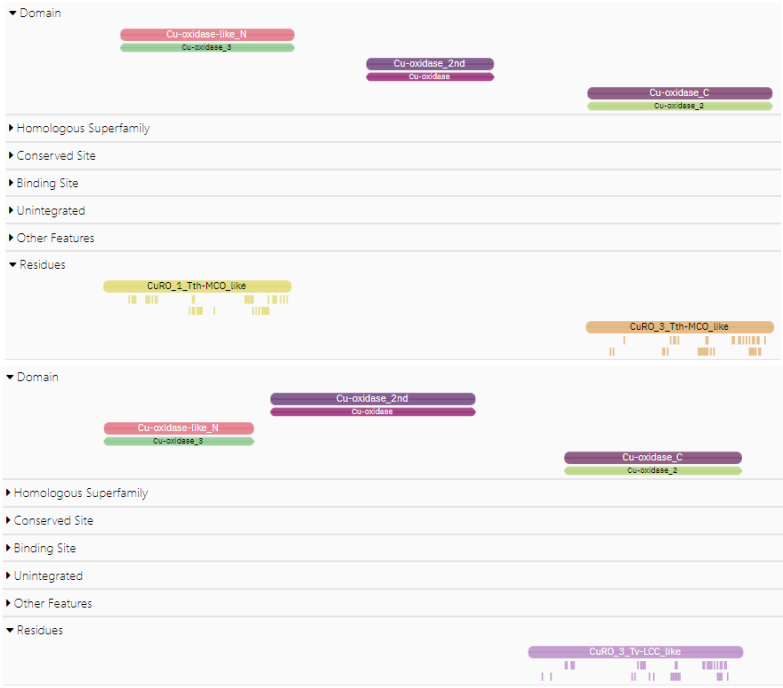

Figure 13. Respectively, *Synechocystis* and *Pleurotus ostreatus* laccases

Also here docking simulations were done on all enzymes resulting from BLAST research. This time, docking simulations were done principally to confirm our enzyme choices that were already made in the BLAST search phase, which *are E. coli* laccase (P36649) and *Synechocystis* laccase (MEB3229074.1) (Figure 14).

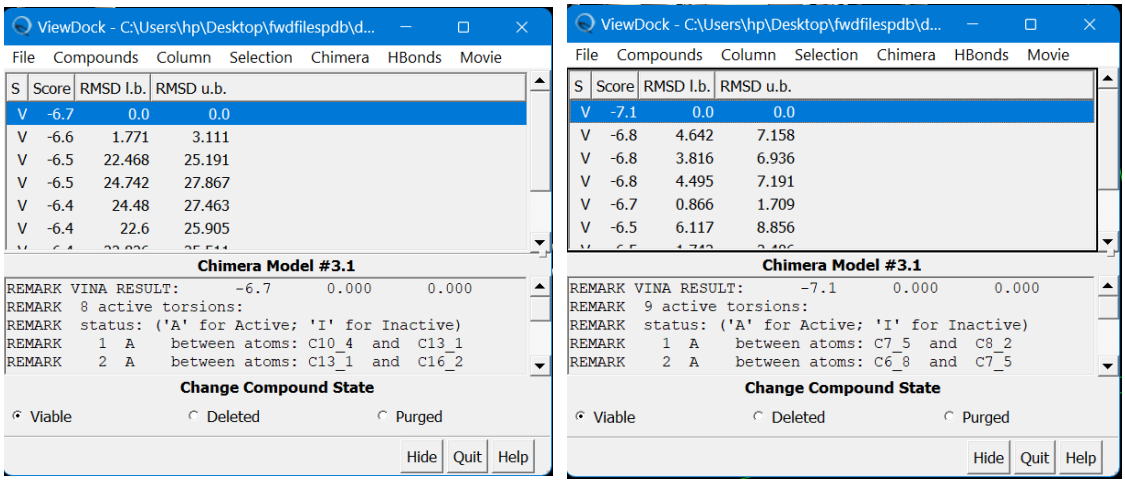

Figure 14. Docking results for *E. coli* laccase with PFOA on the left and PFOS on the right

5.3. Discussion

In the end, we managed to identify all four enzyme candidates: two dehalogenases and two laccases. These enzymes are needed in the laboratory for their surface expression in *E. coli* cells. These enzymes were searched inside

microorganisms that exhibited a better survival rate or some degradation ability in PFAS-contaminated environments, meaning they could have achieved possible adaptation to them. The idea of the surface expression comes from the fact that channels for PFAS internalisation are still unknown, as the one for externalisation after degradation. Some theories indicate an easier PFAS internalisation in plants due to the entrance from aquaporins, but it is even more unknown for microorganisms. Also, another problem with internalisation is the eventual accumulation of fluoride inside the cell, related to PFAS defluorination, because high fluoride concentrations are dangerous to the cell metabolism and can lead to its death.

For dehalogenases, we decided to maintain one out of the five dehalogenases from *D. acidovorans*. In docking simulations, DeHa2 (WP_011137954.1) obtained the highest score. This bioinformatically confirmed previous evidence that showed DeHa2 to be the most efficient among all other dehalogenases from *D. acidovorans* and DeHa2 from *D. acidovorans* was selected as our first dehalogenase candidate. The second dehalogenase was searched with BLAST to find orthologues that could be suited for our purpose. In particular, one dehalogenase from *Synechocystis* (WP_010872272.1) resulted as the best hit from the BLAST multiple alignment research as an orthologue of DeHa4 from *D. acidovorans*. The same dehalogenase also obtained the highest scores in docking simulations among all the orthologues found. For this reason, WP_010872272.1 is the second dehalogenase that was chosen.

The path was less linear for the two laccases, as we found some problems along the way. The main problem with laccases is that they are still largely uncharacterized; in addition, they present a high degree of sequence and structural divergence. For this reason, our preliminary BLAST searches, performed in many bacterial families presented many difficulties. The sequences appeared to be very different from each other, presenting gaps or additional segments even in regions that should have been conserved. This made clustering impossible. We started again from an *E. coli* laccase (P36649) that had been used in other iGEM projects for various bioremediation purposes, and we decided to test it for PFAS bioremediation as well. So, we selected it as our first laccase candidate. We then tried to obtain some orthologues in *Synechocystis* to use one as the second candidate, given that this organism is PFAS resistant, and it was used for preliminary PFAS bioremediation.

We opted for a *Synechocystis* laccase (MEB3229074.1) which obtained good scores for all the BLASTs we performed. Another problem was that from InterProScan, it seemed that *E. coli* was the only one with all three CuRo domains, while the *Synechocystis* laccase was missing CuRo_2. In reality, this problem might be due to the information from the databases.

6. Conclusions

PFAS are a problem of increasing concern, due to the continuous discoveries of health consequences, their increasing accumulation, and the lack of a final solution for degradation. Our MUTANS group decided to contribute to this field of research and to present our results at the iGEM international competition of synthetic biology. The competition requires choosing a local issue, and PFAS contamination in the Veneto region is one of the worst cases ever. Additionally, iGEM encourages participants to consider the human perspective and engage with those directly affected by the problem. For some months now, we have been in contact with local people and experts who have worked with them for years. This has helped us immerse ourselves in and understand the topic better, increasing our commitment to the project's success.

The main objective of our project, called *surPFAS*, is to combine populations of *E. coli* expressing laccase and dehalogenase enzymes on their surface. In the research field, the two have never been combined, but we believe that, due to their individual properties, a system presenting both might represent an innovative idea and lead to new interesting results. We also understand that, for this very reason, we may need to reconsider some steps of the project, such as the enzyme choice. In the future, we might discover that new orthologues from other microorganisms or completely new classes of enzymes could be more suitable. For example, we had some doubts about certain aspects of laccases and considered peroxidases as a potential alternative, since they are another group of enzymes widely used in bioremediation.

7. Bibliography

Berhanu, A., Mutanda, I., Taolin, J., Qaria, M. A., Yang, B., & Zhu, D. (2023). A review of microbial degradation of per- and polyfluoroalkyl substances (PFAS): Biotransformation routes and enzymes. *Science of The Total Environment*, *859*, 160010. https://doi.org/10.1016/j.scitotenv.2022.160010

Giglioli, S., Colombo, L., & Azzellino, A. (2023). Cluster and multivariate analysis to study the diffuse contamination of emerging per- and polyfluoroalkyl substances (PFAS) in the Veneto Region plain (North-eastern Italy). *Chemosphere*, *319*, 137916. https://doi.org/10.1016/j.chemosphere.2023.137916

Harris, J. D., Coon, C. M., Doherty, M. E., McHugh, E. A., Warner, M. C., Walters, C. L., Orahood, O. M., Loesch, A. E., Hatfield, D. C., Sitko, J. C., Almand, E. A., & Steel, J. J. (2022). Engineering and characterization of dehalogenase enzymes from Delftia acidovorans in bioremediation of perfluorinated compounds. *Synthetic and Systems Biotechnology*, *7*(2), 671–676. https://doi.org/10.1016/j.synbio.2022.02.005

Luo, Q., Liang, S., & Huang, Q. (2018). Laccase induced degradation of perfluorooctanoic acid in a soil slurry. *Journal of Hazardous Materials*, *359*, 241–247. https://doi.org/10.1016/j.jhazmat.2018.07.048

Luo, Q., Lu, J., Zhang, H., Wang, Z., Feng, M., Chiang, S.-Y. D., Woodward, D., & Huang, Q. (2015). Laccase-Catalyzed Degradation of Perfluorooctanoic Acid. *Environmental Science & Technology Letters*, *2*(7), 198–203. https://doi.org/10.1021/acs.estlett.5b00119

Marchetto, F., Roverso, M., Righetti, D., Bogialli, S., Filippini, F., Bergantino, E., & Sforza, E. (2021). Bioremediation of Per- and Poly-Fluoroalkyl Substances (PFAS) by Synechocystis sp. PCC 6803: A Chassis for a Synthetic Biology Approach. *Life*, *11*(12). https://doi.org/10.3390/life11121300

Martin, E., Dubessay, P., Record, E., Audonnet, F., & Michaud, P. (2024). Recent advances in laccase activity assays: A crucial challenge for applications on complex substrates. *Enzyme and Microbial Technology*, *173*, 110373. https://doi.org/10.1016/j.enzmictec.2023.110373

Meegoda, J. N., Bezerra de Souza, B., Casarini, M. M., & Kewalramani, J. A. (2022). A Review of PFAS Destruction Technologies. *International Journal of Environmental Research and Public Health*, *19*(24). https://doi.org/10.3390/ijerph192416397

Scott, C., & Miao, H. (2024). Toward the development of a molecular toolkit for the microbial remediation of per-and polyfluoroalkyl substances. *Applied and Environmental Microbiology*, *90*(4), e00157-24. https://doi.org/10.1128/aem.00157-24

Steel, J. J. (2021). *USAFA iGEM 2021—Results!* iGEM. https://2021.igem.org/Team:USAFA/Results

Vascon, F. (2019). *A Synthetic Biology Approach to PFAS Bioremediation: Preliminary Computational and Molecular Studies on Fluoroacetate Dehalogenase* [Tesi di Laurea Magistrale in Biotecnologie Industriali]. Università degli Studi di Padova.

Wackett, L. P. (2021). Why Is the Biodegradation of Polyfluorinated Compounds So Rare? *mSphere*, *6*(5), 10.1128/msphere.00721-21. https://doi.org/10.1128/msphere.00721-21

Zhang, Z., Sarkar, D., Biswas, J. K., & Datta, R. (2022). Biodegradation of per- and polyfluoroalkyl substances (PFAS): A review. *Bioresource Technology*, *344*, 126223. https://doi.org/10.1016/j.biortech.2021.126223

8. Appendix

Appendix 1. BLAST results using D. acidovorans dehalogenases inside E. coli

DeHa1	EFA2408061.1
	WP_001213584.1
DeHa2	WP_097755876.1
	WP_097755881.1
	EIQ7173218.1
	HDH9217445.1
	MCF1957348.1
DeHa4	EFM9948120.1
	MCV2963353.1
	MDI4533050.1
DeHa5	HAI9593615.1

Appendix 2. Lists of dehalogenase orthologues in Synechocystis used for multiBLAST research

25–39% to Burkholderia sp. Fac-dex	
WP_009633998.1	WP_038531141.1
WP_009634094.1	AIE75980.1
WP_028949175.1	WP_010872160.1
WP_010872272.1	QHV01713.1
WP_028946933.1	WP_009630128.1
WP_028954141.1	WP_010873707.1
WP_010871769.1	WP_028948001.1
WP_106907929.1	WP_038021267.1
WP_009634549.1	QHV01001.1
WP_009631931.1	

22–40% to Burkholderia sp. Fac-dex					
WP_194019009.1	WP_028947035.1	WP_010871750.1	WP_194072888.1	WP_199303557.1	WP_080504060.1
WP_194072249.1	MBE9175246.1	WP_010871770.1	WP_009630376.1	WP_194018983.1	WP_162329611.1
WP_190599725.1	WP_009631087.1	WP_194020945.1	WP_162327789.1	WP_190599278.1	WP_038530520.1
WP_162327898.1	WP_190598236.1	WP_194017334.1	WP_194018987.1	WP_194013777.1	

WP_194013416.1	WP_194019539.1	WP_194015193.1	WP_194014523.1	WP_194019704.1	
WP_190598114.1	WP_162327488.1	WP_028947752.1	WP_199303425.1	MBE9175978.1	
WP_194020943.1	WP_194073349.1	MBE9176742.1	WP_190599423.1	WP_028946321.1	
WP_194017332.1	WP_162327505.1	WP_194020915.1	MBE9176468.1	WP_199309777.1	
WP_162327504.1	WP_009630182.1	MBD2655166.1	WP_194072577.1	WP_194014536.1	
WP_038020775.1	WP_190598115.1	WP_190599431.1	WP_010872152.1	WP_162327783.1	

Appendix 3. Dehalogenases resulting from the multiBLAST using D. acidovorans enzymes

DeHa4	UPI0002ABF79E
	WP_194019009.1
	WP_194072249.1
	WP_190599725.1
	WP_194013416.1
	UPI00000C10BF
	WP_162327898.1
	UPI000411AC78
	WP_162327789.1
	UPI00000D34C4
	UPI0002ACC833

Appendix 4. Dehalogenases resulting from the multiBLAST using FAc-DEX from Burkholderia

FAc-DEX FA1	UPI0002ABF79E
	UPI0002ACC833
	WP_028947035.1
	MBE9175246.1
	WP_162327488.1
	WP_010871750.1
	WP_162327505.1
	WP_194020915.1
	MBD2655166.1