EXTREMOPHILES AS MICROBIAL TOOLS TO FACE ENVIRONMENTAL PROBLEMS ASSOCIATED TO ANTHROPOGENIC ACTIVITIES

Miriam Carbonaro

Dottorato in Biotecnologie –XXXVI ciclo

Università di Napoli Federico II

Dottorato in Biotecnologie – XXXVI ciclo

Università di Napoli Federico II

EXTREMOPHILES AS MICROBIAL TOOLS TO FACE ENVIRONMENTAL PROBLEMS ASSOCIATED TO ANTHROPOGENIC ACTIVITIES

Miriam Carbonaro

Dottorando: Miriam Carbonaro

Relatore: Prof. Gabriella Fiorentino

Coordinatore: Prof. Marco Moracci

Settore Scientifico Disciplinare Bio10

Alla mia famiglia e a tutti coloro che mi hanno sostenuto in questo percorso

"What's the use of doing all this work if we don't get some fun out of this?"

Rosalind Elsie Franklin

INDEX

RIASSUNTO

La costruzione di un'economia circolare, definita come "un insieme di modelli sostenibili di produzione e consumo" dall'Agenda 2030 delle Nazioni Unite, emerge come una priorità per la comunità scientifica globale. Negli ultimi anni, la ricerca si è concentrata sulla creazione di nuovi approcci biotecnologici volti a monitorare o ridurre le sostanze inquinanti, e a ottimizzare processi industriali sostenibili. Nel contesto della lotta all'inquinamento e dell'utilizzo dei materiali di scarto per la produzione di risorse rinnovabili e di prodotti di valore aggiunto, gli studi sulla fisiologia, ecologia e biochimica dei microorganismi presenti in vari ecosistemi sono fondamentali.

In particolare, in questo lavoro di tesi intitolato "*Extremophiles as microbial tools to face environmental problems associated to anthropogenic activities*" l'attenzione si è maggiormente focalizzata sullo studio delle potenzialità di un gruppo di microrganismi estremofili, i termofili, e dei loro enzimi, i termozimi. I microrganismi termofili prosperano ad alte temperature e molto spesso dimostrano anche una capacità di sopravvivenza in ambienti caratterizzati da elevate concentrazioni saline, valori di pH estremi e presenza di sostanze tossiche. La prima parte di questa tesi (**Capitolo 2**) è un lavoro di sinossi sulle caratteristiche uniche dei microrganismi estremofili, in particolare i termofili, e dei loro enzimi, evidenziandone l'interesse significativo nel contesto biotecnologico. Sono state esplorate le diverse applicazioni di tali microrganismi in vari settori, con particolare rilevanza nell'ambito industriale e ambientale. La straordinaria capacità di adattamento di questi microrganismi nel crescere su scarti e/o inquinanti li rende strumenti importanti non solo nella tutela ambientale, ad esempio nella gestione dell'inquinamento da metalli pesanti e composti organici, ma anche nella valorizzazione delle biomasse lignocellulosiche per la produzione di prodotti ad alto valore aggiunto. In questa ottica diventa essenziale disporre di strumenti adeguati all'ingegnerizzazione di tali microrganismi al fine di massimizzarne le prestazioni. Quindi, si è cercato di delineare quali sono stati i recenti sviluppi tecnologici nell'ambito dell'ingegneria genomica applicata ai microorganismi termofili.

La fase sperimentale iniziale di questo lavoro di tesi (**Capitolo 3**) si è inserita in un progetto di ricerca che mirava al completamento della caratterizzazione del sistema di resistenza all'arsenico del batterio Gram-negativo *Thermus thermophilus* HB27. *Tt*ArsM, una arsenito metiltransferasi SAM-dipendente, è stata identificata come nuovo membro di questo sistema di resistenza attraverso un approccio di *pull-down* utilizzando *Tt*SmtB (il repressore trascrizionale del sistema)

come esca, mettendo in luce la loro interazione *in vitro*. Le interazioni proteina-proteina tra *Tt*SmtB e *Tt*ArsM sono state confermate da esperimenti di co-immunoprecipitazione e saggi EMSA, evidenziando la formazione di complessi multimerici che potrebbero influire sul controllo trascrizionale o post-trascrizionale. Ulteriori analisi EMSA hanno suggerito un meccanismo regolatorio della trascrizione di *Tt*ArsM da parte di *Tt*SmtB analogo a quello osservato per altri geni di resistenza all'arsenico già caratterizzati nello stesso microorganismo. Il ruolo di *Tt*ArsM nella resistenza all'arsenico è stato poi dimostrato *in vivo* in cellule di *E. coli* che esprimevano *Tt*ArsM in maniera ricombinante. L'enzima, che si contraddistingue per la presenza di una singola cisteina nella sua sequenza amminoacidica, mostra un'attività arsenito metiltransferasica specifica e produce come prodotti primari arsenico mono- e di-metilato. Analisi cromatografiche di esclusione molecolare hanno poi dimostrato come l'interazione di *Tt*ArsM con il substrato arsenito sia in grado di influenzarne la struttura quaternaria, da omodimero a monomero. Per valutare il ruolo di questo enzima *in vivo* è stata sviluppata una strategia di editing genomico con tecnologia CRISPR-Cas9 direttamente in *T. thermophilus* HB27. In questo modo sono stati prodotti tre ceppi mutanti; **i)** un ceppo knock-out per il gene *Tt*arsM, **ii)** un ceppo knock-out del gene *Tt*arsX (il trasportatore di membrana del sistema di resistenza) al cui posto è stato inserito il gene codificante per la proteina fluorescente gialla termotollerante, sYFP, **iii)** un ceppo doppio knock-out per i geni *TtarsM* e *TtarsX* e knock-in per il gene *syfp*. Il ceppo *ΔTtarsM-ΔTtarsX (syfp)* si è rivelato il meno resistente agli ioni As(III) e As(V), ed in grado di emettere fluorescenza a basse concentrazioni di As(III) e As(V), confermando il suo notevole potenziale come *whole-cell bioreporter*. In conclusione, il lavoro fornisce dettagli sulle proprietà biochimiche di *Tt*ArsM, contribuendo alla comprensione del sistema di resistenza all'arsenico in *T. thermophilus* HB27 e fornendo uno strumento promettente da implementare per applicazioni nel biomonitoraggio di arsenico nell'ambiente.

Nella sezione successiva di questa tesi (**Capitolo 4**), sono stati effettuati studi su *Alicyclobacillus mali* FL18, un batterio termo-acidofilo Gram-positivo isolato a Pisciarelli, un'area idrotermale della Solfatara di Pozzuoli vicino Napoli (Italia) resistente ad alcuni metalli pesanti e a determinati antibiotici. Durante questo lavoro di tesi è stato indagato se questo batterio possedesse enzimi in grado di degradare materiali lignocellulosici. In particolare, l'analisi genomica di *A. mali* FL18 ha identificato 73 geni codificanti enzimi potenzialmente attivi sui

carboidrati (CAZymi), corrispondenti a circa il 2,5% del totale dei geni codificanti proteine. Le glicoside idrolasi (GH) costituiscono la famiglia più numerosa con 32 enzimi annotati, i quali appaiono prevalentemente coinvolti nella decomposizione di specifici componenti della lignocellulosa. In particolare, è stato caratterizzato un membro della famiglia 9 delle glicoside idrolasi (GH9), chiamato AmCel9, ed è stato messo in luce il suo ruolo significativo nel processo di idrolisi della cellulosa. In generale, i membri di questa famiglia presentano una struttura quaternaria altamente eterogenea, infatti, recenti analisi metagenomiche hanno rivelato 19 moduli diversi nelle GH9, caratteristica altamente correlata alla loro multifunzionalità. Nello specifico, l'analisi strutturale *in silico* dell'enzima AmCel9 ha mostrato la tipica triade catalitica della famiglia e un dominio Ig-like N-terminale, comune anche nella struttura di altre GH9. AmCel9 purificata (prodotta in modo ricombinante in cellule di *E. coli*) mostra una notevole resistenza ad un ampio spettro di valori di pH, differenziandosi significativamente da altri membri della famiglia GH9. Questo enzima, inoltre, presenta valori molto simili di attività specifica tra i 40 e i 70 °C e mantiene il 60% di attività a 50 °C fino a 16 ore. In termini di specificità del substrato e proprietà catalitiche, AmCel9 è in grado di idrolizzare efficacemente la carbossilmetilcellulosa (CMC) e il lichenano con alti valori di k_{cat} e di attività specifica. Questo enzima inoltre presenta una notevole attività su substrati di cellulosa insolubile, come Avicel PH101 e carta da filtro Whatman®, caratteristica non comunemente riscontrata in numerosi membri di questa famiglia enzimatica. Queste caratteristiche rendono AmCel9 un biocatalizzatore multifunzionale molto promettente da testare per la degradazione di materiali lignocellulosici senza l'esecuzione di pretrattamenti dannosi per l'ambiente.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

SUMMARY

This thesis, entitled "*Extremophiles as microbial tools to face environmental problems associated with anthropogenic activities*", aims to demonstrate the feasibility of sustainable solutions in biotechnology based on the use of extremophiles, with a particular focus on thermophiles and their enzymes. The study of these extraordinary microorganisms can provide new tools for many applications among which environmental issues and valorization of renewable resources. **Chapter 2** offers a comprehensive overview of these microorganisms emphasizing their importance in various biotechnological fields. Their adaptability to thrive in pollutants and biomass wastes makes them ideal for environmental preservation, including heavy metal remediation and biomass valorization. The potential for engineering these microorganisms for improved performance is also discussed.

Chapter 3 focuses on the identification of *Tt*ArsM, a SAM-dependent arsenite methyltransferase with a role in the arsenic resistance system of the Gram-negative bacterium *Thermus thermophilus* HB27. *In vivo* studies on enzyme function have been possible thanks to the set-up of CRISPR-Cas9 genome editing technology for *T. thermophilus*; the tool has also been used to develop a whole-cell arsenic bioreporter system. In the following section (**Chapter 4**) *Alicyclobacillus mali* FL18, a Gram-positive thermo-acidophilic bacterium, is examined with an emphasis on its capacity to produce lignocellulolytic enzymes. Genomic analysis reveals genes for 73 putative carbohydrate-active enzymes (CAZymes), including the multifunctional GH9 family member, AmCel9, exhibiting exceptional cellulolytic activity. In the second section of **Chapter 4**, the study delves into *Geobacillus stearothermophilus* GF16, a Gram-positive thermophilic bacterium, and the ability of its secretome to degrade lignocellulose. The bacterium, grown on minimal medium enriched with xylan, has significant xylanolytic activity. The secretome of *G. stearothermophilus* GF16 effectively hydrolyzes various agri-food residues, highlighting its potential for an eco-friendly lignocellulose degradation strategy. The concluding **Chapter 5** focuses on three glucuronoyl esterases (GEs). These enzymes, belonging to CE15 family, have been characterised from the endophyte *Dyadobacter fermentans* NS114^T and play a crucial role in cleaving ester bonds within lignin-carbohydrate complexes. Their diverse catalytic efficiencies and structural features have been explored, with the goal of deepening their potential for overcoming recalcitrance in lignocellulosic waste utilization.

Chapter 1 – General introduction

The Earth is in a biological and geochemical equilibrium where environmental resources and biosphere elements are replenished through large-scale biochemical cycles. Since the beginning of time, humans have modified the environment in which they live through agriculture, farming, urbanization, and the use of natural resources. However, since the Industrial Revolution in the 19th century, and especially in the last 30 years, advances in chemistry, agriculture, medicine, and massive industrial production have disturbed the ecological balance to an unprecedented degree [1]. Anthropogenic emission of CO_2, depletion of non-renewable resources, production of non-recyclable wastes left in the environment and release of toxic compounds into groundwater are only some aspects of the failure of the current economic system [2].
In addition, the COVID-19 crisis has devastated economies and societies, reducing the likelihood of achieving adequate environmentally sustainable growth in line with the 2030 Agenda and the Paris Climate Agreement [3]. For this reason, an important goal for the scientific community is to find alternatives to classical industrial processes to face the depletion of non-renewable resources and environmental pollution. In fact, the sustainability of environmental resources is one of the fundamental points of the bioeconomy, which aims to reform current processes towards an environmentally friendly circular economy strategy [4].

Importance of extremophiles: focus on thermophilic features
In the last decades, a good part of scientific research has been focused on the use of microbial enzymes and microbial metabolic pathways for the development of biomonitoring and bioremediation strategies to solve environmental problems and more recently for the enhancement of biorefineries designed to convert biomass into chemical compounds. Harsh conditions, which are not optimal for the growth of microorganisms and their enzymes, usually prevail in some steps of these industrial bioprocesses. From this perspective, microbes inhabiting habitats considered inhospitable from the human point of view have some advantages over their mesophilic counterparts [5]; in fact, many bacteria or archaea can tolerate extreme pH values, high or low temperatures, high concentrations of salts, toxic substances and/or heavy metals. These microorganisms, called extremophiles, are valuable sources of biomolecules and bioprocesses for many biotechnological applications in agriculture,

chemistry, biomedicine, etc [6]. Depending on the temperature range in which they grow, extremophiles can be classified as: psychrophiles (-15° to 20°C), mesophiles (20-45°C), thermophiles (45-80°C) and hyperthermophiles (≥80°C) [7]. They populate volcanic regions, hot springs, deep-sea hydrothermal vents, geothermal areas, but also compost piles and waste treatment plants [8]; studies on this group of microorganisms began in the 1960s with the isolation of *Thermus aquaticus* [9] and have been gradually intensified until today.
Unique thermophilic features include smaller genome size, reverse gyrase for DNA protection, chaperones/chaperonins for macromolecular folding, and saturated fatty acids maintaining cell membrane rigidity at high temperature [10]. Due to the temperature range in which they work optimally, many industrial bioprocesses utilize these microbes for their numerous advantages over mesophilic counterparts, such as faster reaction rates, increased solubility of substrates, reduced viscosity, and elimination of microbial contamination [11]. They are also a source of heat-stable catalysts, called thermozymes, which can also show stability and activity under extreme conditions such as high pressure, salt presence, low pH, and stability in detergents and organic solvents [12]. However, although they are ideal platforms for expressing proteins/enzymes or metabolic pathways at high temperatures under harsh and toxic conditions, the practical use of thermophiles/hyperthermophiles could be hindered by difficulties in culturing and genome manipulation [13,14]. It is still needed to deepen the studies on these microorganisms to fully exploit their potential in various biotechnological fields. This was the premise of the work carried out in this thesis project.

Heavy metals and bioreporter strategies

Nowadays, the presence of high levels of heavy metals in the environment is a worldwide problem. In ecotoxicology, the term "heavy metals" refers to metals with relatively high densities, atomic weights, or atomic numbers that are toxic at low concentrations; they occur naturally in soils, rocks, sediments, air, and water, and many metals (iron, copper, manganese, and zinc) are essential for the metabolism of living species [15,16]. Nonetheless, uncontrolled anthropic activities (mining, smelting, waste incineration, fossil fuel combustion, etc.) have severely altered the natural biogeochemical cycles of heavy metals, causing various types of consequences on the ecosystem, such as biodiversity imbalance, biomass reduction, modification of microbial community structure, and reduction of enzyme activity [17]. In fact, this type of contaminant cannot be

degraded and accumulates in the biota through bioaccumulation and biomagnification processes. Therefore, humans experience the highest potential impact because they are at the apex of the food chain, making them susceptible to long-term adverse health effects and clinical manifestations, the severity of which depends on the specific metal involved [18,19]. Five heavy metals elements [arsenic (As), cadmium (Cd), chromium (Cr), mercury (Hg), and lead (Pb)] exert toxic effects at very low concentrations and represent the most widespread heavy metals, called toxic heavy metals (THMs) [20]. Among them, arsenic (As) is one of the most toxic semi-metallic elements of global concern.

Generally, traditional analytical methods for soil and water contaminant analysis, such as atomic absorption spectroscopy (AAS), gas chromatography-mass spectrometry (GC-MS), high performance liquid chromatography (HPLC), inductively coupled plasma-mass spectrometry (ICP-MS), and X-ray absorption spectroscopy have been used [21]. These traditional chemical approaches, however, are primarily used to assess the chemical availability of soil contaminants rather than their bioavailability [22]. In this context, the concept of chemical availability, does not consider critical aspects such as the uptake of contaminants by organisms, their subsequent accumulation in their tissues, and the resulting toxic effects. This limitation is in contrast with the broader concept of bioavailability, which embraces the dynamic and complex processes by which contaminants interact with the environment and the ecosystem [23]. The divergence between these concepts underscores the necessity for a more comprehensive understanding, prompting the exploration of biological approaches that address the interactive dynamics of contaminants within the environment and organisms. Hence, the use of bioreporters and biosensors, becomes essential.

Engineered whole-cell bioreporters are microbial cells, genetically modified able to detect and signal the presence of specific contaminants. A whole-cell bioreporter typically consists of several key components that enable it to function as a biosensor for specific environmental stimuli; these components provide a rapid and sensitive tool of monitoring contaminants in environmental samples and include **i)** a Sensor Element, **ii)** a Genetic Logic Circuit, **iii)** a Reporter Gene, and **iv)** an Actuator (Figure 1) [24].

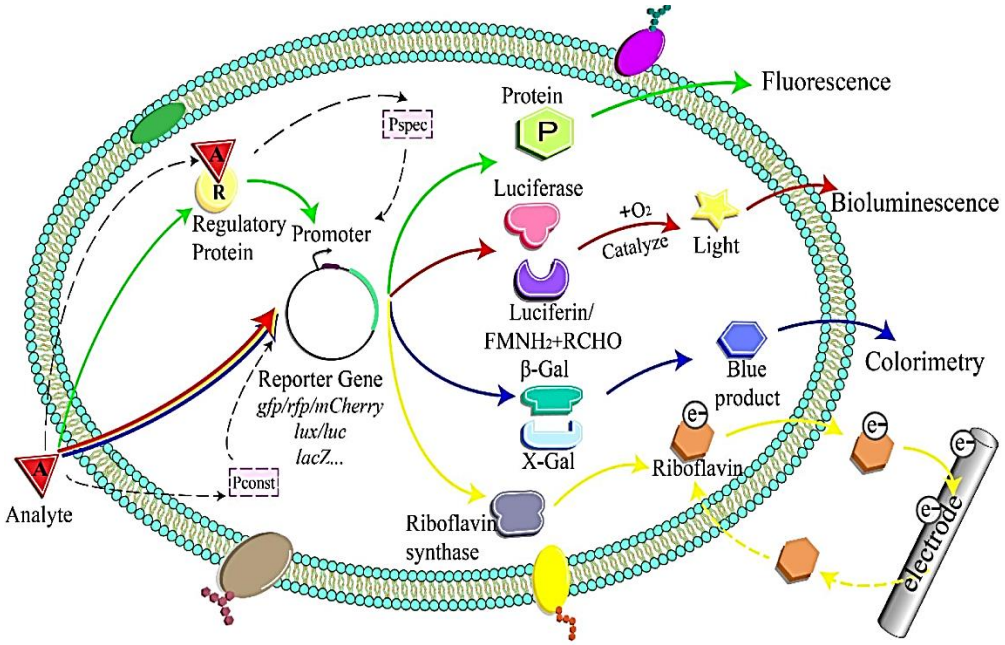

Figure 1. Whole-cell microbial bioreporter-based detection principle. Expression of the specific promoter and downstream reporter gene is activated or inhibited by the analyte [24].

This approach complements traditional chemical analysis to provide a more in-depth assessment of the impact of contaminants on the ecosystem and living organisms. For this reason, a deeper understanding of bacterial resistance pathways and mechanisms is critical to advancing the development and optimization of technologies for assessing the environmental risks associated with heavy metals soil and/or water contamination.

Arsenic resistance in Thermus thermophilus HB27

While several prokaryotes can use arsenic oxidation for energy, life could not have emerged without arsenic resistance systems, as the abundance of arsenic in the environment provided the intense selective pressure that was the driving force for the evolution of mechanisms related to arsenic detoxification and metabolism in Gram-positive and Gram-negative bacteria [25–27]. These mechanisms can classically include **i)** Arsenic Resistance (ars) System, responsible for the detoxification of inorganic arsenic and involving ars operons located in the chromosomes or plasmids of bacteria; **ii)** Arsenic Methylation and Related Pathways, organo-arsenicals are detoxified by this pathway; **iii)** Arsenite Oxidation (aio/arx) and **iv)** Arsenate Reduction (arr) Systems which participate

in the energy-producing respiratory biotransformation between As(III) and As(V) [28,29].

The Gram-negative thermophilic microorganism *Thermus thermophilus* HB27 originally isolated from a natural thermal environment in Japan by Oshima and Kazutumo has a well-characterized arsenic resistance system; it grows at temperatures ranging from 50 ° to 82 °C with an optimum of 70 °C and pH 7.0 [30], and it is resistant to arsenate and arsenite up to concentrations of 44 mM and 40 mM. Its arsenic resistance machinery is interspersed in the chromosome and is composed by an arsenate reductase (*Tt*ArsC) [31], an ATP-dependent membrane transporter (*Tt*ArsX) [32] and *Tt*SmtB, a transcriptional regulator belonging to the ArsR/SmtB family (Figure 2) [33,34].

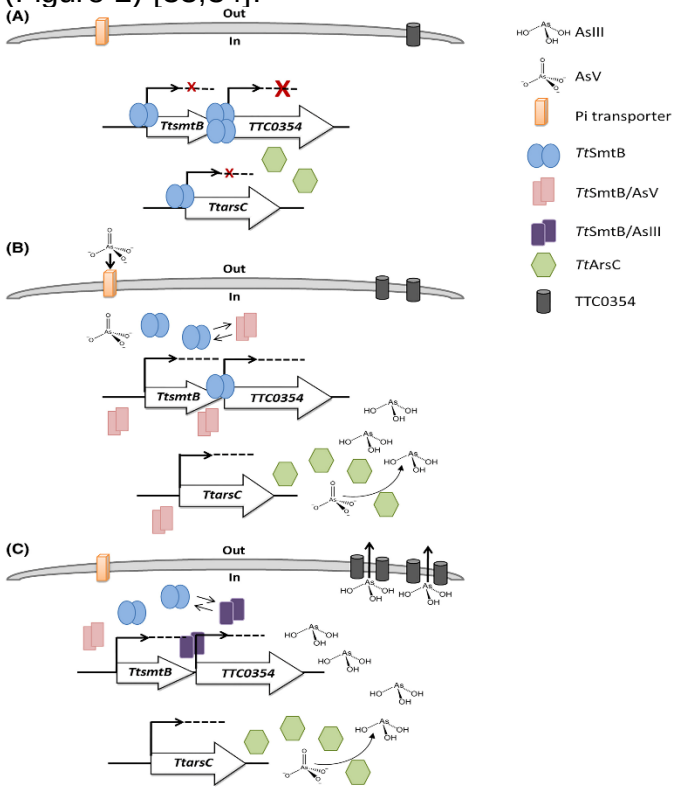

Figure 2. Scheme of arsenic resistance system in *T. thermophilus* HB27. (**A**) In the absence of arsenic, *Tt*SmtB represses gene transcription, especially *TtarsX* (*TT0354*). (**B**) Upon arsenate entry, *Tt*SmtB binds arsenic and represses itself and *TtarsC* expression, allowing arsenate to be converted to arsenite. (**C**) As arsenite accumulates, *Tt*ArsX is released and activated to expel arsenite. This process lowers intracellular arsenate, efficiently exports arsenite, and restores transcriptional repression [33].

Detailed molecular studies of the arsenic resistance system have enabled the development of tools for the detection of heavy metals. In fact, innovative biosensors capable of detecting arsenic have been obtained either using the enzyme *Tt*ArsC absorbed onto stabilized gold nanoparticles or immobilizing on gold electrodes chimeric proteins derived from the genetic fusion of the thermophilic arsenate reductase with hydrophobin from *Pleurotus ostreatus* [35–37].
T. thermophilus HB27 possesses many different genes of biotechnological interest, such as proteases/peptidases, starch/oligosaccharide-hydrolysing enzymes or esterases [38]; it also represents a huge source of interesting thermostable and thermoactive enzymes that can be exploited in environmental bioremediation of persistent organic pollutants [39,40]. These features coupled with its great metabolic versatility, high growth rates, high biomass yields, and a constitutive very efficient natural competence apparatus for the introduction of heterologous DNA [41–43], make *T. thermophilus* HB27 a good model laboratory strain for the generation of engineered strains for industrial production of miscellaneous biomolecules and the development of whole-cell bioreporter system.

Lignocellulosic materials - a renewable resource: intrinsic value and drawbacks

As mentioned above, extremophilic microorganisms are of great interest in the field of biotechnology because of their potential applications in various industrial sectors, such as energy production and the production of biopolymers from renewable sources. Among renewable resources, lignocellulosic waste materials are of particular importance [44].
These materials have enormous potential as a sustainable raw material, providing a valuable alternative to non-renewable and fossil resources. Indeed, lignocellulose is extremely abundant; it is estimated that approximately 181.5 billion tons of lignocellulosic waste is generated annually, and only 9% of biomass contributes to total energy supply, according to a recent report by the International Energy Agency (IEA) [44]. Various types of lignocellulosic wastes are generated from agricultural activities, including sources such as animal manure, post-harvest forestry residues, non-marketable products, low or unprocessed vegetable waste, and by-products from urban and industrial processing. These lignocellulosic materials can be used as feedstock for bioenergy production and to produce bio-based chemicals, materials and polymers, supporting waste valorization and carbon sequestration to offset emissions to fossil

fuels use (Figure 3) [45]. The term lignocellulose refers to a highly heterogeneous biopolymer synthesized by plant cells, consisting mainly of cellulose, hemicellulose, and lignin, which are present in varying amounts and ratios, depending on the origin of the biomass. It also contains small amount of proteins, pectins and extractives (tannins, lipids, resins, steroids, terpenes, terpenoids, flavonoids, and phenolic compounds) [44].

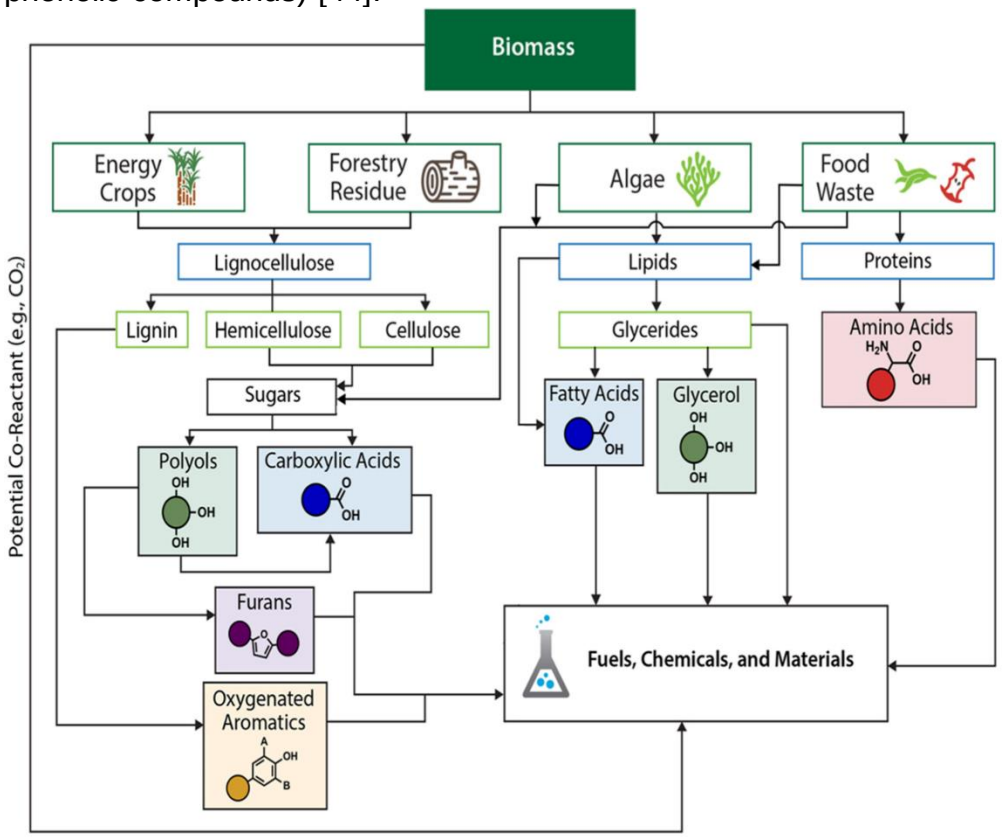

Figure 3. Many value-added products derived from the major lignocellulosic components of common biomass sources [45,46].

The main obstacle to the full utilization of this sustainable resource is its recalcitrance, which refers to its resistance to degradation to smaller molecules to be further converted into biofuels or high-value chemicals. This recalcitrance is primarily due to the complex and tightly bound structure of lignocellulose, which is difficult to access, to separate and to break down into its components. Several factors contribute to the recalcitrance of lignocellulosic materials: **i)** its structural and chemical heterogeneity which depends on the vegetal source, and also poses challenges in the development of universal conversion processes; **ii)** the lignin content and structure, in fact lignin

is a complex and rigid polymer that acts as a barrier, coating cellulose and hemicellulose and making them scarcely accessible to enzymes or chemicals for degradation; **iii)** Lignin-Carbohydrate Complexes (LCCs), covalent bonds between lignin and carbohydrates, particularly hemicellulose, further impede the accessibility of cellulose to conversion processes; **iv)** the crystalline structure of cellulose, which is very resistant to enzymatic or chemical degradation and requires additional pretreatments; **v)** the great heterogeneity of hemicellulose, characterised by different linkages and substitutions, making it even more tricky than cellulose to degrade [44].

The integrated biorefinery concepts aims to the simultaneous valorization of a lignocellulosic biomass maximizing efficiency and minimizing waste to produce a diversified range of products such as biofuels, high-value chemicals and bioproducts. Such a sustainable process involves several critical steps.

For example, physicochemical conversion processes, which include thermochemical methods such as pyrolysis and gasification, provide pathways for the production of biobased chemicals and syngas [45,46]. Biochemical conversion includes pre-treatment methods used to break down the complex lignocellulosic structure and improve the accessibility of cellulose and hemicellulose for further processing; pre-treatments may involve physical, chemical, physicochemical, or biological methods [47,48].

Subsequent enzymatic hydrolysis and fermentation strategies facilitate the conversion of these components into fermentable sugars, biofuels, and high-value chemicals. Meanwhile, lignin undergoes valorization processes that contribute to the production of other valuable chemicals and materials (Figure 3) [48,49].

To improve the efficiency of biofuel and high-value chemical production, novel advances in biotechnology, enzyme engineering, and process optimization are required, including ones that address the recalcitrance of lignocellulosic materials.

Synergistic enzymatic systems for lignocellulose degradation and the advantages of thermophilic microorganisms

The wide variety of lignocellulose components requires the synergistic action of many enzymes with multiple and different activities. In nature several enzymatic systems for lignocellulose degradation have been found, and the Carbohydrate-Active enZymes database (the CAZy database, http://www.cazy.org/) [50] provides a basic classification of these biocatalysts. Ligninolytic enzymes such as laccases, manganese-dependent peroxidases (MnP), lignin peroxidases (LiP),

versatile peroxidases (VP) and aryl alcohol oxidases cause lignin modification and consumption. The hydrolytic enzymes responsible for holocellulose degradation are a group of enzymes generally referred to as cellulases and hemicellulases, but oxidative enzymatic system, esterases and pectinases are also involved. Cellulases break down cellulose, a major component of lignocellulose, into glucose and other simple sugars. Cellulases include endoglucanases, exoglucanases, and β-glucosidases. Hemicellulases, including xylanases, mannanases, and arabinofuranosidases, degrade hemicellulose [51].

Lytic Polysaccharide Monooxygenases (LPMOs) are enzymes capable of oxidative cleavage of cellulose and other polysaccharides using oxygen or hydrogen peroxide as a co-substrate and a reducing agent as an electron donor. Esterases cleave ester bonds in lignocellulose, releasing acetic acid and other organic acids. Esterases are important for the efficient breakdown of hemicellulose and are a relatively new class of enzymes that have been shown to play a crucial role in lignocellulose degradation [52].

The degradation of lignocellulose presents challenges, including temperature and solvents requirements, and the production of inhibitors and toxic compounds. These factors can inhibit enzyme activity, reduce degradation efficiency, and pose risks to microorganisms' viability necessary for the fermentation process. Also in this context, the use of thermophiles or thermophilic enzymes offers some advantages over other treatments, including higher efficiency, faster process, reduced inhibitory effects and great specificity [52,53].

Thermophiles are also beneficial in the efficient breakdown of organic matter, as observed in composting processes, where they contribute to pathogen elimination and accelerated decomposition, due to the higher temperatures that can be used. These capabilities highlight the environmental benefits of using thermophiles and thermozymes in lignocellulosic degradation. Their ability to operate at elevated temperatures reduces the reliance on chemical treatments and harsh conditions, thereby mitigating negative environmental impacts [11,54].

Overall, the use of thermophilic microorganisms and their enzymes is an attractive choice for sustainable and environmentally friendly applications in lignocellulose degradation.

Aim of the work

The overall aim of this project has been focused on the exploitation of extremophiles and their enzymes for the set-up of biomonitoring and lignocellulose valorization strategies. The first goal was to develop whole-cell arsenic bioreporters and the second one was to utilize new thermophilic strains and/or their carbohydrate-active enzymes for the processing of raw materials into high value-added products.

In **Chapter 2**, the comprehensive Review Article documents updated literature regarding the peculiarities of extremophiles and their biocatalysts as well as main applications in biotechnology.

In **Chapter 3**, the Paper 1 deepens the comprehension of the arsenic resistance mechanisms of *T. thermophilus* HB27 and describes a strategy that led to the identification of a novel member within this system. The extensive study of the molecular determinants that regulate the resistance mechanism along with the set-up of a novel CRISPR-Cas9 genome editing tool for the bacterium allowed the development of a whole-cell bioreporter strain able to detect micromolar concentrations of arsenic ions.

The second objective of this thesis is described in the **Chapter 4** and **Chapter 5**. One published paper and one manuscript in preparation examine the lignocellulose degradation activities of two recently identified strains and are included in **Chapter 4**:

1. Insight into CAZymes of *Alicyclobacillus mali* FL18: Characterization of a New Multifunctional GH9 Enzyme (Paper 2-I)
2. Genomic mining of *Geobacillus stearothermophilus* GF16 for xylose production from hemicellulose-rich biomasses using secreted enzymes (Manuscript 2-II).

The concluding **Chapter 5** contains the submitted Paper 3 that is focused on another under explored class of enzymes, the Glucuronoyl Esterases (GEs), considered extremely important to overcome the recalcitrance of lignocellulosic raw material; in particular, the work explores the structural-functional relationship in three different members of the CE15 family highlighting peculiarities of the active site, differences in substrate specificity and giving new insights in the relevance of this class for the degradation of lignocellulosic feedstocks .

References:

1. de Lorenzo, V.; Marlière, P.; Solé, R. Bioremediation at a Global Scale: From the Test Tube to Planet Earth. *Microb Biotechnol* **2016**, doi:10.1111/1751-7915.12399.
2. Barnosky, A.D.; Hadly, E.A.; Bascompte, J.; Berlow, E.L.; Brown, J.H.; Fortelius, M.; Getz, W.M.; Harte, J.; Hastings, A.; Marquet, P.A.; et al. Approaching a State Shift in Earth's Biosphere. *Nature* 2012.
3. Rogelj, J.; Den Elzen, M.; Höhne, N.; Fransen, T.; Fekete, H.; Winkler, H.; Schaeffer, R.; Sha, F.; Riahi, K.; Meinshausen, M. Paris Agreement Climate Proposals Need a Boost to Keep Warming Well below 2 °c. *Nature* 2016.
4. Sarsaiya, S.; Jain, A.; Kumar Awasthi, S.; Duan, Y.; Kumar Awasthi, M.; Shi, J. Microbial Dynamics for Lignocellulosic Waste Bioconversion and Its Importance with Modern Circular Economy, Challenges and Future Perspectives. *Bioresour Technol* 2019.
5. Podar, M.; Reysenbach, A.L. New Opportunities Revealed by Biotechnological Explorations of Extremophiles. *Curr Opin Biotechnol* 2006, *17*, 250–255.
6. Dumorné, K.; Córdova, D.C.; Astorga-Eló, M.; Renganathan, P. Extremozymes: A Potential Source for Industrial Applications. *J Microbiol Biotechnol* 2017.
7. Stetter, K.O. Hyperthermophilic Procaryotes. *FEMS Microbiol Rev* **1996**, *18*, 149–158, doi:10.1111/j.1574-6976.1996.tb00233.x.
8. Mehta, D.; Satyanarayana, T. Diversity of Hot Environments and Thermophilic Microbes. In *Thermophilic Microbes in Environmental and Industrial Biotechnology: Biotechnology of Thermophiles*; 2013 ISBN 9789400758995.
9. Brock, T.D.; Freeze, H. Thermus Aquaticus Gen. n. and Sp. n., a Nonsporulating Extreme Thermophile. *J Bacteriol* **1969**, doi:10.1128/jb.98.1.289-297.1969.
10. Gomes, J.; Steiner, W. The Biocatalytic Potential of Extremophiles and Extremozymes. *Food Technol Biotechnol* 2004.
11. Bruins, M.E.; Janssen, A.E.M.; Boom, R.M. Thermozymes and Their Applications: A Review of Recent Literature and Patents. *Applied Biochemistry and Biotechnology - Part A Enzyme Engineering and Biotechnology* **2001**, doi:10.1385/ABAB:90:2:155.
12. Atomi, H. Recent Progress towards the Application of Hyperthermophiles and Their Enzymes. *Curr Opin Chem Biol* 2005.
13. Goodenough, P.W.; Jenkins, J.A. Protein Engineering to Change Thermal Stability for Food Enzymes. In Proceedings of the Biochemical Society Transactions; 1991.
14. Elleuche, S.; Schäfers, C.; Blank, S.; Schröder, C.; Antranikian, G. Exploration of Extremophiles for High Temperature Biotechnological Processes. *Curr Opin Microbiol* 2015.
15. Nies, D.H. Microbial Heavy-Metal Resistance. *Appl Microbiol Biotechnol* 1999.
16. Gallo, G.; Puopolo, R.; Limauro, D.; Bartolucci, S.; Fiorentino, G. Metal-Tolerant Thermophiles: From the Analysis of Resistance Mechanisms to Their Biotechnological Exploitation. *Open Biochem J* **2018**, doi:10.2174/1874091x01812010149.
17. Vitousek, P.M.; Mooney, H.A.; Lubchenco, J.; Melillo, J.M. Human Domination of Earth's Ecosystems. *Science (1979)* **1997**, doi:10.1126/science.277.5325.494.

18. Singh, R.; Gautam, N.; Mishra, A.; Gupta, R. Heavy Metals and Living Systems: An Overview. *Indian J Pharmacol* 2011.
19. Pérez-De-Mora, A.; Burgos, P.; Madejón, E.; Cabrera, F.; Jaeckel, P.; Schloter, M. Microbial Community Structure and Function in a Soil Contaminated by Heavy Metals: Effects of Plant Growth and Different Amendments. *Soil Biol Biochem* **2006**, doi:10.1016/j.soilbio.2005.05.010.
20. Clemens, S.; Ma, J.F. Toxic Heavy Metal and Metalloid Accumulation in Crop Plants and Foods. *Annu Rev Plant Biol* 2016.
21. Brányiková, I.; Lucáková, S.; Kuncová, G.; Trögl, J.; Synek, V.; Rohovec, J.; Navrátil, T. Estimation of Hg (II) in Soil Samples by Bioluminescent Bacterial Bioreporter E. Coli ARL1, and the Effect of Humic Acids and Metal Ions on the Biosensor Performance. *Sensors* **2020**, *20*, 3138.
22. Hou, Q.-H.; Ma, A.-Z.; Lv, D.; Bai, Z.-H.; Zhuang, X.-L.; Zhuang, G.-Q. The Impacts of Different Long-Term Fertilization Regimes on the Bioavailability of Arsenic in Soil: Integrating Chemical Approach with Escherichia Coli ArsRp:: Luc-Based Biosensor. *Appl Microbiol Biotechnol* **2014**, *98*, 6137–6146.
23. Council, N.R. *Bioavailability of Contaminants in Soils and Sediments: Processes, Tools, and Applications*; National Academies Press, 2003; ISBN 0309086256.
24. Zeng, N.; Wu, Y.; Chen, W.; Huang, Q.; Cai, P. Whole-Cell Microbial Bioreporter for Soil Contaminants Detection. *Front Bioeng Biotechnol* **2021**, *9*, 622994.
25. Zhu, Y.G.; Yoshinaga, M.; Zhao, F.J.; Rosen, B.P. Earth Abides Arsenic Biotransformations. *Annu Rev Earth Planet Sci* 2014.
26. Yang, H.C.; Rosen, B.P. New Mechanisms of Bacterial Arsenic Resistance. *Biomed J* 2016.
27. Oremland, R.S.; Stolz, J.F. The Ecology of Arsenic. *Science (1979)* 2003.
28. Hedges, R.W.; Baumberg, S. Resistance to Arsenic Compounds Conferred by a Plasmid Transmissible between Strains of Escherichia Coli. *J Bacteriol* **1973**, doi:10.1128/jb.115.1.459-460.1973.
29. Yan, G.; Chen, X.; Du, S.; Deng, Z.; Wang, L.; Chen, S. Genetic Mechanisms of Arsenic Detoxification and Metabolism in Bacteria. *Curr Genet* **2019**, *65*, 329–338.
30. Oshima, T.; Imahori, K. Description of Thermus Thermophilus (Yoshida and Oshima) Comb. Nov., a Nonsporulating Thermophilic Bacterium from a Japanese Thermal Spa. *Int J Syst Bacteriol* **1974**, doi:10.1099/00207713-24-1-102.
31. del Giudice, I.; Limauro, D.; Pedone, E.; Bartolucci, S.; Fiorentino, G. A Novel Arsenate Reductase from the Bacterium Thermus Thermophilus HB27: Its Role in Arsenic Detoxification. *Biochim Biophys Acta Proteins Proteom* **2013**, doi:10.1016/j.bbapap.2013.06.007.
32. Antonucci, I.; Gallo, G.; Limauro, D.; Contursi, P.; Ribeiro, A.L.; Blesa, A.; Berenguer, J.; Bartolucci, S.; Fiorentino, G. Characterization of a Promiscuous Cadmium and Arsenic Resistance Mechanism in Thermus Thermophilus HB27 and Potential Application of a Novel Bioreporter System. *Microb Cell Fact* **2018**, doi:10.1186/s12934-018-0918-7.
33. Antonucci, I.; Gallo, G.; Limauro, D.; Contursi, P.; Ribeiro, A.L.; Blesa, A.; Berenguer, J.; Bartolucci, S.; Fiorentino, G. An ArsR/SmtB Family Member Regulates Arsenic Resistance Genes Unusually Arranged in Thermus Thermophilus HB27. *Microb Biotechnol* **2017**, doi:10.1111/1751-7915.12761.

34. Gallo, G.; Antonucci, I.; Pirone, L.; Amoresano, A.; Contursi, P.; Limauro, D.; Pedone, E.; Bartolucci, S.; Fiorentino, G. A Physicochemical Investigation on the Metal Binding Properties of TtSmtB, a Thermophilic Member of the ArsR/SmtB Transcription Factor Family. *Int J Biol Macromol* **2019**, *138*, doi:10.1016/j.ijbiomac.2019.07.174.
35. Politi, J.; Spadavecchia, J.; Fiorentino, G.; Antonucci, I.; de Stefano, L. Arsenate Reductase from Thermus Thermophilus Conjugated to Polyethylene Glycol-Stabilized Gold Nanospheres Allow Trace Sensing and Speciation of Arsenic Ions. *J R Soc Interface* **2016**, doi:10.1098/rsif.2016.0629.
36. Politi, J.; Spadavecchia, J.; Fiorentino, G.; Antonucci, I.; Casale, S.; de Stefano, L. Interaction of Thermus Thermophilus ArsC Enzyme and Gold Nanoparticles Naked-Eye Assays Speciation between As(III) and As(V). *Nanotechnology* **2015**, *26*, 435703, doi:10.1088/0957-4484/26/43/435703.
37. Puopolo, R.; Sorrentino, I.; Gallo, G.; Piscitelli, A.; Giardina, P.; le Goff, A.; Fiorentino, G. Self-Assembling Thermostable Chimeras as New Platform for Arsenic Biosensing. *Sci Rep* **2021**, *11*, 1–12.
38. López-López, O.; Cerdán, M.-E.; González-Siso, M.-I. Thermus Thermophilus as a Source of Thermostable Lipolytic Enzymes. *Microorganisms* **2015**, *3*, 792–808.
39. Raddadi, N.; Cherif, A.; Daffonchio, D.; Neifar, M.; Fava, F. Biotechnological Applications of Extremophiles, Extremozymes and Extremolytes. *Appl Microbiol Biotechnol* 2015.
40. Aulitto, M.; Fusco, S.; Fiorentino, G.; Limauro, D.; Pedone, E.; Bartolucci, S.; Contursi, P. Thermus Thermophilus as Source of Thermozymes for Biotechnological Applications: Homologous Expression and Biochemical Characterization of an α-Galactosidase. *Microb Cell Fact* **2017**, doi:10.1186/s12934-017-0638-4.
41. Sandmann, G. Carotenoid Biosynthesis and Biotechnological Application. In Proceedings of the Archives of Biochemistry and Biophysics; 2001.
42. Yokoyama, A.; Shizuri, Y.; Hoshino, T.; Sandmann, G. Thermocryptoxanthins: Novel Intermediates in the Carotenoid Biosynthetic Pathway of Thermus Thermophilus. *Arch Microbiol* **1996**, doi:10.1007/s002030050336.
43. Cava, F.; Hidalgo, A.; Berenguer, J. Thermus Thermophilus as Biological Model. *Extremophiles* 2009.
44. Tišma, M.; Bucić-Kojić, M.; Planinić, M. Bio-Based Products from Lignocellulosic Waste Biomass: A State of the Art. *Chem Biochem Eng Q* **2021**, *35*, 139–156.
45. Mujtaba, M.; Fraceto, L.; Fazeli, M.; Mukherjee, S.; Savassa, S.M.; de Medeiros, G.A.; Santo Pereira, A. do E.; Mancini, S.D.; Lipponen, J.; Vilaplana, F. Lignocellulosic Biomass from Agricultural Waste to the Circular Economy: A Review with Focus on Biofuels, Biocomposites and Bioplastics. *J Clean Prod* **2023**, 136815.
46. Lucas, F.W.S.; Grim, R.G.; Tacey, S.A.; Downes, C.A.; Hasse, J.; Roman, A.M.; Farberow, C.A.; Schaidle, J.A.; Holewinski, A. Electrochemical Routes for the Valorization of Biomass-Derived Feedstocks: From Chemistry to Application. *ACS Energy Lett* **2021**, *6*, 1205–1270.
47. Valdivia, M.; Galan, J.L.; Laffarga, J.; Ramos, J. Biofuels 2020: Biorefineries Based on Lignocellulosic Materials. In *Microbial biotechnology*; Wiley Online Library, 2016; Vol. 9, pp. 585–594.

48. Aich, S.; Datta, S. Efficient Utilization of Lignocellulosic Biomass: Hydrolysis Methods for Biorefineries. In *Biorefineries: A Step Towards Renewable and Clean Energy*; Verma, P., Ed.; Springer Singapore: Singapore, 2020; pp. 273–295 ISBN 978-981-15-9593-6.
49. Squinca, P.; Bilatto, S.; Badino, A.C.; Farinas, C.S. Nanocellulose Production in Future Biorefineries: An IntegratedApproach Using Tailor-Made Enzymes. *ACS Sustain Chem Eng* **2020**, *8*, 2277–2286, doi:10.1021/acssuschemeng.9b06790.
50. Drula, E.; Garron, M.-L.; Dogan, S.; Lombard, V.; Henrissat, B.; Terrapon, N. The Carbohydrate-Active Enzyme Database: Functions and Literature. *Nucleic Acids Res* **2022**, *50*, D571–D577.
51. Andlar, M.; Rezić, T.; Marđetko, N.; Kracher, D.; Ludwig, R.; Šantek, B. Lignocellulose Degradation: An Overview of Fungi and Fungal Enzymes Involved in Lignocellulose Degradation. *Eng Life Sci* **2018**, *18*, 768–778.
52. Karnaouri, A.; Antonopoulou, I.; Zerva, A.; Dimarogona, M.; Topakas, E.; Rova, U.; Christakopoulos, P. Thermophilic Enzyme Systems for Efficient Conversion of Lignocellulose to Valuable Products: Structural Insights and Future Perspectives for Esterases and Oxidative Catalysts. *Bioresour Technol* **2019**, *279*, 362–372.
53. Finore, I.; Feola, A.; Russo, L.; Cattaneo, A.; Di Donato, P.; Nicolaus, B.; Poli, A.; Romano, I. Thermophilic Bacteria and Their Thermozymes in Composting Processes: A Review. *Chemical and Biological Technologies in Agriculture* **2023**, *10*, 7.
54. Kumar, S.; Dangi, A.K.; Shukla, P.; Baishya, D.; Khare, S.K. Thermozymes: Adaptive Strategies and Tools for Their Biotechnological Applications. *Bioresour Technol* 2019, *278*, 372–382.

Chapter 2 – Extremophiles, a Nifty Tool to face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering

Extremophilic microorganisms are a valuable reservoir of biomolecules naturally stable to harsh conditions that can be applied in various fields. Among them, thermophiles show significant potential for a wide range of industrial and environmental applications requiring high temperatures.

The ability of these microorganisms to degrade waste biomasses and chemical contaminants represents a promising avenue for their employment in biocatalysis and biotransformation processes; in this sense they can contribute not only to environmental protection, but also to enhance the sustainable production of high-value biotechnological products. The increasing demand to improve the performance of these microorganisms has been the driving force to establish new genetic engineering tools. At the top of that CRISPR-Cas technology, can significantly expand their applicability.

This chapter summarizes the pros and cons in the exploration of biotechnological applications of thermophiles, by examining resistance mechanisms to pollutants and their role in bioremediation processes.

International Journal of *Environmental Research and Public Health*

MDPI

Review

Extremophiles, a Nifty Tool to Face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering

Giovanni Gallo [1,2], Rosanna Puopolo [1], Miriam Carbonaro [1], Emanuela Maresca [1] and Gabriella Fiorentino [1,2,*]

[1] Department of Biology, University of Naples Federico II, Via Cinthia 21, 80126 Napoli, Italy; giovanni.gallo2@unina.it (G.G.); rosanna.puopolo@unina.it (R.P.); miriam.carbonaro@unina.it (M.C.); emanuelamaresca95@gmail.com (E.M.)

[2] Consiglio Nazionale delle Ricerche CNR, Institute of Polymers, Composites and Biomaterials (IPCB), Via Campi Flegrei, 34, 80078 Pozzuoli, Italy

* Correspondence: fiogabri@unina.it

Abstract: Extremophiles are microorganisms that populate habitats considered inhospitable from an anthropocentric point of view and are able to tolerate harsh conditions such as high temperatures, extreme pHs, high concentrations of salts, toxic organic substances, and/or heavy metals. These microorganisms have been broadly studied in the last 30 years and represent precious sources of biomolecules and bioprocesses for many biotechnological applications; in this context, scientific efforts have been focused on the employment of extremophilic microbes and their metabolic pathways to develop biomonitoring and bioremediation strategies to face environmental pollution, as well as to improve biorefineries for the conversion of biomasses into various chemical compounds. This review gives an overview on the peculiar metabolic features of certain extremophilic microorganisms, with a main focus on thermophiles, which make them attractive for biotechnological applications in the field of environmental remediation; moreover, it sheds light on updated genetic systems (also those based on the CRISPR-Cas tool), which expand the potentialities of these microorganisms to be genetically manipulated for various biotechnological purposes.

Keywords: extremophiles; environmental pollution; heavy-metal resistance; aromatic-compounds; bioremediation; biosensors; genome-engineering; CRISPR-Cas

Citation: Gallo, G.; Puopolo, R.; Carbonaro, M.; Maresca, E.; Fiorentino, G. Extremophiles, a Nifty Tool to Face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering. *Int. J. Environ. Res. Public Health* **2021**, *18*, 5228. https://doi.org/10.3390/ijerph18105228

Academic Editor: Paul B. Tchounwou

Received: 16 April 2021
Accepted: 9 May 2021
Published: 14 May 2021

1. Introduction

Natural environments on Earth display the most miscellaneous life conditions, and microorganisms are among the few entities that are able to grow in very extreme and inhospitable habitats. Hot springs, volcanic areas, polar regions, saline-alkaline, or acidic lakes and deep-see hydrothermal vents are some examples of natural environments that show temperature, salt concentration, pH, and pressure conditions very harsh for almost all forms of life [1].

Extremophiles are microorganisms that can live in these kinds of natural niches, and based on the conditions in which they thrive, they can be grouped in: acidophiles/alkaliphiles that grow at acid or alkaline pHs, halophiles that can live at high salt concentrations, piezophiles that prosper in high pressure conditions, metallophiles that are able to thrive in presence of metals/heavy metals, psychrophiles which live at low temperatures, and thermophiles/hyperthermophiles that grow at elevated temperatures [2]. Moreover, as a result of anthropogenic activities, some microbes adapted to flourish in polluted environments such as industrial wastewaters and contaminated soils characterized by the presence of toxic substances like pesticides, heavy metals, and different chemicals [3,4]. Therefore, extremophiles possess peculiar biological molecules and metabolic pathways that allow them to face multiple environmental stresses, sometimes simultaneously. For example,

enzymes of thermophiles (called thermozymes), in comparison to their mesophilic counterparts, maintain their activity and their folding at higher temperatures due to a more compact hydrophobic core, and a better distribution of hydrogen bonds and salt bridges at the protein surface [5]. At the same time thermozymes are also stable in the presence of organic solvents, denaturing agents, and high salinity, and more resistant to proteolysis, mirroring the niches where they are found. So far, extremophiles have increasingly received attention for their biotechnological significance and for industrial purposes; the study of these peculiar microorganisms, their metabolisms and catalysts makes possible to develop bioremediation technologies and bio-based energy processes.

In the last years, as a consequence of environmental pollution, global warming, and depletion of non-renewable sources and with the push of the 2030 Agenda for Sustainable Development drawn up by the United Nations [6], many research efforts have been focused either on the optimization of green and sustainable industrial processes (biorefineries), and on the setup of biotechnological methods to monitor and remove pollutants from the environment (biomonitoring and bioremediation, respectively) [7,8]. In this context, investigation on the biology, ecology, and physiology of microorganisms is a necessary prerequisite to set up white and green biotechnologies [9] in the field of industrial processes, energy generation [10], prevention of environmental pollution by detection and/or removal of contaminants [11], and production of biopolymers from renewable resources [12,13]. For example, thermophilic microorganisms can find applications to reduce pollution in industrial wastewaters that are often characterized by higher temperatures and highly dissolved heavy metals [14,15]. Moreover, thermophiles are more advantageous than mesophiles in biorefineries which require high-temperature steps [16].

This review analyses the metabolic strategies adopted by extremophilic microorganisms, with major emphasis on thermophilic ones, to face three classes of compounds with high impact on environmental pollution: heavy metals, organic compounds, and lignocellulosic biomasses, as well as their exploitation for application in bioremediation, biosensing, and biorefinery. Moreover, it provides updates regarding available genetic systems to engineer these microorganisms, in order to use them as platforms for metabolic engineering and production of valuable compounds.

2. Heavy Metals

The term "heavy metals" is widely referred to a group of metals and metalloids associated with potential toxicity or ecotoxicity. Generally, these metals possess relatively high densities, atomic weights, or atomic numbers. The criteria used for this classification vary, depending on the author and the context [17]; a recent paper reported by the *International Journal of Environmental Research and Public Health*, proposed to refer to them as "potentially toxic elements" [18]. Heavy metals are among the most persistent and toxic pollutants in the environment (Figure 1); they are non-biodegradable, and even in small concentrations, can threaten human and environmental health [19]. Heavy metals naturally occur in soils, rocks, sediments, air, and waters and microbial communities affect their speciation and mobility in the environment, because they are actively involved in metal geochemical cycles [20]. In traces, several heavy metals are essential for life; almost half of all enzymes require the presence of a metal atom to function [21]. Some of them as iron, copper, nickel, manganese, and zinc play key roles as functional centers in proteins and enzymes (i.e., metalloproteins) allowing biological transformations that are exceptionally unlikely to proceed spontaneously [22], as manganese in manganese-peroxidases or copper in laccases [23]; in fact these metallozymes are often employed as industrial biocatalysts (see Section 4.1, Lignin degrading thermozymes). The uncontrolled urbanization and the anthropogenic activities have much altered metal amounts in the environment; in fact, heavy metals are released from mining activities and industrial wastes, vehicle emissions, microplastics floating in the world's oceans or they come from common devices as lead-acid batteries, fertilizers, paints [24–26]. On the other hand, their use is expected to increase over time,

since many heavy metals, like copper or nickel, have been identified by the European Commission as critical raw materials for the transition to green energy technologies [27,28].

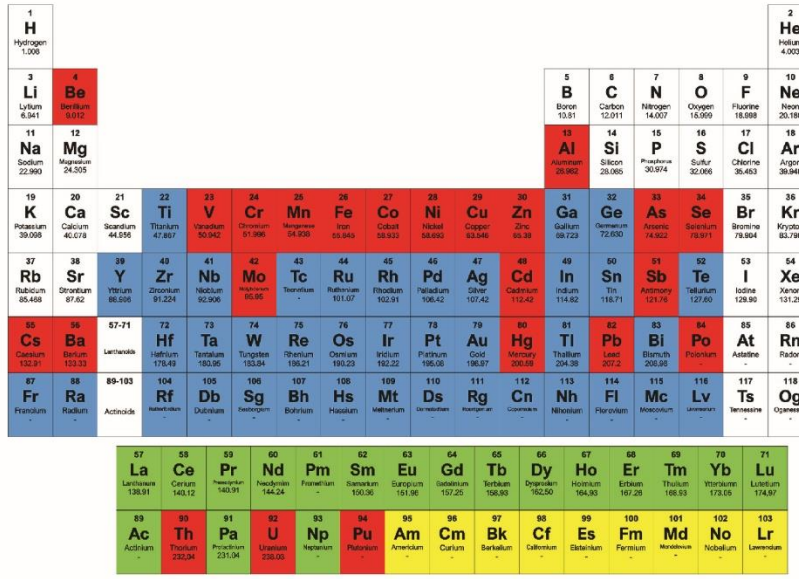

Figure 1. Periodic table of elements. Metals/metalloids are highlighted on the basis of the main characteristic that define them as "heavy": density > 5 g/cm^3 (**blue**); toxic (**red**); rare (**green**); synthetic (**yellow**).

Microorganisms have evolved resistance systems to cope with these toxic metals that usually rely on a balance between uptake and efflux processes [29]; many of these systems are also common in mesophiles, but in thermophilic Bacteria/Archaea they can present peculiar features [30]. The comprehension of the heavy metals resistance systems in thermophiles is increasingly supported by genome analyses, which allow to individuate their putative molecular determinants [31,32]. To date, at least four mechanisms of heavy metal resistance have been described: extracellular barrier; active transport of metal ions (efflux); enzymatic reduction of metal ions; intracellular sequestration [33–39]. Some bacteria are able to form complexes or chelates with extracellular polymers that reduce the permeability of metals [40]. However, heavy metals can escape this system and enter the cell thanks to the uptake systems of elements essential for life, for example, arsenic enters the cell via the phosphate or the glucose transporters [41]. Usually, the resistance to heavy metals is due to the coordinated work of intracellular enzymatic oxido-reduction and heavy metal efflux systems which generally consuming ATP, pushes the toxic metal outside the cell [42,43]. The resistance genes are usually organized in operons that also guarantee the expression of a transcription factor that regulates the whole system. Regarding the intracellular sequestration of heavy metals, the general mechanism foresees that some proteins, rich in cysteine residues, form complexes with the metals by exploiting the thiol groups [30].

Sometimes the same microorganism owns more metal resistance mechanisms; for example, *Escherichia coli* possesses either a copper active transport system (CopA) and another system based on multicopper oxidases (CueO) and CusCFBA transport system for periplasmic copper detoxification [23,44]. These resistance mechanisms are often activated as stress response [45,46]; understanding their underpinning molecular basis is crucial for application in the environmental monitoring of metal contamination (biosensing) and/or to set up bioremediation processes [14,47,48].

Biometallurgy is the branch of biotechnology that exploits the interaction between microorganisms (or their components) and metals or metal-bearing minerals (Figure 2) [49]. It includes microbial processes as metal biosorption, bioaccumulation or biomining (described below); these processes play a crucial role, on one hand, in the supply of critical raw materials, because can offer eco-efficient alternatives to classical pyro- or hydrometallurgical processes [50,51], and on the other hand in the set-up of strategies for metal biomonitoring and bioremediation (Table 1). In this context, the exploitation of thermophiles offers several advantages related to their ability to survive under harsh conditions and to degrade recalcitrant mineral species. Furthermore, in principle they could be successfully used in situ for metal bioremediation and/or biorecovery in any environment [52].

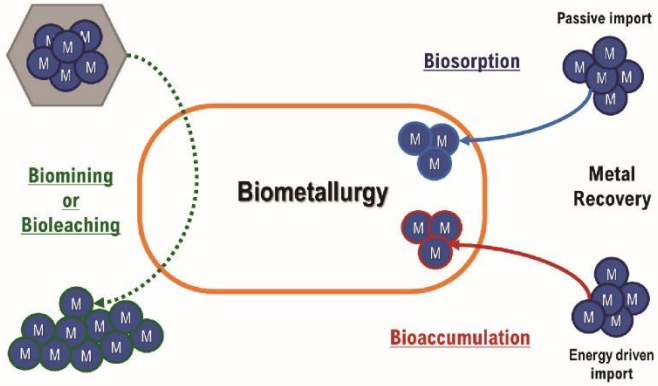

Figure 2. Schematic representation of metal bioprocesses in biometallurgy. Metals (blue circles) can be extracted from ores through biomining/bioleaching (green arrow) and they can be recovered into the cell by passive import (biosorption—blue arrow) or energy driven transport (bioaccumulation—red arrow).

Table 1. Examples of thermophiles exploited in biometallurgy.

Heavy Metals Tolerant Thermophiles				
Application	**Target**	**Microorganism**	**Temperature**	**Ref.**
Bioleaching	Cu, Zn, Ni, Cd, Al, Cr, Pb	Consortium of *Sulfobacillus thermosulfidooxidans* and *Thermoplasma acidophilum*	45 °C	[53]
Biosorption	Ag, Cd, Co, Cr, Cu, Fe, Pb, Zn	*Geobacillus thermodenitrificans*	60 °C	[3]
Biomineralization and Bioaccumulation	Eu	*T. scotoductus* SA-01	65 °C	[54]
Biosensing	Cd, As	*T. thermophilus* HB27	70 °C	[55]

Table 1. *Cont.*

Heavy Metals Tolerant Thermophiles				
Application	**Target**	**Microorganism**	**Temperature**	**Ref.**
Biosensing	Ni, Zn, Co, Hg, Mn, Cr, Cu, Fe, Cd	*A. amylolyticus*	60 °C	[56]
Biosorption	Cd, Cu, Co, Mn	*Geobacillus thermodenitrificans* and *A. amylolyticus*	60 °C	[57]
Biosorption	Cd	*Geobacillus stearothemophilus*	60 °C	[58]
Bioaccumulation	Cd, Cu, Ni, Mn, Zn	*Geobacillus toebii* subsp. *Decanicus* *Geobacillus thermoleovorans* subsp. *stromboliensis*	60 °C	[59]
Biosorption	U, Th	*B. cereus* SO-14	65 °C	[60]
Biosensing	As, Cd, Hg, Pb	*Acidibacillus ferrooxidans*	45 °C	[61]

The possibility to combine metal biorecovery with bioremediation represents an intriguing challenge to reduce process costs: microorganisms can recover metals from polluted sites, contemporarily reducing pollution, and producing valuable elements [62]. The microbial pathways that can be exploited to remove toxic metals from an environment are those related to biosorption/bioaccumulation; they consist into the ability of microorganisms to sequester heavy metals on the cell surface or intracellularly. In particular, the term "biosorption" is referred to passive processes that follow a kinetic equilibrium, while "bioaccumulation" to energy driven processes which require active metabolism [63]. For example, heavy metals metabolic pathways have been widely investigated within the *Geobacillus* genus, since many members of this species are highly tolerant to various heavy metals (As, Ag, Cd, Co, Cr, Cu, Fe, Pb, U, Zn) [57,59,64]. In particular, their biosorption and bioaccumulation mechanisms have been analyzed and applied on environmental samples to remove unwanted metals [3,58,59].

On the other hand, biomining (called "bioleaching" if metals are solubilized during the process) consists in the ability of microorganisms to extract and recover metals from ores and waste concentrates [65–67]. Several microbial species, as *Sulfobacillus* sp. and *Ferroplasma* sp., are well known for their ability to solubilize Fe(II) [31,68]; usually they are acidophilic chemolithotrophs (autotrophs or mixotrophs) presenting iron and/or sulfur oxidizing pathways [69]. In some cases, they are used as part of microbial *consortia*, which can perform the bioleaching of different metals simultaneously [53].

Microorganisms able to extract and accumulate metals from ores or geothermal sources are very interesting for their potential application in the biorecovery of rare-earth metals, for example, the Europium (Eu), which is widely used for the production of modern devices (solar cells, mobile phones and computers, biomedical instruments); *Thermus scotoductus* SA-01 can survive in the presence of high levels (up to 1 mM) of Eu, a concentration hundred times higher than that typically found in the environment and is able to extract and accumulate it from geothermal fluids [54]. The biorecovery can be useful also for monitoring metals at low concentration in the environment: Özdemir S. and co-workers [60] set up a preconcentration method with *Bacillus cereus* SO-14 to increase sensitivity in the detection of U(VI) and Th(IV) by ICP-OES (Inductively Coupled Plasma—Optical Emission Spectrometry).

Microorganisms have also been exploited for biomonitoring as *whole-cell* biosensors. In *Thermus thermophilus* HB27, the arsenic responsive transcriptional repressor *Tt*SmtB regulates the expression of the arsenic efflux protein *Tt*ArsX, in particular, *Tt*SmtB responds to variation of Cd(II), As(III) and As(V) concentrations [68,69]; therefore, Antonucci and co-workers engineered *T. thermophilus* to express a reporter gene from the *Ttars*X promoter [55].

In the set-up of systems for metal biomonitoring, it is also possible to follow a decrease of enzymatic activity as a toxicological indicator of heavy metals: Poli and co-workers observed a decrease in the α-amylase activity of *Anoxybacillus amylolyticus*, in the presence

of heavy metals [56]. In contrast, Shih-Hung and co-workers followed the inhibition of the iron-oxidizing activity of an *Acidibacillus ferrooxidans* strain [61].

In the fields of heavy metals bioremediation and biomonitoring many thermophilic oxidoreductases have also been characterized. For example, quinone oxidoreductase, chromate reductase, and superoxide dismutase from different *Anoxibacillus* species have been employed for Pb and Cr bioremoval [70,71]; moreover, the arsenate reductase from *T. thermophilus* HB27 (*Tt*ArsC) has been exploited as the biological recognition element for the development of different arsenic biosensors [72–74].

3. Organic Pollutants

Organic pollutants are a wide class of chemically different organic compounds released in the environment as toxic wastes [75]. They originate from domestic sewage, urban run-off, industrial effluents, and agricultural wastewater and include pesticides, fertilizers, hydrocarbons, phenols, plasticizers, biphenyls, detergents, oils, greases, and pharmaceuticals [76]. Therefore, the organic pollutants are a very heterogeneous group: the main constituents of the persistent organic pollutants (POPs) are organochlorinated pesticides (OCPs), polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCB), dioxins, and dibenzofurans; they are persistent because they remain intact in the environment for extended periods (years or decades in soil/sediment) [77,78]. These compounds are released in air and soil and even though they are scarcely soluble in water, can be biomagnified in living organisms and cause adverse effects to human health [79]. Among the most common environmental pollutants of the marine environment, there are petroleum hydrocarbons that contaminate the sea through natural oil spills, like reservoirs and volcanic processes in the deep ocean, and artificial oil spills, as oil tanker accidents, oil transportation processes, or oil refineries. In most cases, this last process represents the primary way to contaminate the sea with crude oil (Figure 3) [80,81].

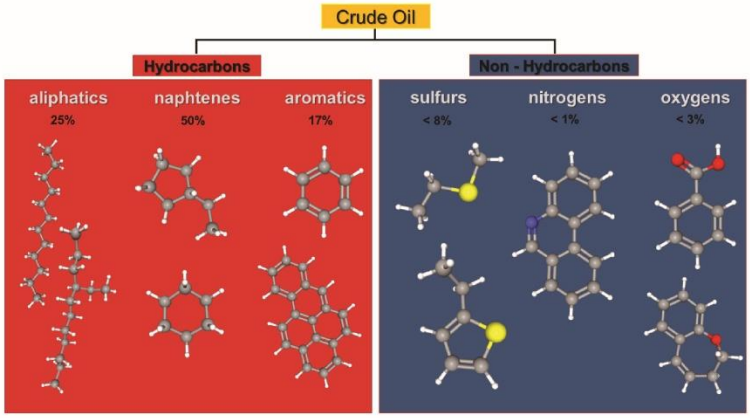

Figure 3. Schematic representation of organic compounds in crude oil. Atoms are reported in grey (C), white (H), yellow (S), blue (N) and red (O).

Microbial activities on anthropogenic organic compounds usually arise from evolution of previously existing enzymes and metabolic pathways. Microorganisms have evolved effective catalysts for detoxification of toxic compounds, as result of a selective pressure [82]. Generally, toxic compounds are converted into metabolites entering central metabolic

pathways: for example, pathways responsible for the biodegradation of aliphatic and alicyclic carboxylic acids include β-oxidation, combined α- and β-oxidation, and aromatization pathways [83]. In *Sulfolobus solfataricus* members of the multiple antibiotic resistance regulators family (MarR-family) are involved in detoxification of aromatic compounds, as benzaldehyde and salicylate [84,85]. Aromatic hydrocarbon dioxygenases, belonging to the large family of Rieske non-heme iron oxygenases (ROHs), catalyze the initial reaction in the bacterial biodegradation of a diverse array of aromatic and polyaromatic hydrocarbons, aromatic acids, chlorinated aromatics, and heterocyclic aromatic compounds [86]. They are attractive in biotechnology for bioremediation as well as for the production of industrially and medically important chiral chemicals; for example, toluene dioxygenase catalyze the oxidation of benzene to benzene *cis*-diol [87], and many thermophilic bacteria distributed mainly among *Chloroflexi*, *Deinococcus–Thermus*, and *Firmicutes* have been identified as sources of these appealing enzymes [88].

Bioremediation of toxic compounds represents an effective and sustainable technology, compared with physical and chemical remediation technologies, based on microbial activities that in an ideal bioprocess degrade all the substances to CO_2 and H_2O (complete mineralization) [89]. In crude-oil bioremediation processes, the exploitation of thermophiles can be considered an optimal choice; in fact, at higher temperature there is a decrease in the oil viscosity that increases the diffusion rates of organic compounds making them more accessible to microbial degradation [90]. Since the degradation activity can be often substrate-specific, biodegradation processes can be optimized using microbial consortia, thus expanding the spectrum of action [91,92]; for example, two strains of *Geobacillus jurassicus* and *Geobacillus subterraneus*, isolated from the Dagang high temperature oil field in China, can grow on benzoate but not phenol; therefore, to degrade crude oil, they need the presence of complementary phenol degrading activities (in this specific case a *G. stearothermophilus*, which can use phenol but not benzoate) [93].

In addition to biodegradation processes (Table 2), thermophiles also produce macromolecules that can be considered useful for the bioremediation of organic pollutants: *Bacillus licheniformis* and *Anaerophaga thermohalophila* have been characterized for the production, under anaerobic conditions, of low molecular weight peptides, which are surface-active compounds, exploitable as biosurfactants for oil removal [94].

Table 2. Examples of thermophiles exploited in bioremediation processes of organic compounds.

Organic Compounds Degrading Thermophiles				
Bioprocess	**Target**	**Microorganisms**	**Temperature**	**Ref.**
Biodegradation	Crude oil	Consortium of *Bacillus, Geobacillus* and *Clostridium*	55 °C	[91]
Biodegradation	Hydrocarbons	*Geobacillus pallidus*	30–70 °C	[4]
Biofilter	Volatile Organic Compounds (VOCs)	Consortium of 25 genera belonging to Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, Deltaproteobacteria, Flavobacteriia, Sphingobacteriia, and Bacilli classes	50–60 °C	[92]
Biodegradation	Phenolic compounds	*Bacillus thermoleovorans* sp. A2	65 °C	[95]
Biodegradation	Hydrocarbons	Consortium of *Geobacillus* and *Thermoattinomyces* spp.	60 °C	[93]

The substrate specificity and the stability of detoxifying thermozymes make them also exploitable as recognition elements of biosensors, especially those which require electrochemical detection; for example, the haloacid dehalogenase, L-HAD_{ST} of *Sulfolobus tokodaii* was immobilized an N-hydroxysuccinimidyl Sepharose resin and used for the detection of halogenated organic compounds, retaining 70% of its initial activity after storage at 4 °C for 6 months [95,96].

Toxic Dyes

Industrialization has represented one of the main causes of water pollution. Wastewaters can be rich in recalcitrant, mutagen and carcinogenic compounds [97]. Dyes are a class of very toxic pollutants that are released in the wastewaters of textile manufacturing. These recalcitrant compounds change both the pH and the chemical oxygen demand (COD) and biochemical oxygen demand (BOD) in aquatic ecosystems [98,99]. The dyes are classified on the basis of their chemical structure or industrial uses. The most employed dyes are acid dye, synthetic dye, and direct dye. Even if they can are complex organic molecules, each dye has a characteristic chromophore: for example, the acid and synthetic azo dyes are typical for their azo linkage (–N = N–), the central chromophore of the anthraquinone dyes derives from the oxidation of anthracene, the indigoid dyes derive from indoles [100].

There are several technologies for colored wastewater remediation: physical, chemical, and biological. The physical treatments include screening, coagulation, precipitation, adsorption and membrane filtration; the chemical treatments comprise coagulation–flocculation, oxidation, ozonation, Fenton oxidation, photocatalytic oxidation, ion exchange, and electrochemical treatments; the biological methods are aerobic, anaerobic, and anaerobic–aerobic treatments in which the contaminated organic compounds are converted into safe and stable compounds [101,102]. Each technology has advantages and inconveniences; in fact, to date the typical method for wastewater remediation of colored waters is physical–chemical flocculation combined with biological treatment [101,102].

Interestingly, the analysis of microbial communities of these waters revealed the occurrence of several bacteria able to decompose the dyes: they possess intra/extra cellular oxidoreductases, such as dye decoloration peroxidases (DyP), laccases, azoreductases [103]. Many laccases and azoreductases have been characterized in several thermophiles like *T. thermophilus*, *Geobacillus*, *Anoxibacillus*, *Thermosediminibacter* species, and others (Table 3). Their enzymes are expected to be more stable to extreme temperatures and pHs in comparison to those of mesophilic bacteria.

Table 3. Examples of dye-decolorizing thermophiles.

Microorganisms	Substrates	Ref.
Anoxybacillus pushchinoensis, *Anoxybacillus kamchatkensis and Anoxybacillus flavithermus*	Reactive Black 5	[104]
Anoxybacillus sp.	Congo red	[105]
G. stearothermophilus	Remazol Brilliant Blue R, Methyl Orange, Malachite Green (MG) and Indigo Carmine	[106]
T. thermophilus HB27	Dye orange, Acid red dye, green dye, naphthol brilliant blue, Remazol brilliant blue, congo red	[107]
T. thermophilus SG0.5JP17-16	Congo Red, Reactive Black B and Reactive Black WNN, and Remazol Brilliant Blue R	[108]
Thermus sp. 2.9	Xylidine, RBBR, Gentian Violet, Methyl Orange	[109]
Geobacillus sp. JS12	Congo red, Malachite green	[110]
Anoxybacillus ayderensis SK3-4	Direct blue 6, acid black 1, direct green 6, direct black 19, and acid blue 93	[111]
Thermosediminibacter oceani	Malachite green (MG) and Congo red	[112]

In the genome of *T. thermophilus* HB27, a laccase that is able to oxidize six different dyes (dye orange, acid red dye, green dye, naphthol brilliant blue, Remazol brilliant blue, Congo red), has been characterized; the enzyme requires an electron shuffle, that supports this reaction for some dyes [107]. Also *Geobacillus* sp. JS12 contains a laccase, LacG, that can decolor these artificial compounds at 70 °C [110]. In the alkaliphilic and thermophilic bacterium *Anoxybacillus* sp. strain UARK-01, the UARK 01 laccase can oxidize the Congo red substrate, one of the most toxic dyes [105]. In addition to the laccases, these extremophiles also have a striking azoreductase, active on different dyes; for example, *Anoxibacillus* sp. PDR2 acts towards the direct black G [113]; moreover, the

degradation of the same dye can be obtained by a thermophilic microflora, consisting of facultative aerobic (*Anoxybacillus flavithermus* strain 52-1A, *Tepidiphilus thermophilus* strain JHK30, *Tepidiphilus succinatimandens* strain 4BON, *Brevibacillus aydinogluensis* strain PDF25, *Bacillus thermoamylovorans* strain DKP and *Geobacillus thermoleovorans* strain NP1) and exclusively anaerobic bacteria (*Thermoanaerobacterium thermosaccharolyticum* strain DSM 571, *Thermoanaerobacterium thermostercoris* strain Buff, and *Caloramator proteoclasticus* strain Uruguayensis) [114]. These examples give a generic view on the potential application of thermophilic oxidoreductases in biological detoxification.

4. Lignocellulosic Biomasses

In the last decades, there is a growing interest in the use of microbes in industrial processing to break waste food and lignocellulose biomasses to produce biofuels and bioproducts. Among renewable resources, non-food lignocellulosic waste biomasses are currently considered among the most promising materials, since they are present in large quantities and at low cost [115]. Every year, a significant amount of lignocellulosic residues is generated worldwide from agricultural wastes, food industry, household garbage, non-food seeds, etc. (see Table 4), causing an increase in environmental pollution. Lignocellulosic wastes are also often improperly stored and recalcitrant to different disposal treatments; moreover, when burnt, they provoke environmental pollution problems. Thus, the reuse and exploitation of such wastes in industrial biotechnology to produce interesting chemicals allows to bypass a part of disposal treatments [116].

Table 4. Percentage content of most common lignocellulosic wastes [117,118].

Lignocellulosic Wastes	Lignin (%)	Hemicellulose (%)	Cellulose (%)
Softwood stems	25–35	25–35	45–50
Hardwood stems	18–25	24–40	40–55
Miscellaneous Corn stover	19	22	39
Wheat straw	15	50	30
Rice straw	18	24	33
Nutshells	30–40	25–30	25–30
Peels	14–20	11	4
Shells	26–30	20–25	40–45
Sorted refuse	20	20	60
Swine waste	n/a	28	6
Solid cattle manure	2.7–5.7	1.4–3.3	1.6–4.7
Grass	10–25	35–50	30–40
Cotton seed hairs	0	5–20	80–95
Leaves	0	80–85	15–20
Sawdust	14–34	71–89	31–64
Paper	0–15	0	85–99
Newspaper	14–19	25–40	40–55
Wastepaper from chemical pulps	5–10	10–20	60–70
Primary wastewater solids	24–29	n/a	8–15

Lignocellulose is a significant component of plant biomass and it consists of cellulose, hemicellulose, and lignin. Cellulose and hemicellulose are polymers of different sugars; in particular, the principal constituent of lignocellulosic biomass is cellulose, a polysaccharide composed of β-1,4-linked D-glucose units, widely employed for paper and cardboard production. Instead, hemicellulose is a complex branched polysaccharide formed by a mixture of xylans, mannans, β-glucans, and xyloglucans, depending on the type of wood. In softwood, hemicellulose mainly consists of galactoglucomannan, composed of β-1,4-linked D-glucose and D-galactose units. In contrast, xylan is the constituent of hemicellulose in hardwood, and it is formed of β-1,4-linked D-xylose units, which can be substituted with other monosaccharides [119]. Hemicellulose is closely associated with cellulose filaments and covalently attached to lignin, forming a matrix.

Lignin is an aromatic heteropolymer composed of ether and C–C bonds that link phenylpropanoid aryl-C3 units. The percentage composition of these polymers in waste lignocellulosic biomasses varies as shown in the table below; the content of cellulose, hemicellulose and lignin also changes in a single plant depending on the age, stage of growth, and other conditions [120].

A high number of microorganisms, belonging to both bacterial and archaeal kingdom, possesses complex metabolic pathways able to decompose lignocellulose. Their enzymes can be utilized as biocatalysts for green approaches in several industrial fields, such as paper and pulp industry, food processing and textile sector, agriculture, animal food production, etc. (Figure 4). Furthermore, several microbial based technologies for the exploitation of lignocellulosic wastes as raw materials for producing bioproducts and biofuels were set up. In the specific case of bioethanol production, from early 70 to 2000s, a "first-generation technology" (1G) was developed, in which the biorefinery systems were based on the use of starch/sugar crops (sugar beet, maize, and sugar). However, 1G has several unsustainability issues bound to the great request of crops subtracted to the food chain and the cultivation of large areas (destined for this purpose) that causes deforestation and decrease of biodiversity [121]. As an alternative, in the "second-generation" technologies (2G), lignocellulosic materials are employed as feedstocks; in this respect, lignocellulose is cheap and immediately available in a large amount [122], but the development of tailored technologies is necessary to exploit more recalcitrant components. For example, both 1G and 2G technologies have to simultaneously maximize production yield and reduce costs and environmental impact; in both cases, exploitation of microbial mechanisms and biocatalysis supports the process. The bioethanol production process consists of different phases: biomass pre-treatment, saccharification, fermentation, and distillation. The main difference between the two technologies lies in the complexity of the starting raw material; in the first-generation technologies after pre-treatment of sugarcane and maize, a chemically homogenous material (sucrose and starch) can be easily broken into sugar units by a limited number of enzymes like amylases, amylopullulanases or glucosidases [123]. For the improvement of 1G processes, investigation on α-amylases, α-D-glucosidases, pullulanases and amylopullulanases of thermophilic bacteria and archaea have been carried out. For example, *T. thermophilus* HB27, *Thermoanaerobacter ethanolicus* 39E, *Geobacillus thermoleovorans* NP33, *Rhodothermus marinus*, *Clostridium thermosulfurogenes*, *Clostridium thermocellum*, *Desulfurococcus mucosus*, *Fervidobacterium pennavorans*, *Bacillus stearothermophilus*, *Thermotoga maritima* and some species of the genera *Pyrococcus*, *Thermoanaerobacter*, and *Thermococcus* have been studied since they are able to produce starch degrading enzymes. However, some of their biocatalysts have limited activity with high starch concentration (>30%). Therefore, mesophilic hosts are still the preferred ones for bioethanol production [124,125]. In fact, standardized methodologies for saccharification and fermentation have been mainly optimized in engineered mesophilic yeasts or microbes (*Saccharomyces* and *Zymomonas* spp.) with a high bioethanol production yield [126].

On the other hand, the degradation of lignocellulose is more complicated because the starting matrix is heterogeneous. The lignin removal step requires a significant amount of energy (acid hydrolysis or steam explosion) to release sugar polymers for the subsequent saccharification step [123]. Moreover, in 2G, either in saccharification and fermentation, additional steps are necessary to achieve the complete production of bioethanol from cellulose and hemicellulose; therefore, despite its cheapness and availability, lignocellulosic material implicates a more elaborate treatment process.

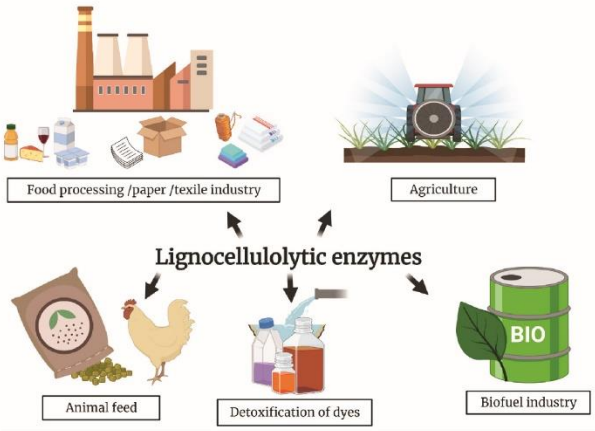

Figure 4. Industrial applications of lignocellulolytic enzymes. In food processes cellulases and xylanases are employed to improve the shelf life of dairy products and to hydrolyze monosaccharides in milk processing; they are also used in winery industries and to decrease viscosity of fruit juice. Also, laccases are used to increase quality of beverages and food, for example eliminating toxic substances. In pulp/paper industry laccases and xylanases enhance pulp bleaching for paper manufacturing, and cellulases improve flexibility and softness of fibers. Lignocellulolytic enzymes can be employed for improving nutrient digestibility of animal feeds, for enhancing color and surface brightness of fabric in textile industry, for textile dye bleaching, for synthesis of complex polymers. In agriculture, they are involved in fruit ripening and defense mechanisms against insects. Laccases are employed in wastewater treatment of colored waters. Lignocellulolytic enzymes are used in biorefinery systems to produce biofuels [103,127,128] (Created with BioRender.com (accessed on 20 January 2021)).

In this context, the use of thermophiles in lignocellulosic biomass degradation has the advantage that higher temperatures and organic solvents can be used, reducing either risks of microbial contamination or energy consumption (because the cooling steps are not necessary), and increasing rates of hydrolysis and product yields [129]. Another attractive progress on the use of renewable lignocellulosic biomass to produce bioethanol or feedstock chemicals consists in setting up microbial based bioprocesses that exploit the synergistic degradative capabilities of thermophilic microorganisms or consortia [130].

4.1. Lignin Degrading Thermozymes

Enzymes that depolymerize lignin are isolated principally from the white-rote fungi; some of these enzymes are manganese peroxidase (MnP), lignin peroxidase (LiP), versatile peroxidase (VP), and laccase. In particular, laccases are a heterogeneous subfamily of multicopper oxidases (MCOs) that can be involved in several biological processes like lignolysis and detoxification of dyes (see above) [131].

Recent studies have described very promising thermostable laccases able to degrade lignin derived from thermophilic microorganisms such as *Bacillus* sp. PC-3 and sp. FNT with optimum activity temperatures of 60 and 70 °C, respectively [132,133]. The first archaeal laccase was characterized from the halotolerant *Haloferax volcanii*, a promising microorganism for the lignin break-down purposes; this archaeon grows up to 50 °C and it possesses a stable glycoprotein, the laccase LccA that acts on several substrates at elevated temperature (55 °C), high salt concentrations (0.1 to 1.4 M) and it can maintain its activity also in organic solvents [134]. Also, many *Thermus* species have laccases that can be employed in lignin degradation; laccase of *Thermus* sp. 2.9 can retain 80% of its activity at 70 °C for 16 h and is able to successfully delignify *Eucalyptus* biomass [135].

Two strains of *T. thermophilus* (HJ6 and HB27) produce laccases capable of maintaining an optimal temperature range of activity at 85–90 °C for reactions up to 1 h [107,136]. Moreover, two hyperthermophilic bacteria present laccases with remarkable heat stability; a chemolithoautotrophic bacterium, *Aquifex aeolicus*, expresses a multicopper oxidase with an optimal temperature of 75 °C but that preserves its activity at 80 and 90 °C for up to 9 and 5 h, respectively [137]. Furthermore, the laccase-like multi-copper oxidase of *Thermobaculum terrenum* is extremely thermostable with a half-time of inactivation of 2.24 days at 70 °C and 350 min at 80 °C and pH 7 [138].

These features suggest that an impressive compromise between thermostability and lasting activity exists for these enzymes; for this reason, they are considered as promising tools to degrade the lignin component of biomasses.

4.2. Cellulose and Hemicellulose Degrading Thermozymes

Several thermophilic archaea and bacteria are able to produce a considerable amount of promising cellulose/hemicellulose degrading enzymes. These thermophilic enzymes, differently from their mesophilic counterparts, have the advantage that they can be added immediately after the thermochemical pre-treatment of biomass, making the cooling steps not necessary, increasing conversion efficiency, and saving time [139].

Whereas cellulose can be completely depolymerized through endoglucanases, exoglucanases and β-D-glucosidases, the total degradation of hemicellulose requires a wider pool of enzymatic activities (i.e., β-xylosidases, β-xylanases, β-glucuronidases, β-mannanase, β-mannosidase, α-galactosidase, etc.).

Several thermophilic cellulose degrading bacteria have been isolated from distinct environments like hot springs, compost systems and soil. They include different species belonging to the genera *Actinomadura*, *Alicyclobacillus*, *Anoxybacillus*, *Acidothermus*, *Bacillus*, *Caldanaerovirga*, *Caldicellulosirupto*, *Cellulomonas*, *Clostridium*, *Dictyoglomus*, *Geobacillus*, *Paenibacillus*, *Nesterenkonia*, *Paenibacillus*, *Pyrococcus*, *Rhodothermus*, *Sulfolobus*, *Thermoanaerobacterium*, and *Thermotoga*; they can produce both cellulose and hemicellulose degrading enzymes that can raise the rates of biomass hydrolysis if they are used in industrial bioprocesses [140].

Some examples of remarkable thermophiles include the following ones.

Acidothermus cellulolyticus 11B, isolated from a hot spring in Yellowstone National Park, produces a tri-functional enzyme that can break down birchwood xylan with high efficiency; in fact, this enzyme has endo-xylanase, arabinofuranosidase, and acetyl-xylan esterase activities [141].

Two thermoalkaline species of *Anoxybacillus* (*kamchatkensis* NASTPD13 and sp. 3M) express many xylanases and β-xylosidases, respectively, highly resistant to alkaline and acidic pHs, denaturing agents and organic solvents [142,143]. Also the facultative anaerobic *Bacillus coagulans* MA-13, which lives at an optimal temperature of 55 °C, secretes an endo-1,4-β-glucanase which can act from 37 to 60 °C. *B. coagulans* MA-13 is also able to ferment sugars derived from pre-treatment of lignocellulose to lactic acid in the presence of inhibitors; in fact, it was proved that this bacterium can grow and ferment in bioreactors containing 95% hydrolysate [144]. At the same time, seed culture pre-adaptation of *B. coagulans* MA-13, before simultaneous saccharification and fermentation step, can improve the production of lactic acid; again, it has a pool of interesting intra- and extracellular enzymes with glycosyl hydrolyzing activities that make *B. coagulans* MA-13 useful for increasing nutritional value of food [145,146].

Clostridium thermocellum has a non-enzymatic scaffolding protein bound with different enzymatic subunits that simultaneously degrade cellulose and hemicellulose [147]. In 2018 a new cellulolytic strain was identified in the *Chryseobacterium* genus, which produces an enzyme with a double cellulase/xylanase activity working either on carboxymethylcellulose and birchwood xylan [148].

Instead, the anaerobic *Caldicoprobacter* sp. CL-2, isolated from bovine manure compost, has a xylanase activity showing a modular structure with a glycoside hydrolase

domain coupled with a carbohydrate binding module [149]. Another hyperthermophilic microorganism from geothermal springs that produce miscellaneous glycoside hydrolases is *T. maritima*; it has an endoglucanase enzyme (Tm_Cel5A) with an optimum T of 80 °C and pH of 4.8. Tm_Cel5A is a peculiar GH5 (glycoside hydrolase family 5) enzyme with an unusual activity because it can act both on glucan and mannan based polysaccharides, while the other GH5 hydrolysing enzymes degrade either cellulose or mannans [150].

Dictyoglomus turgidum is another thermophilic microorganism that displays a set of genes encoding putative enzymes with glycosyl hydrolyse activity; this not yet well characterized thermophile has an endo-1,4-β-mannanase, *Dtur*CelB, with a high thermoresistance (Tm of 88 °C) and a good thermal and pH stability; it is also resistant to chemicals and has been analyzed in an enzymatic cocktail able to cut-off cellulose and hemicellulose [151,152]. In fact, different thermophilic biocatalysts can be utilized synergistically for the complete breakdown of hemicellulose sugars (pentose and hexose) of lignocellulosic material. The two recombinant thermophilic enzymes, the above mentioned *Dtur*CelB from *D. turgidum* and the α-galactosidase from *T. thermophilus*, can be tested in the lignocellulose pre-hydrolyzing step right before the saccharification step [153].

Other two thermophiles, *Thermotoga neapolitana* 5068 and *T. thermophilus*, display the most thermoactive (~100 °C) and thermostable (half-life of 30 h at 70 °C) α-galactosidase activities, respectively [154,155].

Archaeal glycoside hydrolyzing enzymes have also been exploited to improve biomass degradation processes. For example, the hyperthermophilic archaeon *Sulfolobus shibatae* encodes an endo-1,4-β-D-glucanase that accomplishes the break-down of carboxymethylcellulose, xylan and barley β-glucan [156]. *Pyrococcus furiosus* produces extracellular endoglucanases, intracellular glucosidases, and different intra- and extracellular amylases with a high thermostability in the range 80–100 °C; one of these enzymes, a β-glycosidase, is immobilized and used in industrial process of lactulose production [157]. Furthermore, *Saccharolobus solfataricus* expresses a membrane-bound xylanase, an extracellular endoglucanase, intra- and extracellular galactosidases, an extracellular xylosidase and an intracellular mannosidase. In the case of *S. solfataricus*, its thermozymes have been used as model for engineering mesophilic enzymes in order to improve their thermostability; for instance, the β-glycosidase of this archaeon, that shows a maximal activity above 95 °C, represents a fine example of an efficacious heterologous production in a yeast expression system [158].

In recent years, several studies are focused on taking advantage on thermophilic communities that can provide a high hydrolyzation rate of lignocellulosic material (Table 5). For example, consortia formed by bacterial and fungal microorganisms such as Alcaligenaceae, Burkholderiacea, *Thermoamylovorans*, Xanthomonadaceae, *Mycobacterium*, *Talaromyces* and *Rubrobacter* can decompose biomasses with a high content of lignin [159].

Furthermore, high throughput genome sequencing, transcriptomics, proteomics, metagenomic, and other omics techniques together with metabolic engineering strategies and bioinformatic tools have contributed significantly to explore a considerable amount of novel thermophilic lignocellulolytic microorganisms and enzymes.

Table 5. Examples of thermophilic bacteria and archaea able to hydrolyze the lignocellulose.

Lignocellulosic Component	Microorganism	Temperature	Ref.
Lignin	*A. aeolicus*	89 °C	[137]
	Bacillus sp. PC-3	55–92 °C	[132]
	Bacillus sp. FNT	50–55 °C	[133]
	H. volcanii	50 °C	[160]
	Thermus sp. 2.9	65 °C	[135]
	T. thermophilus HJ6	80 °C	[136]
	T. thermophilus HB27	70 °C	[107]
	T. terrenum	67 °C	[138]
	Fungal and bacterial consortium	55 °C	[159]

Table 5. *Cont.*

Lignocellulosic Component	Microorganism	Temperature	Ref.
Cellulose and hemicellulose	*A. cellulolyticus* 11B	70 °C	[141]
	A. kamchatkensis NASTPD13	60 °C	[142]
	Anoxybacillus sp. 3M	55 °C	[143]
	B. coagulans MA-13	55 °C	[144]
	Brevibacillus borstelensis SDM	50 °C	[161]
	C. thermocellum	60 °C	[147]
	Chryseobacterium sp.	55 °C	[148]
	Caldicoprobacter sp. CL-2	60–75 °C	[149]
	D. turgidum	75–80 °C	[151]
	T. maritima	80 °C	[150]
	T. neapolitana 5068	70–80 °C	[162]
	T. thermophilus HB27	70 °C	[107]
	P. furiosus	100 °C	[158]
	S. shibatae	80 °C	[156]
	S. solfataricus	80 °C	[158]

The development of genome editing tools also represents a new approach to address biomass degradation by microorganisms; in fact in a next future, the genome manipulation of thermophilic bacteria will make possible to develop fine bioprocessing microbial strains, that will be capable of better performing degradation of lignocellulose [163]. Thus, thermophiles have a great potential to be considered as a suitable platform for metabolic engineering to produce various biomolecules and/or valuable chemicals from lignocellulosic biomasses.

5. Engineering of Thermophiles

Thermophilic microorganisms have unique biochemical and physiological characteristics with important biotechnological implications. Thermophilic microorganisms can be used in numerous applications, such as biocatalysis, or as sources of thermoactive or thermostable enzymes. However, unfortunately, their employment as *whole-cell* systems is limited by the lack of easily usable genetic systems. This situation has changed recently, with unprecedented progress in genetic tools for extremophilic microorganisms, and the use of these microorganisms as platforms has become possible.

Significant studies have been made to develop and improve molecular genetic techniques for thermophilic microorganisms in the past decade, either belonging to the bacterial or archaeal kingdom. A significant challenge for genetic modification in thermophiles is the choice of a selectable marker to screen positive transformants. The antibiotics typically used in mesophiles often target cell components specific to bacteria and are ineffective against the archaeal species. Even in cases where antibiotics are useful, both the antimicrobial compound and the gene product that confers resistance must be stable at elevated temperatures. Due to the low efficiency of the heat-resistant antibiotic selection markers, usually nutritional selection systems such as enzymes essential for the synthesis of amino acids can be used. To date, genetic techniques have been obtained for ten such archaea, including *Metallosphaera*, *Sulfolobus*, *Thermococcus*, and *Pyrococcus* species [164–167].

The creation of genome editing tools enabling stable integration of genetic elements into host chromosomes is crucial for industrial applications, where plasmid instability becomes problematic and volumes of antibiotics on an industrial scale are very polluting. In principle, two alternative approaches for developing genome editing tools can be adopted for thermophilic bacteria; one is by adapting mesophilic protocols to function at elevated temperatures. The second is to seek alternative means of genome editing from thermophilic springs. Several examples describe the use of homologous recombination to knock out or replace chromosomal genes in thermophilic bacteria. In 2012, Suzuki and co-workers developed a *pyrF*/*pyr*R counterselection system for *Geobacillus kaustophilus*, enabling marker-free genome editing at 60 °C [168].

Another widely used system to obtain genetic manipulation of thermophilic bacteria is the *Cre/loxP* site-specific recombination [169]. This recombination is performed between two *loxP* sites using a Cre recombinase, *loxP* is a 34 bp consensus DNA sequence with a central spacing region of 8 bp, which defines its orientation, flanked by two 13 bp palindromic sequences, which are the Cre binding sites. The *Cre/lox* system's effectiveness in a broad spectrum of biological species and a wide variety of applications has made this technology indispensable for in vivo genetic manipulation. This system allows various recombination types, such as conditional recombination, intermolecular recombination and time and space specific recombination [169]. Recently, a *Cre/lox* system was developed for the thermophilic bacterium *T. thermophilus* HB27 [170], leading to the development of a highly efficient method of destroying multiple genes to facilitate genetic manipulation of this bacterium. The most important advantage that made easier to develop genetic tools for *T. thermophilus* is the constitutive expression of a natural competence system in several strains [171]. Several plasmids have been developed to transform *T. thermophilus*, and some of these, suitably modified using regions of homology to the chromosome, have been used to stimulate homologous recombination, obtain deletions of genes, thus allowing the study of the in vivo function of specific proteins [172].

For these reasons, *T. thermophilus* is considered a biological model for functional studies and a right candidate for biotechnological applications. However, its efficient defense system against the exogenous DNA can be an impairment since it can destroy the cloning vectors used for transformation; in 2014, Daan C. Swarts and co-workers identified *Tt*Ago, a protein belonging to the Argonaute family as the protein responsible for the prevention of the uptake and propagation of foreign DNA [173]. The researchers observed that the protein generally attacks the AT-rich regions of double-stranded DNA, leading to the complete plasmid degradation by other nucleases [173].

The Rise of the CRISPR-Cas Era

Until 2013, the principal genome editing tools were the zinc finger nucleases, the transcription activator-like effector nucleases and intrinsic homologous recombination systems [174–176]. These systems use artificial fusion proteins composed by an engineered DNA-binding domain fused to the non-specific nuclease domain of the restriction enzyme FokI. These systems were extensively used for the genome editing of eukaryotic microorganisms. A new technology for genome editing rose based on RNA-guided engineered nucleases (CRISPR-Cas9 system) in the last decades. Although the CRISPR array was discovered in the late 1980s [177], its function remained unknown until 2005 [178]. Only in 2007, it was concluded that it represented a bacterial innate immunity system [179,180]. The transition of the CRISPR/Cas system from a biological phenomenon to a tool for genome engineering occurred when it was shown that the target DNA sequence could be reprogrammed simply by changing 20 nucleotides in the crisprRNA (crRNA) and that the targeting specificity of the crRNA could be combined with the structural properties of the tracrRNA (trans-activating crisprRNA) in a chimeric single guide RNA (sgRNA) [181] (Figure 5).

Furthermore, the evidence that sgRNAs with different specificities could be produced made it possible to modify more loci simultaneously, giving a connection to the so-called CRISPR-mania [182]. The various genome editing applications pioneered in human and animal cells have recently been transferred back to bacteria to carry out genome editing and transcriptional control, as well as genome-wide screens. In fact, the CRISPR-Cas system was used to obtain some genetically modified prokaryotes. However, one problem for applying this genetic editing tool to thermophilic microorganisms, is that it is based on a mesophilic system. In recent years the research has been going towards the search for Cas proteins from thermophiles; in fact, a thermostable genome editing tool was developed based on a thermophilic Cas9, that can be used up to 55 °C and contains everything necessary for genome editing in a single plasmid; with the advent of the ThermoCas9,

genome manipulation in moderate thermophilic bacteria becomes possible, making the editing process much more comfortable and less time-consuming [183].

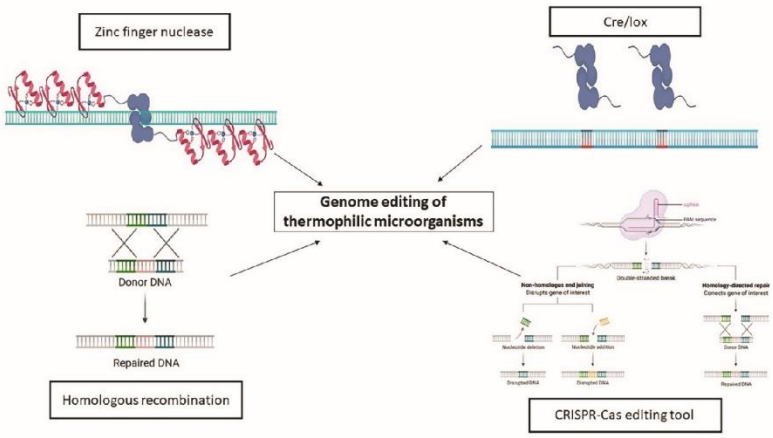

Figure 5. Schematic representation of the genome editing tools for thermophilic microorganisms. The Zinc finger nuclease is the first genome editing tool setup for thermophilic microorganisms. The most common is the spontaneous homologous recombination. In the last decades, the Cre/lox and CRISPR-Cas based tools were adapted for thermophiles (Created with BioRender.com (accessed on 20 January 2021)).

A genome-editing tool was recently developed for moderate thermophilic bacteria obtained using the Cas12a from *Francisella novicida* [184]; this system allowed to obtain knockout mutants in less than one week with high editing efficiencies. *Fn*Cas12a has an interesting potential for the genome editing of many thermophilic bacteria and archaea.

Cas9 and Cas12a are multidomain CRISPR-associated nucleases that can cleave complementary DNA targets using a guide RNA. The Cas9 belongs to type II-a, while the Cas12a to type V-a. The first enzyme is the best characterized and utilizes nuclease for genome-editing purposes. In the last years, Cas12a has emerged as a potential alternative. These two enzymes have distinct evolutionary origins and present different structural architectures, resulting in specific molecular mechanisms; in fact, the nuclease activities of Cas9 and Cas12a and the resulting DNA repair outcomes are affected by circumstantial factors such as cell type, target sequence, and genomic context [185]. Their biological differences influence their application as genome editing tools: in some cases, the Cas9 activity is more suitable for some organisms, in other cases the best option is to use Cas12a. Instead, Cas9 and Cas12a and their engineered variants are highly complementary in their properties and together build up a powerful and versatile toolkit.

6. Conclusions

Extremophiles represent a class of microorganisms very interesting for their ability to live in harsh conditions, not only high temperature, but also extreme pHs and high salinity concentrations. These peculiar characteristics make them and their biocatalysts very promising tools for industrial and environmental applications. Thermophilic extremophiles stand out in biometallurgy for biomonitoring and bioremediation, as well as in degradation of organic biomasses to transform them into resources ready to be re-used. In fact, in

addition to the biofuel production, hydrolyzing extremozymes have a wide range of applications in the food, feed, beverage, textile, pulp and paper industry.

Improved knowledge in *omic-era* and the increasing need to address environmental pollution with green processes drive biotechnological research in search of microorganisms that can replace chemical processes. The newly available thermophilic genome editing tools based on the CRISPR-Cas system, open the way for the complete achievement of these goals in various industrial fields. In fact, the rise of the "CRISPR-Cas *era*" makes possible the application of engineered extremophiles as a *whole-cell* platform.

Author Contributions: Conceptualization, G.G., R.P., M.C., E.M. and G.F.; investigation, G.G., R.P., M.C. and E.M.; data curation, G.G., R.P., M.C. and E.M.; figure preparation, R.P. and M.C.; writing—original draft preparation, G.G., R.P., M.C., E.M. and G.F.; writing—review and editing, G.G., R.P., M.C., E.M. and G.F.; visualization, G.G., R.P., M.C., E.M. and G.F.; supervision, G.F.; project administration, G.F.; funding acquisition, G.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by PRIN 2017-GoodbyWaste: Obtain GOOD products–exploit BY-products–reduce WASTE, funded by MIUR (Ministero dell'Istruzione dell'Università e della Ricerca) and by ERA-NET Cofund MarTERA who funded the project: "FLAshMoB: Functional Amyloid Chimera for Marine Biosensing".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Arora, N.K.; Panosyan, H. Extremophiles: Applications and roles in environmental sustainability. *Environ. Sustain.* **2019**, *2*, 217–218. [CrossRef]
2. Dumorné, K.; Córdova, D.C.; Astorga-Eló, M.; Renganathan, P. Extremozymes: A potential source for industrial applications. *J. Microbiol. Biotechnol.* **2017**, *27*, 649–659. [CrossRef] [PubMed]
3. Chatterjee, S.K.; Bhattacharjee, I.; Chandra, G. Biosorption of heavy metals from industrial waste water by Geobacillus thermodenitrificans. *J. Hazard. Mater.* **2010**, *175*, 117–125. [CrossRef]
4. Zheng, C.; He, J.; Wang, Y.; Wang, M.; Huang, Z. Hydrocarbon degradation and bioemulsifier production by thermophilic Geobacillus pallidus strains. *Bioresour. Technol.* **2011**, *102*, 9155–9161. [CrossRef] [PubMed]
5. Finch, A.; Kim, J. Thermophilic Proteins as Versatile Scaffolds for Protein Engineering. *Microorganisms* **2018**, *6*, 97. [CrossRef] [PubMed]
6. DESA UN (Department of Economic and Social Affairs of United Nations). Transforming Our World: The 2030 Agenda for Sustainable Development. In *A New Era in Global Health*; Springer: New York, NY, USA, 2015. [CrossRef]
7. Pakshirajan, K.; Rene, E.R.; Ramesh, A. Biotechnology in Environmental Monitoring and Pollution Abatement 2015. *Biomed Res. Int.* **2015**, *2015*. [CrossRef] [PubMed]
8. Schneider, F.; Kläy, A.; Zimmermann, A.B.; Buser, T.; Ingalls, M.; Messerli, P. How can science support the 2030 Agenda for Sustainable Development? Four tasks to tackle the normative dimension of sustainability. *Sustain. Sci.* **2019**, *14*, 1593–1604. [CrossRef]
9. Barcelos, M.C.S.; Lupki, F.B.; Campolina, G.A.; Nelson, D.L.; Molina, G. The colors of biotechnology: General overview and developments of white, green and blue areas. *FEMS Microbiol. Lett.* **2018**, *365*, fny239. [CrossRef]
10. Yadav, M.; Sehrawat, N.; Singh, M.; Kumar, V.; Sharma, A.K.; Kumar, S. Thermophilic microbes-based fuel cells: An eco-friendly approach for sustainable energy production. In *Bioremediation for Environmental Sustainability*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 235–246.
11. Hsieh, M.-C.; Chung, Y.-C. Measurement of biochemical oxygen demand from different wastewater samples using a mediator-less microbial fuel cell biosensor. *Environ. Technol.* **2014**, *35*, 2204–2211. [CrossRef]
12. Straub, C.T.; Counts, J.A.; Nguyen, D.M.N.; Wu, C.-H.; Zeldes, B.M.; Crosby, J.R.; Conway, J.M.; Otten, J.K.; Lipscomb, G.L.; Schut, G.J.; et al. Biotechnology of extremely thermophilic archaea. *FEMS Microbiol. Rev.* **2018**, *42*, 543–578. [CrossRef]
13. Lanzilli, M.; Esercizio, N.; Vastano, M.; Xu, Z.; Nuzzo, G.; Gallo, C.; Manzo, E.; Fontana, A.; D'Ippolito, G. Effect of Cultivation Parameters on Fermentation and Hydrogen Production in the Phylum Thermotogae. *Int. J. Mol. Sci.* **2020**, *22*, 341. [CrossRef]
14. Sar, P.; Kazy, S.K.; Paul, D.; Sarkar, A. Metal Bioremediation by Thermophilic Microorganisms. In *Thermophilic Microbes in Environmental and Industrial Biotechnology: Biotechnology of Thermophiles*; Satyanarayana, T., Littlechild, J., Kawarabayasi, Y., Eds.; Springer: Dordrecht, The Netherlands, 2013; pp. 171–201, ISBN 978-94-007-5899-5.

15. Satyanarayana, T.; Littlechild, J.; Kawarabayasi, Y. Thermophilic microbes in environmental and industrial biotechnology. *Biotechnol. Thermoph.* **2013**, *3*. [CrossRef]
16. Rekadwad, B.; Gonzalez, J.M. Multidisciplinary involvement and potential of thermophiles. *Folia Microbiol.* **2019**, *64*, 389–406. [CrossRef] [PubMed]
17. Duffus, J.H. "Heavy metals"—A meaningless term? (IUPAC Technical Report). *Pure Appl. Chem.* **2002**, *74*, 793–807. [CrossRef]
18. Pourret, O.; Hursthouse, A. It's Time to Replace the Term "Heavy Metals" with "Potentially Toxic Elements" When Reporting Environmental Research. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4446. [CrossRef]
19. Mathew, B.B.; Singh, H.; Biju, G.; Krishnamurthy, N.B. Classification, Source, and Effect of Environmental Pollutants and Their Biodegradation. *J. Environ. Pathol. Toxicol. Oncol.* **2017**, *36*, 55–71. [CrossRef]
20. Rensing, C.; Rosen, B.P. Heavy Metals Cycle (Arsenic, Mercury, Selenium, others). In *Encyclopedia of Microbiology*; Elsevier Inc.: Amsterdam, The Netherlands, 2009; pp. 205–219, ISBN 9780123739445.
21. Finkelstein, J. Metalloproteins. *Nature* **2009**, *460*, 813. [CrossRef]
22. Boer, J.L.; Mulrooney, S.B.; Hausinger, R.P. Nickel-dependent metalloenzymes. *Arch. Biochem. Biophys.* **2014**, *544*, 142–152. [CrossRef] [PubMed]
23. Janusz, G.; Pawlik, A.; Świderska-Burek, U.; Polak, J.; Sulej, J.; Jarosz-Wilkołazka, A.; Paszczyński, A. Laccase properties, physiological functions, and evolution. *Int. J. Mol. Sci.* **2020**, *21*, 966. [CrossRef]
24. Men, C.; Liu, R.; Wang, Q.; Guo, L.; Shen, Z. The impact of seasonal varied human activity on characteristics and sources of heavy metals in metropolitan road dusts. *Sci. Total Environ.* **2018**, *637–638*, 844–854. [CrossRef]
25. Vhahangwele, M.; Khathutshelo, L.M. Environmental Contamination by Heavy Metals. In *Heavy Metals*; Hosam, S., Ed.; IntechOpen: London, UK, 2018; pp. 115–133, ISBN 978-1-78923-361-2.
26. Cole, M.; Lindeque, P.; Halsband, C.; Galloway, T.S. Microplastics as contaminants in the marine environment: A review. *Mar. Pollut. Bull.* **2011**, *62*, 2588–2597. [CrossRef]
27. Horta Arduin, R.; Mathieux, F.; Huisman, J.; Blengini, G.A.; Charbuillet, C.; Wagner, M.; Baldé, C.P.; Perry, N. Novel indicators to better monitor the collection and recovery of (critical) raw materials in WEEE: Focus on screens. *Resour. Conserv. Recycl.* **2020**, *157*, 104772. [CrossRef] [PubMed]
28. European Commission. *Raw Materials Demand for Wind and Solar PV Technologies in the Transition towards a Decarbonised Energy System*; European Commission: Brussels, Belgium, 2020.
29. Pedone, E.; Bartolucci, S.; Fiorentino, G. Sensing and adapting to environmental stress: The archaeal tactic. *Front. Biosci.* **2004**, *9*, 2909–2926. [CrossRef] [PubMed]
30. Ranawat, P.; Rawat, S. Metal-tolerant thermophiles: Metals as electron donors and acceptors, toxicity, tolerance and industrial applications. *Environ. Sci. Pollut. Res.* **2017**, 1–29. [CrossRef]
31. Panyushkina, A.E.; Babenko, V.V.; Nikitina, A.S.; Selezneva, O.V.; Tsaplina, I.A.; Letarova, M.A.; Kostryukova, E.S.; Letarov, A. V Sulfobacillus thermotolerans: New insights into resistance and metabolic capacities of acidophilic chemolithotrophs. *Sci. Rep.* **2019**, *9*, 15069. [CrossRef]
32. Aulitto, M.; Gallo, G.; Puopolo, R.; Mormone, A.; Limauro, D.; Contursi, P.; Piochi, M.; Bartolucci, S.; Fiorentino, G. Genomic Insight of Alicyclobacillus mali FL18 Isolated From an Arsenic-Rich Hot Spring. *Front. Microbiol.* **2021**, *12*. [CrossRef]
33. Bruins, M.R.; Kapil, S.; Oehme, F.W. Microbial resistance to metals in the environment. *Ecotoxicol. Environ. Saf.* **2000**, *45*, 198–207. [CrossRef] [PubMed]
34. Cánovas, D.; Cases, I.; De Lorenzo, V. Heavy metal tolerance and metal homeostasis in Pseudomonas putida as revealed by complete genome analysis. *Environ. Microbiol.* **2003**, *5*, 1242–1256. [CrossRef]
35. Cazorla, F.M.; Arrebola, E.; Sesma, A.; Pérez-García, A.; Codina, J.C.; Murillo, J.; de Vicente, A. Copper Resistance in Pseudomonas syringae Strains Isolated from Mango Is Encoded Mainly by Plasmids. *Phytopathology* **2002**, *92*, 909–916. [CrossRef]
36. Cervantes, C.; Campos-García, J.; Devars, S.; Gutiérrez-Corona, F.; Loza-Tavera, H.; Torres-Guzmán, J.C.; Moreno-Sánchez, R. Interactions of chromium with microorganisms and plants. *FEMS Microbiol. Rev.* **2001**, *25*, 335–347. [CrossRef]
37. Cervantes, C.; Gutierrez-Corona, F. Copper resistance mechanisms in bacteria and fungi. *FEMS Microbiol. Rev.* **1994**, *14*, 121–137. [CrossRef] [PubMed]
38. Cha, J.S.; Cooksey, D.A. Copper resistance in Pseudomonas syringae mediated by periplasmic and outer membrane proteins. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 8915–8919. [CrossRef] [PubMed]
39. Lin, Y.-F.; Walmsley, A.R.; Rosen, B.P. An arsenic metallochaperone for an arsenic detoxification pump. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15617–15622. [CrossRef] [PubMed]
40. El-Helow, E.R.; Sabry, S.A.; Amer, R.M. Cadmium biosorption by a cadmium resistant strain of Bacillus thuringiensis: Regulation and optimization of cell surface affinity for metal cations. *Biometals* **2000**, *13*, 273–280. [CrossRef]
41. Yang, H.-C.; Rosen, B.P. New mechanisms of bacterial arsenic resistance. *Biomed. J.* **2016**, *39*, 5–13. [CrossRef]
42. Schelert, J.; Dixit, V.; Hoang, V.; Simbahan, J.; Drozda, M.; Blum, P. Occurrence and Characterization of Mercury Resistance in the Hyperthermophilic Archaeon Sulfolobus solfataricus by Use of Gene Disruption. *J. Bacteriol.* **2004**, *186*, 427–437. [CrossRef]
43. Smith, A.T.; Smith, K.P.; Rosenzweig, A.C. Diversity of the metal-transporting P 1B-type ATPases. *JBIC J. Biol. Inorg. Chem.* **2014**, *19*, 947–960. [CrossRef]
44. Bondarczuk, K.; Piotrowska-Seget, Z. Molecular basis of active copper resistance mechanisms in Gram-negative bacteria. *Cell Biol. Toxicol.* **2013**, *29*, 397–405. [CrossRef]

45. Bartolucci, S.; Contursi, P.; Fiorentino, G.; Limauro, D.; Pedone, E. Responding to toxic compounds: A genomic and functional overview of Archaea (383 views visite). *Front. Biosci.* **2013**, *18*, 165–189.
46. Ranawat, P.; Rawat, S. Stress response physiology of thermophiles. *Arch. Microbiol.* **2017**, *199*, 391–414. [CrossRef]
47. Gallo, G.; Puopolo, R.; Limauro, D.; Bartolucci, S.; Fiorentino, G. Metal-Tolerant Thermophiles: From the Analysis of Resistance Mechanisms to their Biotechnological Exploitation. *Open Biochem. J.* **2018**, *12*, 149–160. [CrossRef]
48. Vasudevan, N.; Jayshree, A. Extremozymes and Extremoproteins in Biosensor Applications. In *Encyclopedia of Marine Biotechnology*; Wiley Online Books; Wiley: Hoboken, NJ, USA, 2020; pp. 1711–1736, ISBN 9781119143802.
49. Zhuang, W.Q.; Fitts, J.P.; Ajo-Franklin, C.M.; Maes, S.; Alvarez-Cohen, L.; Hennebel, T. Recovery of critical metals using biometallurgy. *Curr. Opin. Biotechnol.* **2015**, *33*, 327–335. [CrossRef]
50. Hennebel, T.; Boon, N.; Maes, S.; Lenz, M. Biotechnologies for critical raw material recovery from primary and secondary sources: R&D priorities and future perspectives. *New Biotechnol.* **2015**, *32*, 121–127. [CrossRef]
51. Garole, D.J.; Hossain, R.; Garole, V.J.; Sahajwalla, V.; Nerkar, J.; Dubal, D.P. Recycle, Recover and Repurpose Strategy of Spent Li-ion Batteries and Catalysts: Current Status and Future Opportunities. *ChemSusChem* **2020**, *13*, 3079–3100. [CrossRef]
52. Castro, C.; Urbieta, M.S.; Cazón, J.P.; Donati, E.R. Metal biorecovery and bioremediation: Whether or not thermophilic are better than mesophilic microorganisms. *Bioresour. Technol.* **2019**, *279*, 317–326. [CrossRef] [PubMed]
53. Ilyas, S.; Lee, J.C.; Kim, B.S. Bioremoval of heavy metals from recycling industry electronic waste by a consortium of moderate thermophiles: Process development and optimization. *J. Clean. Prod.* **2014**, *70*, 194–202. [CrossRef]
54. Maleke, M.; Valverde, A.; Vermeulen, J.G.; Cason, E.; Gomez-Arias, A.; Moloantoa, K.; Coetsee-Hugo, L.; Swart, H.; Van Heerden, E.; Castillo, J. Biomineralization and bioaccumulation of europium by a thermophilic metal resistant bacterium. *Front. Microbiol.* **2019**, *10*, 81. [CrossRef]
55. Antonucci, I.; Gallo, G.; Limauro, D.; Contursi, P.; Ribeiro, A.L.A.L.; Blesa, A.; Berenguer, J.; Bartolucci, S.; Fiorentino, G. Characterization of a promiscuous cadmium and arsenic resistance mechanism in Thermus thermophilus HB27 and potential application of a novel bioreporter system. *Microb. Cell Fact.* **2018**, *17*. [CrossRef]
56. Poli, A.; Salerno, A.; Laezza, G.; di Donato, P.; Dumontet, S.; Nicolaus, B. Heavy metal resistance of some thermophiles: Potential use of α-amylase from Anoxybacillus amylolyticus as a microbial enzymatic bioassay. *Res. Microbiol.* **2009**, *160*, 99–106. [CrossRef] [PubMed]
57. Özdemir, S.; Kılınç, E.; Poli, A.; Nicolaus, B. Biosorption of Heavy Metals (Cd^{2+}, Cu^{2+}, Co^{2+}, and Mn^{2+}) by Thermophilic Bacteria, Geobacillus thermantarcticus and Anoxybacillus amylolyticus: Equilibrium and Kinetic Studies. *Bioremediat. J.* **2013**, *17*, 86–96. [CrossRef]
58. Hetzer, A.; Daughney, C.J.; Morgan, H.W. Cadmium ion biosorption by the thermophilic bacteria Geobacillus stearothermophilus and G. thermocatenulatus. *Appl. Environ. Microbiol.* **2006**, *72*, 4020–4027. [CrossRef] [PubMed]
59. Özdemir, S.; Kilinc, E.; Poli, A.; Nicolaus, B.; Güven, K. Cd, Cu, Ni, Mn and Zn resistance and bioaccumulation by thermophilic bacteria, Geobacillus toebii subsp. decanicus and Geobacillus thermoleovorans subsp. stromboliensis. *World J. Microbiol. Biotechnol.* **2012**, *28*, 155–163. [CrossRef] [PubMed]
60. Ozdemir, S.; Kılınc, E.; Yalcin, M.S.; Soylak, M.; Sen, F. A new magnetized thermophilic bacteria to preconcentrate uranium and thorium from environmental samples through magnetic solid-phase extraction. *J. Pharm. Biomed. Anal.* **2020**, *186*, 113315. [CrossRef] [PubMed]
61. Yang, S.H.; Cheng, K.C.; Liao, V.H.C. A novel approach for rapidly and cost-effectively assessing toxicity of toxic metals in acidic water using an acidophilic iron-oxidizing biosensor. *Chemosphere* **2017**, *186*, 446–452. [CrossRef] [PubMed]
62. Gherghe, S.L. Considerations about recovery of critical metals using bio-metallurgy. *E3S Web Conf.* **2017**, *18*, 01020. [CrossRef]
63. Hlihor, R.-M.; Apostol, L.-C.; Gavrilescu, M. Environmental Bioremediation by Biosorption and Bioaccumulation: Principles and Applications. In *Enhancing Cleanup of Environmental Pollutants: Volume 1: Biological Approaches*; Anjum, N.A., Gill, S.S., Tuteja, N., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 289–315, ISBN 978-3-319-55426-6.
64. Puopolo, R.; Gallo, G.; Mormone, A.; Limauro, D.; Contursi, P.; Piochi, M.; Bartolucci, S.; Fiorentino, G. Identification of a new heavy-metal-resistant strain of Geobacillus stearothermophilus isolated from a hydrothermally active volcanic area in southern Italy. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2678. [CrossRef]
65. Johnson, D.B.; Du Plessis, C.A. Biomining in reverse gear: Using bacteria to extract metals from oxidised ores. *Miner. Eng.* **2015**, *75*, 2–5. [CrossRef]
66. Donati, E.R.; Castro, C.; Urbieta, M.S. Thermophilic microorganisms in biomining. *World J. Microbiol. Biotechnol.* **2016**, *32*, 179. [CrossRef]
67. Dunbar, W.S. Biotechnology and the Mine of Tomorrow. *Trends Biotechnol.* **2017**, *35*, 79–89. [CrossRef]
68. Hawkes, R.B.; Franzmann, P.D.; Plumb, J.J. Moderate thermophiles including "Ferroplasma cupricumulans" sp. nov. dominate an industrial-scale chalcocite heap bioleaching operation. *Hydrometallurgy* **2006**, *83*, 229–236. [CrossRef]
69. Zhou, S.; Gan, M.; Zhu, J.; Liu, X.; Qiu, G. Assessment of Bioleaching Microbial Community Structure and Function Based on Next-Generation Sequencing Technologies. *Minerals* **2018**, *8*, 596. [CrossRef]
70. Jardine, J.L.; Stoychev, S.; Mavumengwana, V.; Ubomba-Jaswa, E. Screening of potential bioremediation enzymes from hot spring bacteria using conventional plate assays and liquid chromatography—Tandem mass spectrometry (Lc-Ms/Ms). *J. Environ. Manag.* **2018**, *223*, 787–796. [CrossRef] [PubMed]

71. Opperman, D.J.; Piater, L.A.; Van Heerden, E. A novel chromate reductase from Thermus scotoductus SA-01 related to old yellow enzyme. *J. Bacteriol.* **2008**, *190*, 3076–3082. [CrossRef]
72. Politi, J.; Spadavecchia, J.; Fiorentino, G.; Antonucci, I.; De Stefano, L. Arsenate reductase from *Thermus thermophilus* conjugated to polyethylene glycol-stabilized gold nanospheres allow trace sensing and speciation of arsenic ions. *J. R. Soc. Interface* **2016**, *13*, 20160629. [CrossRef] [PubMed]
73. Politi, J.; Spadavecchia, J.; Fiorentino, G.; Antonucci, I.; Casale, S.; De Stefano, L. Interaction of Thermus thermophilus ArsC enzyme and gold nanoparticles naked-eye assays speciation between As(III) and As(V). *Nanotechnology* **2015**, *26*, 435703. [CrossRef]
74. Puopolo, R.; Sorrentino, I.; Gallo, G.; Piscitelli, A.; Giardina, P.; Le Goff, A.; Fiorentino, G. Self-assembling thermostable chimeras as new platform for arsenic biosensing. *Sci. Rep.* **2021**, *11*, 2991. [CrossRef] [PubMed]
75. Cui, Z.; Tian, W.; Qin, H.; Wang, X.; Zhao, W. Optimal design and control of Eastman organic wastewater treatment process. *J. Clean. Prod.* **2018**, *198*, 333–350. [CrossRef]
76. Rashed, M.N. Adsorption Technique for the Removal of Organic Pollutants from Water and Wastewater. In *Organic Pollutants—Monitoring, Risk and Treatment*; InTech: West Palm Beach, FL, USA, 2013; pp. 167–194.
77. Érseková, A.; Hilscherová, K.; Klánová, J.; Giesy, J.P.; Novák, J. Effect-based assessment of passive air samples from four countries in Eastern Europe. *Environ. Monit. Assess.* **2014**, *186*, 3905–3916. [CrossRef]
78. Ashraf, M.A. Persistent organic pollutants (POPs): A global issue, a global challenge. *Environ. Sci. Pollut. Res.* **2017**, *24*, 4223–4227. [CrossRef]
79. Gaur, N.; Narasimhulu, K.; PydiSetty, Y. Recent advances in the bio-remediation of persistent organic pollutants and its effect on environment. *J. Clean. Prod.* **2018**, *198*, 1602–1631. [CrossRef]
80. Hassanshahian, M.; Amirinejad, N.; Behzadi, M.A. Crude oil pollution and biodegradation at the Persian Gulf: A comprehensive and review study. *J. Environ. Health Sci. Eng.* **2020**. [CrossRef] [PubMed]
81. Neamah, A.I. Separation of the Petroleum System. *Hilltop Rev.* **2014**, *7*, 11.
82. Copley, S.D.; Chem, N.; Author, B. Evolution of Efficient Pathways for Degradation of Anthropogenic Chemicals. *Nat. Chem. Biol.* **2009**, *5*, 559–566. [CrossRef] [PubMed]
83. Whitby, C. Microbial Naphthenic Acid Degradation. *Adv. Appl. Microbiol.* **2010**, *70*, 93–125.
84. Fiorentino, G.; Ronca, R.; Cannio, R.; Rossi, M.; Bartolucci, S. MarR-like transcriptional regulator involved in detoxification of aromatic compounds in Sulfolobus solfataricus. *J. Bacteriol.* **2007**, *189*, 7351–7360. [CrossRef] [PubMed]
85. Fiorentino, G.; Del Giudice, I.; Bartolucci, S.; Durante, L.; Martino, L.; Del Vecchio, P. Identification and physicochemical characterization of BldR2 from Sulfolobus solfataricus, a novel archaeal member of the MarR transcription factor family. *Biochemistry* **2011**, *50*, 6607–6621. [CrossRef] [PubMed]
86. Parales, R.E.; Resnick, S.M. Aromatic hydrocarbon dioxygenases. In *Biodegradation and Bioremediation*; Springer: Berlin, Germany, 2004; pp. 175–195.
87. Gibson, D.T.; Parales, R.E. Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr. Opin. Biotechnol.* **2000**, *11*, 236–243. [CrossRef]
88. Chakraborty, J.; Suzuki-Minakuchi, C.; Okada, K.; Nojiri, H. Thermophilic bacteria are potential sources of novel Rieske non-heme iron oxygenases. *AMB Express* **2017**, *7*, 1–15. [CrossRef]
89. Yemashova, N.A.; Murygina, V.P.; Zhukov, D.V.; Zakharyantz, A.A.; Gladchenko, M.A.; Appanna, V.; Kalyuzhnyi, S.V. Biodeterioration of crude oil and oil derived products: A review. *Rev. Environ. Sci. Bio/Technology* **2007**, *6*, 315–337. [CrossRef]
90. Koshlaf, E.; Ball, A.S. Soil bioremediation approaches for petroleum hydrocarbon polluted environments. *AIMS Microbiol.* **2017**, *3*, 25–49. [CrossRef]
91. Gu, G.; Li, Z.; Zhao, D.; Zhao, C. Isolation and characterization of a thermophilic oil-degrading bacterial consortium. *China Pet. Process. Petrochemical Technol.* **2013**, *15*, 82–90.
92. Yang, K.; Li, L.; Ding, W.; Liu, J.; Xue, S. A full-scale thermophilic biofilter in the treatment of sludge drying exhaust: Performance, microbial characteristics and bioaerosol emission. *J. Chem. Technol. Biotechnol.* **2018**, *93*, 2216–2225. [CrossRef]
93. Nazina, T.N.; Sokolova, D.S.; Shestakova, N.M.; Grigoryan, A.A.; Mikhailova, E.M.; Babich, T.L.; Lysenko, A.M.; Tourova, T.P.; Poltaraus, A.B.; Feng, Q.; et al. The Phylogenetic Diversity of Aerobic Organotrophic Bacteria from the Dagang High-Temperature Oil Field. *Microbiology* **2005**, *74*, 343–351. [CrossRef]
94. Karlapudi, A.P.; Venkateswarulu, T.C.; Tammineedi, J.; Kanumuri, L.; Ravuru, B.K.; Dirisala, V.r.; Kodali, V.P. Role of biosurfactants in bioremediation of oil pollution-a review. *Petroleum* **2018**, *4*, 241–249. [CrossRef]
95. Feitkenhauer, H.; Schnicke, S.; Müller, R.; Märkl, H. Determination of the kinetic parameters of the phenol-degrading thermophile Bacillus themoleovorans sp. A2. *Appl. Microbiol. Biotechnol.* **2001**, *57*, 744–750. [CrossRef] [PubMed]
96. Bachas-Daunert, P.G.; Sellers, Z.P.; Wei, Y. Detection of halogenated organic compounds using immobilized thermophilic dehalogenase. *Anal. Bioanal. Chem.* **2009**, *395*, 1173–1178. [CrossRef] [PubMed]
97. Yien Jun, L.; Sie Yon, L.; Mubarak, N.; Han Bing, C.; Pan, S.; Danquah, M.K.; Abdullah, E.C.; Khalid, M. An overview of immobilized enzyme technologies for dye and phenolic removal from wastewater. *J. Environ. Health Sci. Eng.* **2019**. [CrossRef]
98. Maier, J.; Kandelbauer, A.; Erlacher, A.; Cavaco-Paulo, A.; Gübitz, G.M. A New Alkali-Thermostable Azoreductase from Bacillus sp. Strain SF. *Appl. Environ. Microbiol.* **2004**, *70*, 837–844. [CrossRef]
99. Pandey, A.; Singh, P.; Iyengar, L. Bacterial decolorization and degradation of azo dyes. *Int. Biodeterior. Biodegradation* **2007**, *59*, 73–84. [CrossRef]

100. El Harfi, S.; El Harfi, A. Classifications, properties and applications of textile dyes: A review. *Appl. J. Environ. Eng. Sci.* **2017**, *3*, 3.
101. Ghoreishi, S.M.; Haghighi, R. Chemical catalytic reaction and biological oxidation for treatment of non-biodegradable textile effluent. *Chem. Eng. J.* **2003**, *95*, 163–169. [CrossRef]
102. Crini, G. Non-conventional low-cost adsorbents for dye removal: A review. *Bioresour. Technol.* **2006**, *97*, 1061–1085. [CrossRef] [PubMed]
103. Lončar, N.; Colpa, D.I.; Fraaije, M.W. Exploring the biocatalytic potential of a DyP-type peroxidase by profiling the substrate acceptance of Thermobifida fusca DyP peroxidase. *Tetrahedron* **2016**, *72*, 7276–7281. [CrossRef]
104. Deive, F.J.; Domínguez, A.; Barrio, T.; Moscoso, F.; Morán, P.; Longo, M.A.; Sanromán, M.A. Decolorization of dye Reactive Black 5 by newly isolated thermophilic microorganisms from geothermal sites in Galicia (Spain). *J. Hazard. Mater.* **2010**, *182*, 735–742. [CrossRef] [PubMed]
105. Al-Balawi, T.H.A.-K.; Wood, A.L.; Solis, A.; Cooper, T.; Barabote, R.D. Anoxybacillus sp. strain UARK-01, a new thermophilic soil bacterium with hyperthermostable alkaline laccase activity. *Curr. Microbiol.* **2017**, *74*, 762–771. [CrossRef]
106. Gianolini, J.E.; Britos, C.N.; Mulreedy, C.B.; Trelles, J.A. Hyperstabilization of a thermophile bacterial laccase and its application for industrial dyes degradation. *3 Biotech* **2020**, *10*, 1–7. [CrossRef]
107. Kumari, A.; Kishor, N.; Guptasarma, P. Characterization of a mildly alkalophilic and thermostable recombinant Thermus thermophilus laccase with applications in decolourization of dyes. *Biotechnol. Lett.* **2018**, *40*, 285–295. [CrossRef] [PubMed]
108. Liu, H.; Cheng, Y.; Du, B.; Tong, C.; Liang, S.; Han, S.; Zheng, S.; Lin, Y. Overexpression of a novel thermostable and chloride-tolerant laccase from Thermus thermophilus SG0. 5JP17-16 in Pichia pastoris and its application in synthetic dye decolorization. *PLoS ONE* **2015**, *10*, e0119833.
109. Navas, L.E.; Carballo, R.; Levin, L.; Berretta, M.F. Fast decolorization of azo dyes in alkaline solutions by a thermostable metal-tolerant bacterial laccase and proposed degradation pathways. *Extremophiles* **2020**, *24*, 705–719. [CrossRef] [PubMed]
110. Jeon, S.-J.; Park, J.-H. Refolding, characterization, and dye decolorization ability of a highly thermostable laccase from Geobacillus sp. JS12. *Protein Expr. Purif.* **2020**, *173*, 105646. [CrossRef]
111. Wang, J.; Chang, F.; Tang, X.; Li, W.; Yin, Q.; Yang, Y.; Hu, Y. Bacterial laccase of Anoxybacillus ayderensis SK3-4 from hot springs showing potential for industrial dye decolorization. *Ann. Microbiol.* **2020**, *70*, 1–9. [CrossRef]
112. Sadeghian, I.; Rezaie, Z.; Rahmatabadi, S.S.; Hemmati, S. Biochemical insights into a novel thermo/organo tolerant bilirubin oxidase from Thermosediminibacter oceani and its application in dye decolorization. *Process Biochem.* **2020**, *88*, 38–50. [CrossRef]
113. Chen, G.; An, X.; Feng, L.; Xia, X.; Zhang, Q. Genome and transcriptome analysis of a newly isolated azo dye degrading thermophilic strain Anoxybacillus sp. *Ecotoxicol. Environ. Saf.* **2020**, *203*, 111047. [CrossRef]
114. Chen, Y.; Feng, L.; Li, H.; Wang, Y.; Chen, G.; Zhang, Q. Biodegradation and detoxification of Direct Black G textile dye by a newly isolated thermophilic microflora. *Bioresour. Technol.* **2018**, *250*, 650–657. [CrossRef]
115. Sarsaiya, S.; Jain, A.; Kumar Awasthi, S.; Duan, Y.; Kumar Awasthi, M.; Shi, J. Microbial dynamics for lignocellulosic waste bioconversion and its importance with modern circular economy, challenges and future perspectives. *Bioresour. Technol.* **2019**, *291*, 121905. [CrossRef]
116. Bilal, M.; Wang, Z.; Cui, J.; Ferreira, L.F.R.; Bharagava, R.N.; Iqbal, H.M.N. Environmental impact of lignocellulosic wastes and their effective exploitation as smart carriers–A drive towards greener and eco-friendlier biocatalytic systems. *Sci. Total Environ.* **2020**, *722*, 137903. [CrossRef]
117. Chandra, R.; Yadav, S.; Kumar, V. Microbial Degradation of Lignocellulosic Waste and Its Metabolic Products. *Environ. Waste Manag.* **2015**, 263–312. [CrossRef]
118. Houfani, A.A.; Anders, N.; Spiess, A.C.; Baldrian, P.; Benallaoua, S. Insights from enzymatic degradation of cellulose and hemicellulose to fermentable sugars—A review. *Biomass Bioenergy* **2020**, *134*, 13. [CrossRef]
119. Bugg, T.D.H.; Ahmad, M.; Hardiman, E.M.; Rahmanpour, R. Pathways for degradation of lignin in bacteria and fungi. *Nat. Prod. Rep.* **2011**, *28*, 1883–1896. [CrossRef]
120. Jeffries, T.W. Biodegradation of lignin and hemicelluloses. In *Biochemistry of Microbial Degradation*; Ratledge, C., Ed.; Springer: Dordrecht, The Netherlands, 1994; pp. 233–277, ISBN 978-94-010-4738-8.
121. Valdivia, M.; Galan, J.L.; Laffarga, J.; Ramos, J. Biofuels 2020: Biorefineries based on lignocellulosic materials. *Microb. Biotechnol.* **2016**, *9*, 585–594. [CrossRef] [PubMed]
122. Nair, R.B.; Lennartsson, P.R.; Taherzadeh, M.J. Bioethanol production from agricultural and municipal wastes. In *Current Developments in Biotechnology and Bioengineering*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 157–190.
123. Aditiya, H.B.; Mahlia, T.M.I.; Chong, W.T.; Nur, H.; Sebayang, A.H. Second generation bioethanol production: A critical review. *Renew. Sustain. Energy Rev.* **2016**, *66*, 631–653. [CrossRef]
124. Gomes, I.; Gomes, J.; Steiner, W. Highly thermostable amylase and pullulanase of the extreme thermophilic eubacterium Rhodothermus marinus: Production and partial characterization. *Bioresour. Technol.* **2003**, *90*, 207–214. [CrossRef]
125. Elleuche, S.; Schäfers, C.; Blank, S.; Schröder, C.; Antranikian, G. Exploration of extremophiles for high temperature biotechnological processes. *Curr. Opin. Microbiol.* **2015**, *25*, 113–119. [CrossRef]
126. Bibra, M.; Wang, J.; Squillace, P.; Pinkelman, R.; Papendick, S.; Schneiderman, S.; Wood, V.; Amar, V.; Kumar, S.; Salem, D. Biofuels and value-added products from extremophiles. In *Advances in Biotechnology*; Nawani, N.N., Khetmalas, M., Razdan, P.N., Pandey, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2014; pp. 17–51.

127. Thapa, S.; Mishra, J.; Arora, N.; Mishra, P.; Li, H.; O'Hair, J.; Bhatti, S.; Zhou, S. Microbial cellulolytic enzymes: Diversity and biotechnology with reference to lignocellulosic biomass degradation. *Rev. Environ. Sci. Bio/Technology* **2020**, *19*, 621–648. [CrossRef]
128. Chauhan, P.S.; Goradia, B.; Saxena, A. Bacterial laccase: Recent update on production, properties and industrial applications. *3 Biotech* **2017**, *7*, 1–20. [CrossRef]
129. Finore, I.; Lama, L.; Poli, A.; Di Donato, P.; Nicolaus, B. Biotechnology implications of extremophiles as life pioneers and wellspring of valuable biomolecules. In *Microbial Factories*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 193–216.
130. Rastogi, M.; Shrivastava, S. Recent advances in second generation bioethanol production: An insight to pretreatment, saccharification and fermentation processes. *Renew. Sustain. Energy Rev.* **2017**, *80*, 330–340. [CrossRef]
131. Xu, F. Laccase. In *Encyclopedia of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation*; Flickinger, M.C., Drew, S.W., Eds.; Wiley: New York, NY, USA, 1999; pp. 1545–1554.
132. Sharma, V.; Ayothiraman, S.; Dhakshinamoorthy, V. Production of highly thermo-tolerant laccase from novel thermophilic bacterium Bacillus sp. PC-3 and its application in functionalization of chitosan film. *J. Biosci. Bioeng.* **2019**, *127*, 672–678. [CrossRef] [PubMed]
133. Atalah, J.; Cáceres-Moreno, P.; Espina, G.; Blamey, J.M. Thermophiles and the applications of their enzymes as new biocatalysts. *Bioresour. Technol.* **2019**, *280*, 478–488. [CrossRef]
134. Haque, R.U.; Paradisi, F.; Allers, T. Haloferax volcanii for biotechnology applications: Challenges, current state and perspectives. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 1371–1382. [CrossRef]
135. Navas, L.E.; Martínez, F.D.; Taverna, M.E.; Fetherolf, M.M.; Eltis, L.D.; Nicolau, V.; Estenoz, D.; Campos, E.; Benintende, G.B.; Berretta, M.F. A thermostable laccase from Thermus sp. 2.9 and its potential for delignification of Eucalyptus biomass. *AMB Express* **2019**, *9*, 1–10. [CrossRef] [PubMed]
136. Kim, H.-W.; Lee, S.-Y.; Park, H.; Jeon, S.-J. Expression, refolding, and characterization of a small laccase from Thermus thermophilus HJ6. *Protein Expr. Purif.* **2015**, *114*, 37–43. [CrossRef]
137. Fernandes, A.T.; Soares, C.M.; Pereira, M.M.; Huber, R.; Grass, G.; Martins, L.O. A robust metallo-oxidase from the hyperthermophilic bacterium Aquifex aeolicus. *FEBS J.* **2007**, *274*, 2683–2694. [CrossRef]
138. Brander, S.; Mikkelsen, J.D.; Kepp, K.P. TtMCO: A highly thermostable laccase-like multicopper oxidase from the thermophilic Thermobaculum terrenum. *J. Mol. Catal. B Enzym.* **2015**, *112*, 59–65. [CrossRef]
139. Aulitto, M.; Fusco, S.; Limauro, D.; Fiorentino, G.; Bartolucci, S.; Contursi, P. Galactomannan degradation by thermophilic enzymes: A hot topic for biotechnological applications. *World J. Microbiol. Biotechnol.* **2019**, *35*, 32. [CrossRef]
140. Berger, E.; Ferreras, E.; Taylor, M.P.; Cowan, D.A. Extremophiles and their use in biofuel synthesis. In *Industrial Biocatalysis*; Pan Stanford Publishing Pte Ltd.: Singapore, 2014; pp. 239–282.
141. Shahid, S.; Tajwar, R.; Akhtar, M.W. A novel trifunctional, family GH10 enzyme from Acidothermus cellulolyticus 11B, exhibiting endo-xylanase, arabinofuranosidase and acetyl xylan esterase activities. *Extremophiles* **2018**, *22*, 109–119. [CrossRef]
142. Yadav, P.; Maharjan, J.; Korpole, S.; Prasad, G.S.; Sahni, G.; Bhattarai, T.; Sreerama, L. Production, purification, and characterization of thermostable alkaline xylanase from Anoxybacillus kamchatkensis NASTPD13. *Front. Bioeng. Biotechnol.* **2018**, *6*, 65. [CrossRef] [PubMed]
143. Marcolongo, L.; La Cara, F.; Del Monaco, G.; Paixão, S.M.; Alves, L.; Marques, I.P.; Ionata, E. A novel β-xylosidase from Anoxybacillus sp. 3M towards an improved agro-industrial residues saccharification. *Int. J. Biol. Macromol.* **2019**, *122*, 1224–1234. [CrossRef]
144. Aulitto, M.; Fusco, S.; Bartolucci, S.; Franzén, C.J.; Contursi, P. Bacillus coagulans MA-13: A promising thermophilic and cellulolytic strain for the production of lactic acid from lignocellulosic hydrolysate. *Biotechnol. Biofuels* **2017**, *10*, 1–15. [CrossRef] [PubMed]
145. Aulitto, M.; Fusco, S.; Nickel, D.B.; Bartolucci, S.; Contursi, P.; Franzén, C.J. Seed culture pre-adaptation of Bacillus coagulans MA-13 improves lactic acid production in simultaneous saccharification and fermentation. *Biotechnol. Biofuels* **2019**, *12*, 45. [CrossRef]
146. Aulitto, M.; Strazzulli, A.; Sansone, F.; Cozzolino, F.; Monti, M.; Moracci, M.; Fiorentino, G.; Limauro, D.; Bartolucci, S.; Contursi, P. Prebiotic properties of Bacillus coagulans MA-13: Production of galactoside hydrolyzing enzymes and characterization of the transglycosylation properties of a GH42 β-galactosidase. *Microb. Cell Fact.* **2021**, *20*, 1–17. [CrossRef]
147. Tian, L.; Papanek, B.; Olson, D.G.; Rydzak, T.; Holwerda, E.K.; Zheng, T.; Zhou, J.; Maloney, M.; Jiang, N.; Giannone, R.J. Simultaneous achievement of high ethanol yield and titer in Clostridium thermocellum. *Biotechnol. Biofuels* **2016**, *9*, 1–11. [CrossRef] [PubMed]
148. Tan, H.; Miao, R.; Liu, T.; Yang, L.; Yang, Y.; Chen, C.; Lei, J.; Li, Y.; He, J.; Sun, Q. A bifunctional cellulase–xylanase of a new Chryseobacterium strain isolated from the dung of a straw-fed cattle. *Microb. Biotechnol.* **2018**, *11*, 381–398. [CrossRef]
149. Widyasti, E.; Shikata, A.; Hashim, R.; Sulaiman, O.; Sudesh, K.; Wahjono, E.; Kosugi, A. Biodegradation of fibrillated oil palm trunk fiber by a novel thermophilic, anaerobic, xylanolytic bacterium Caldicoprobacter sp. CL-2 isolated from compost. *Enzyme Microb. Technol.* **2018**, *111*, 21–28. [CrossRef] [PubMed]
150. Pereira, J.H.; Chen, Z.; McAndrew, R.P.; Sapra, R.; Chhabra, S.R.; Sale, K.L.; Simmons, B.A.; Adams, P.D. Biochemical characterization and crystal structure of endoglucanase Cel5A from the hyperthermophilic Thermotoga maritima. *J. Struct. Biol.* **2010**, *172*, 372–379. [CrossRef]

151. Brumm, P.J.; Gowda, K.; Robb, F.T.; Mead, D.A. The complete genome sequence of hyperthermophile Dictyoglomus turgidum DSM 6724^{TM} reveals a specialized carbohydrate fermentor. *Front. Microbiol.* **2016**, *7*, 1979. [CrossRef]
152. Fusco, F.A.; Ronca, R.; Fiorentino, G.; Pedone, E.; Contursi, P.; Bartolucci, S.; Limauro, D. Biochemical characterization of a thermostable endomannanase/endoglucanase from Dictyoglomus turgidum. *Extremophiles* **2018**, *22*, 131–140. [CrossRef]
153. Aulitto, M.; Fusco, F.A.; Fiorentino, G.; Bartolucci, S.; Contursi, P.; Limauro, D. A thermophilic enzymatic cocktail for galactomannans degradation. *Enzyme Microb. Technol.* **2018**, *111*, 7–11. [CrossRef] [PubMed]
154. McCutchen, C.M.; Duffaud, G.D.; Leduc, P.; Petersen, A.R.H.; Tayal, A.; Khan, S.A.; Kelly, R.M. Characterization of extremely thermostable enzymatic breakers (α-1, 6-galactosidase and β-1, 4-mannanase) from the hyperthermophilic bacterium Thermotoga neapolitana 5068 for hydrolysis of guar gum. *Biotechnol. Bioeng.* **1996**, *52*, 332–339. [CrossRef]
155. Aulitto, M.; Fusco, S.; Fiorentino, G.; Limauro, D.; Pedone, E.; Bartolucci, S.; Contursi, P. Thermus thermophilus as source of thermozymes for biotechnological applications: Homologous expression and biochemical characterization of an α-galactosidase. *Microb. Cell Fact.* **2017**, *16*, 1–10. [CrossRef] [PubMed]
156. Bala, A.; Singh, B. Cellulolytic and xylanolytic enzymes of thermophiles for the production of renewable biofuels. *Renew. Energy* **2019**, *136*, 1231–1244. [CrossRef]
157. Mayer, J.; Kranz, B.; Fischer, L. Continuous production of lactulose by immobilized thermostable β-glycosidase from Pyrococcus furiosus. *J. Biotechnol.* **2010**, *145*, 387–393. [CrossRef] [PubMed]
158. Suleiman, M.; Krüger, A.; Antranikian, G. Biomass-degrading glycoside hydrolases of archaeal origin. *Biotechnol. Biofuels* **2020**, *13*, 1–14. [CrossRef] [PubMed]
159. Ceballos, S.J.; Yu, C.; Claypool, J.T.; Singer, S.W.; Simmons, B.A.; Thelen, M.P.; Simmons, C.W.; VanderGheynst, J.S. Development and characterization of a thermophilic, lignin degrading microbiota. *Process Biochem.* **2017**, *63*, 193–203. [CrossRef]
160. Uthandi, S.; Saad, B.; Humbard, M.A.; Maupin-Furlow, J.A. LccA, an archaeal laccase secreted as a highly stable glycoprotein into the extracellular medium by Haloferax volcanii. *Appl. Environ. Microbiol.* **2010**, *76*, 733–743. [CrossRef]
161. Aulitto, M.; Tom, L.M.; Ceja-Navarro, J.A.; Simmons, B.A.; Singer, S.W. Whole-Genome Sequence of Brevibacillus borstelensis SDM, Isolated from a Sorghum-Adapted Microbial Community. *Microbiol. Resource Announc.* **2020**. [CrossRef]
162. Benedetti, M.; Vecchi, V.; Betterle, N.; Natali, A.; Bassi, R.; Dall'Osto, L. Design of a highly thermostable hemicellulose-degrading blend from Thermotoga neapolitana for the treatment of lignocellulosic biomass. *J. Biotechnol.* **2019**, *296*, 42–52. [CrossRef]
163. Lin, L.; Xu, J. Dissecting and engineering metabolic and regulatory networks of thermophilic bacteria for biofuel production. *Biotechnol. Adv.* **2013**, *31*, 827–837. [CrossRef]
164. Farkas, J.A.; Picking, J.W.; Santangelo, T.J. Genetic techniques for the archaea. *Annu. Rev. Genet.* **2013**, *47*, 539–561. [CrossRef]
165. Farkas, J.; Stirrett, K.; Lipscomb, G.L.; Nixon, W.; Scott, R.A.; Adams, M.W.W.; Westpheling, J. Recombinogenic properties of Pyrococcus furiosus strain COM1 enable rapid selection of targeted mutants. *Appl. Environ. Microbiol.* **2012**, *78*, 4669–4676. [CrossRef]
166. Maezato, Y.; Johnson, T.; McCarthy, S.; Dana, K.; Blum, P. Metal resistance and lithoautotrophy in the extreme thermoacidophile Metallosphaera sedula. *J. Bacteriol.* **2012**, *194*, 6856–6863. [CrossRef]
167. Zhang, C.; Cooper, T.E.; Krause, D.J.; Whitaker, R.J. Augmenting the genetic toolbox for Sulfolobus islandicus with a stringent positive selectable marker for agmatine prototrophy. *Appl. Environ. Microbiol.* **2013**, *79*, 5539–5549. [CrossRef]
168. Suzuki, H.; Murakami, A.; Yoshida, K. Counterselection system for Geobacillus kaustophilus HTA426 through disruption of pyrF and pyrR. *Appl. Environ. Microbiol.* **2012**, *78*, 7376–7383. [CrossRef]
169. Nagy, A. Cre recombinase: The universal reagent for genome tailoring. *Genesis* **2000**, *26*, 99–109. [CrossRef]
170. Togawa, Y.; Nunoshiba, T.; Hiratsu, K. Cre/lox-based multiple markerless gene disruption in the genome of the extreme thermophile Thermus thermophilus. *Mol. Genet. Genomics* **2018**, *293*, 277–291. [CrossRef]
171. Hidaka, Y.; Hasegawa, M.; Nakahara, T.; Hoshino, T. The entire population of Thermus thermophilus cells is always competent at any growth phase. *Biosci. Biotechnol. Biochem.* **1994**, *58*, 1338–1339. [CrossRef]
172. Blesa, A.; Baquedano, I.; Quintáns, N.G.; Mata, C.P.; Castón, J.R.; Berenguer, J. The transjugation machinery of Thermus thermophilus: Identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet.* **2017**, *13*, e1006669. [CrossRef] [PubMed]
173. Swarts, D.C.; Jore, M.M.; Westra, E.R.; Zhu, Y.; Janssen, J.H.; Snijders, A.P.; Wang, Y.; Patel, D.J.; Berenguer, J.; Brouns, S.J.J. DNA-guided DNA interference by a prokaryotic Argonaute. *Nature* **2014**, *507*, 258–261. [CrossRef]
174. Kim, Y.G.; Cha, J.; Chandrasegaran, S. Hybrid restriction enzymes: Zinc finger fusions to Fok I cleavage domain. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1156–1160. [CrossRef]
175. Li, T.; Liu, B.; Spalding, M.H.; Weeks, D.P.; Yang, B. High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* **2012**, *30*, 390–392. [CrossRef] [PubMed]
176. Cava, F.; Laptenko, O.; Borukhov, S.; Chahlafi, Z.; Blas-Galindo, E.; Gómez-Puertas, P.; Berenguer, J. Control of the respiratory metabolism of Thermus thermophilus by the nitrate respiration conjugative element NCE. *Mol. Microbiol.* **2007**, *64*, 630–646. [CrossRef] [PubMed]
177. Ishino, Y.; Shinagawa, H.; Makino, K.; Amemura, M.; Nakata, A. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J. Bacteriol.* **1987**, *169*, 5429–5433. [CrossRef]
178. Pourcel, C.; Salvignol, G.; Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **2005**, *151*, 653–663. [CrossRef]

179. Mohanraju, P.; Makarova, K.S.; Zetsche, B.; Zhang, F.; Koonin, E.V.; Van der Oost, J. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **2016**, *353*, aad5147. [CrossRef]
180. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007**, *315*, 1709–1712. [CrossRef] [PubMed]
181. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **2012**, *337*, 816–821. [CrossRef] [PubMed]
182. Pennisi, E. The CRISPR Craze. *Science* **2013**, *341*, 833–836. [CrossRef] [PubMed]
183. Mougiakos, I.; Mohanraju, P.; Bosma, E.F.; Vrouwe, V.; Bou, M.F.; Naduthodi, M.I.S.; Gussak, A.; Brinkman, R.B.L.; Van Kranenburg, R.; Van Der Oost, J. Characterizing a thermostable Cas9 for bacterial genome editing and silencing. *Nat. Commun.* **2017**, *8*, 1–11. [CrossRef]
184. Mohanraju, P.; Mougiakos, I.; Albers, J.; Mabuchi, M.; Fuchs, R.T.; Curcuru, J.L.; Van Kranenburg, R.; Robb, G.B.; Van Der Oost, J. Development of a Cas12a-Based Genome Editing Tool for Moderate Thermophiles. *Cris. J.* **2021**, *4*, 82–91. [CrossRef]
185. Swarts, D.C.; Jinek, M. Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, e1481. [CrossRef]

Chapter 3 – A Hyperthermoactive-Cas9 Editing Tool Reveals the Role of a Unique Arsenite Methyltransferase in the Arsenic Resistance System of *Thermus thermophilus* HB27

This chapter focuses on understanding the role of *Tt*ArsM, a SAM-dependent arsenite methyltransferase, in the arsenic resistance system of *T. thermophilus* HB27. *Tt*ArsM was identified by a pull-down proteomic approach and shown to be involved in arsenic detoxification. The research described in the paper had three main objectives: **i**) the functional characterization of *Tt*ArsM and its regulation, **ii**) the setup of a novel genetic system for *T. thermophilus* based on CRISPR-Cas9 technology **iii**) the development of a whole-cell arsenic bioreporter using mutant strain of *T. thermophilus*.

Biochemical studies revealed the unique catalytic mechanism of *Tt*ArsM and its specific activity in the methylation of arsenite. Regulatory studies elucidated the role of *Tt*SmtB in *Tt*ArsM transcription and revealed protein-protein interactions that influence transcriptional or post-transcriptional control.
A genome editing tool based on CRISPR-Cas9 technology was optimized at 65 °C in *T. thermophilus* HB27 and this method was used to confirm the *in vivo* role of *Tt*ArsM in arsenic tolerance.
The use of the newly developed CRISPR-Cas9 genome editing tool allowed the generation of a double mutant strain expressing the thermostable yellow fluorescent protein (sYFP) with enhanced sensitivity to arsenic ions, demonstrating its potential as a highly responsive bioreporter.

In conclusion, this work provides comprehensive insights into *Tt*ArsM, contributing to a deeper understanding of the arsenic resistance system in *T. thermophilus* HB27 and highlighting the powerful potential of genome editing in both basic and applied research.

AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

A Hyperthermoactive-Cas9 Editing Tool Reveals the Role of a Unique Arsenite Methyltransferase in the Arsenic Resistance System of *Thermus thermophilus* HB27

Giovanni Gallo,[a,d] Ioannis Mougiakos,[b,e] Mauricio Bianco,[b] Miriam Carbonaro,[a] Andrea Carpentieri,[c] Anna Illiano,[f] Pietro Pucci,[f] Simonetta Bartolucci,[a] John van der Oost,[b] Gabriella Fiorentino[a,d,g]

[a]Department of Biology, University of Naples Federico II, Naples, Italy
[b]Laboratory of Microbiology, Wageningen University & Research, Wageningen, The Netherlands
[c]Department of Chemical Sciences, University of Naples Federico II, Naples, Italy
[d]Consiglio Nazionale delle Ricerche CNR, Institute of Polymers, Composites and Biomaterials (IPCB), Pozzuoli, Naples, Italy
[e]Helmholtz Institute for RNA-based Infection Research (HIRI), Helmholtz-Centre for Infection Research (HZI), Würzburg, Germany
[f]CEINGE-Biotecnologie avanzate, Naples, Italy
[g]Task Force on Microbiome Studies, University of Naples Federico II, Naples, Italy

Giovanni Gallo and Ioannis Mougiakos contributed equally to this work. Author order was determined alphabetically.

ABSTRACT Arsenic detoxification systems can be found in a wide range of organisms, from bacteria to humans. In a previous study, we discovered an arsenic-responsive transcriptional regulator in the thermophilic bacterium *Thermus thermophilus* HB27 (*Tt*SmtB). Here, we characterize the arsenic resistance system of *T. thermophilus* in more detail. We employed *Tt*SmtB-based pulldown assays with protein extracts from cultures treated with arsenate and arsenite to obtain an *S*-adenosyl-L-methionine (SAM)-dependent arsenite methyltransferase (*Tt*ArsM). *In vivo* and *in vitro* analyses were performed to shed light on this new component of the arsenic resistance network and its peculiar catalytic mechanism. Heterologous expression of *TtarsM* in *Escherichia coli* resulted in arsenite detoxification at mesophilic temperatures. Although *Tt*ArsM does not contain a canonical arsenite binding site, the purified protein does catalyze SAM-dependent arsenite methylation with formation of monomethylarsenites (MMAs) and dimethylarsenites (DMAs). In addition, *in vitro* analyses confirmed the unique interaction between *Tt*ArsM and *Tt*SmtB. Next, a highly efficient ThermoCas9-based genome-editing tool was developed to delete the *Tt*ArsM-encoding gene on the *T. thermophilus* genome and to confirm its involvement in the arsenite detoxification system. Finally, the *TtarsX* efflux pump gene in the *T. thermophilus* Δ*TtarsM* genome was substituted by a gene encoding a stabilized yellow fluorescent protein (sYFP) to create a sensitive genome-based bioreporter system for the detection of arsenic ions.

IMPORTANCE We here describe the discovery of an unknown protein by using a proteomics approach with a transcriptional regulator as bait. Remarkably, we successfully obtained a novel type of enzyme through the interaction with a transcriptional regulator controlling the expression of this enzyme. Employing this strategy, we isolated *Tt*ArsM, the first thermophilic prokaryotic arsenite methyltransferase, as a new enzyme of the arsenic resistance mechanism in *T. thermophilus* HB27. The atypical arsenite binding site of *Tt*ArsM categorizes the enzyme as the first member of a new arsenite methyltransferase type, exclusively present in the *Thermus* genus. The enzyme methylates arsenite-producing MMAs and DMAs. Furthermore, we developed an hyperthermophilic Cas9-based genome-editing tool, active up to 65°C. The tool allowed us to perform highly efficient, marker-free modifications (either gene deletion or insertion) in the *T. thermophilus* genome. With these modifications, we confirmed the critical role of *Tt*ArsM in the arsenite detoxification system and developed a sensitive whole-cell bioreporter for

Invited Editor Erika Kothe, Friedrich-Schiller-Universitat

Editor Reinhard Fischer, Karlsruhe Institute of Technology (KIT)

Address correspondence to Gabriella Fiorentino, fiogabri@unina.it, or John van der Oost, john.vanderoost@wur.nl.

Received 25 September 2021
Accepted 28 October 2021
Published 7 December 2021

arsenic ions. We anticipate that the developed tool can be easily adapted for editing the genomes of other thermophilic bacteria, significantly boosting fundamental and metabolic engineering in hyperthermophilic microorganisms.

KEYWORDS arsenic resistance, thermoresistance, extremophiles, *Thermus thermophilus*, CRISPR-Cas9 genome editing, genetic tool, bioreporter

Arsenic is the most abundant environmental toxic element which enters the biosphere mainly from geochemical and (to a lesser extent) anthropogenic sources such as herbicides, growth promoters for livestock, and industrial activities (1). Arsenic has two relevant oxidation states, trivalent arsenite [As(III)] and pentavalent arsenate [As(V)]. Methylated arsenicals include mono- (MA), di- (DMA), and tri- (TMA) methylated forms. In general, trivalent states are more toxic than pentavalent ones, and TMAs are more toxic than inorganic arsenite. Although arsenic is not beneficial for life, it can enter cells through transporters such as aquaglyceroporins. Hence, arsenic detoxification systems can be found in a wide range of organisms, from bacteria to humans. Arsenic resistance genes (*ars*) include genes encoding efflux transporters, redox enzymes, methyltransferases, transcriptional repressors, and biosynthetic pathways for arsenosugars and arsenolipids (2, 3). The identification and characterization of these pathways have attracted the attention of fundamental, evolutionary, and biotechnological research (4–6).

Microorganisms have been exposed to arsenic since the origin of life and consequently have evolved arsenic resistance systems, encoded by genes generally clustered in operons (7). The organization and number of the operons of arsenic resistance genes are highly variable between different species (8, 9), reflecting differences in the level of arsenic resistance. The key players of arsenic detoxification are (i) arsenate reductases (ArsC) that reduce intracellular arsenate to arsenite (10), (ii) efflux permeases responsible for arsenite transport outside the cell (11), and (iii) transcriptional repressors that are generally metalloregulatory proteins of the ArsR/SmtB family (12). In addition, arsenite can be methylated by arsenite *S*-adenosylmethionine methyltransferases (ArsM) into MMAs, DMAs, and TMAs (13), after which they can passively leave the cell or be extruded by methylarsenite-specific efflux permease (ArsP) (14).

Based on recent molecular clock analyses, it has been concluded that arsenite efflux and arsenite methylation represented the core of microbial arsenic resistance systems before the rise of atmospheric oxygen (15). In such primordial anoxic environments, methyl-arsenicals could also function as antibiotics against competitor microbes. After the rise of atmospheric oxygen, the ArsM enzymes did become primary components of the arsenic detoxification machinery; nevertheless, in some microorganisms, they maintained their antibiotic activity (16, 17).

Hydrothermal hot springs, which can be considered environments with conditions similar to niches of primordial Earth, may contain large amounts of arsenic. In these cases, these hot springs are niches for arsenic-tolerant microorganisms, which play a critical role in the global arsenic biogeochemical cycle (18). Although the resistance mechanisms to inorganic arsenic have been studied in many microorganisms (19), the contribution of organo-arsenical biotransformation in extreme environments is still at a stage of infancy. In this regard, only two algal thermoactive ArsM enzymes have been characterized to date (20).

The thermophilic bacterium *Thermus thermophilus* HB27, originally isolated from a volcanic hot spring in Japan (21), has an unusual genetic organization of its machinery to cope with arsenic toxicity. The currently identified arsenic resistance genes are randomly scattered in its genome (22), complicating the identification of all the genes involved. In previous studies, we elucidated some components of the arsenic resistance system of *T. thermophilus* HB27. We identified and characterized *Tt*SmtB, the metalloregulatory transcriptional repressor that is responsible for the regulation of the arsenic detoxification system. *Tt*SmtB recognizes and firmly binds to operator sequences

in the promoter regions of the arsenite efflux gene (*TtarsX*) (23) and the arsenate reductase gene (*Tarec*) (24, 25), efficiently repressing their transcription in the absence of arsenic ions. *Tt*SmtB and *Tt*ArsX are also involved in cadmium sensing and export, respectively (23). In two *T. thermophilus* HB27 deletion mutant strains (Δ*TtarsX* and Δ*TtsmtB*), tolerance to arsenate, arsenite, and cadmium was significantly reduced compared to the wild-type strain. Although these analyses confirmed the involvement of *Tt*ArsX and *Tt*SmtB in the promiscuous resistance mechanism, the mutant strains could still grow at concentrations of arsenic up to 3 mM (22, 23). Notably, the genome of *T. thermophilus* HB27 is not predicted to express arsenite methyltransferases or arsenite oxidases, suggesting the existence of unidentified component(s) of the *T. thermophilus* HB27 arsenic resistance system that cannot be predicted by *in silico* approaches, highlighting the need to employ an alternative experimental identification method.

Since members of the ArsR/SmtB family are a group of homodimeric proteins with a common HTH-winged helix DNA binding domain and heterogeneous metal binding domain architectures and interaction modes (26), we hypothesized that *Tt*SmtB could even form protein interactions with unknown, functionally related protein partners.

In this study, using an integrated proteomic, biochemical, and genetic approach, we provide a gain of insight into the arsenic resistance system of *T. thermophilus* HB27. We report the discovery of the first *T. thermophilus* HB27 arsenite methyltransferase, *Tt*ArsM. Moreover, we describe the development of a highly efficient, markerless Cas9-based genome-editing tool at temperatures up to 65°C. Using this ThermoCas9 system, we demonstrated the *in vivo* involvement of *Tt*ArsM in arsenite detoxification. The newly developed genome editing tool was further validated by constructing a very sensitive whole-cell bioreporter system in which the *TtarsX* efflux transporter gene was substituted by a gene encoding a thermo-adapted superfolder yellow fluorescent protein (*syfp*) (27).

RESULTS

Exploring the protein-protein interactions of *Tt*SmtB. A combined comparative and functional proteomic approach was employed to identify putative *Tt*SmtB-interacting proteins with a role in arsenite metabolism/detoxification. Purified recombinant His-tagged *Tt*SmtB was bound to a Ni^{2+}-nitrilotriacetic acid (NTA) resin for protein pull-down assays using *T. thermophilus* HB27 cell extracts (CFE) from cultures exposed either to arsenite or arsenate or untreated CFE cultures used as control. SDS-PAGE separation of the pulled-down proteins eluted with 0.5 M imidazole, followed by liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) (28), and comparative analysis of the acquired data resulted in the identification of 51 cytosolic proteins that interact with *Tt*SmtB (Table S1 in the supplemental material). Only five of these proteins are simultaneously present in CFE from cultures exposed to arsenite and arsenate, but not in control CFE from the nonexposed cultures. Among these proteins, TTC0109 (GenPept accession no. AAS80457; UniProt code Q72LF0) was predicted to be involved in the arsenic detoxification system (based on homology to annotated ArsR family transcriptional regulators) and to contain a C-terminal *S*-adenosyl-L-methionine (SAM)-dependent methyltransferase domain (based on homology to annotated methyltransferase domain-containing proteins), suggesting a role in arsenic methylation. To date, there are no annotated arsenite SAM-dependent methyltransferases in the genome of *T. thermophilus* HB27 or the genomes of other thermophilic bacteria. Hence, we selected TTC0109 for further investigation as a potential novel arsenite methyltransferase.

Bioinformatics analysis of TTC0109. BLASTP analysis of TTC0109 translated sequence with sequenced microbial genomes and evolutionary analysis conducted with MEGA X demonstrated that TTC0109 is highly conserved among the members of the *Thermus* genus (Fig. 1A). Moreover, multiple-sequence alignment of TTC0109 with characterized prokaryotic arsenite methyltransferases (Fig. 1B) showed that all the aligned proteins contain a typical Rossman fold (29). This fold contains a GxGxG motif in a loop region, which presumably interacts with the carboxypropyl moiety of SAM, and a highly conserved aspartic

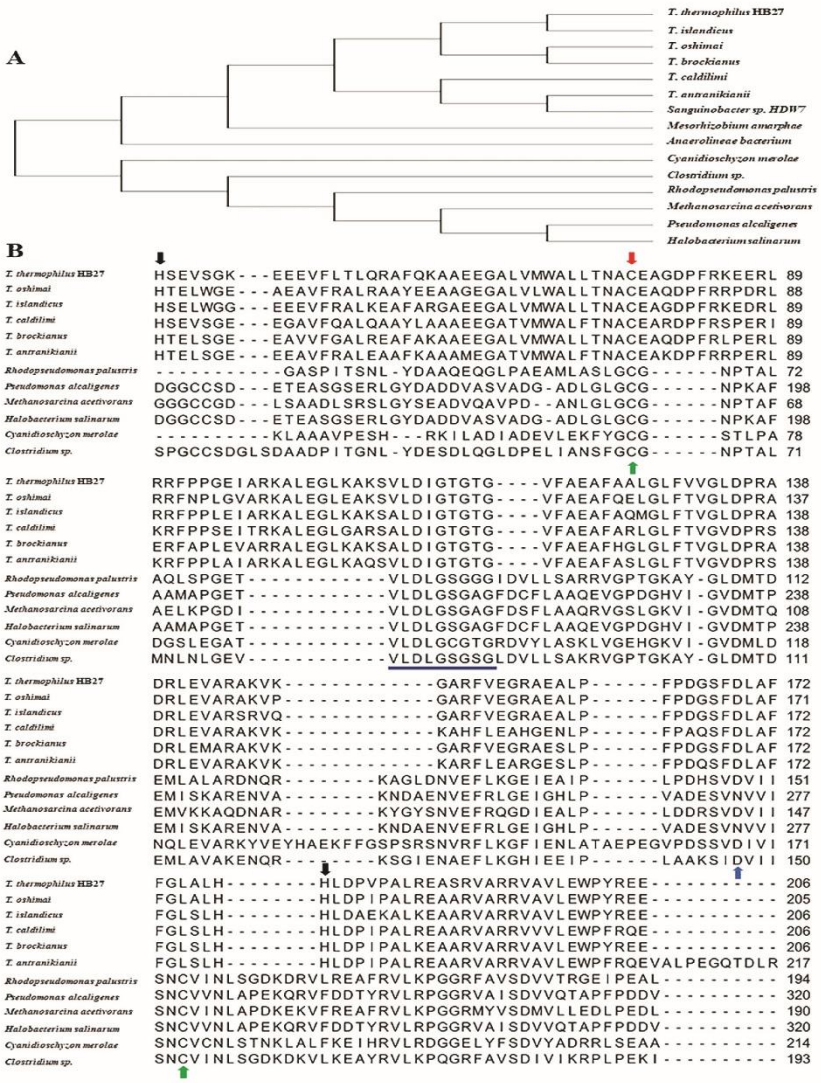

FIG 1 (A) Phylogenetic tree of archaeal and bacterial arsenite methyltransferases, SAM-dependent methyltransferases, and methyltransferase domain-containing proteins. The aminoacidic sequences used for the construction of the phylogenetic tree are *Tt*ArsM from *T. thermophilus* HB27;

(Continued on next page)

acid residue at the end of the β2 strand, which forms hydrogen bonds with the ribose hydroxyls of the cofactor (30). In the case of TTC0109, the predicted GxGxG motif is composed of G114, T115, G116, T117, and G118 residues, and the conserved aspartic residue is D135 (29, 30).

On the other hand, the alignment shows that TTC0109 greatly differs from characterized ArsM proteins in the remaining sequence (Fig. 1B); nonetheless, if TTC0109 is an arsenite methyltransferase, it would be evolutionarily distant from other archaeal and bacterial arsenite methyltransferases, as shown in the phylogenetic tree (Fig. 1A), and therefore, it could belong to a new type of arsenite methyltransferases. Additionally, all the known arsenite methyltransferases, including those in the alignment, possess at least two, usually three, cysteines that are responsible for the binding of arsenite and its subsequent methylation (16, 31). TTC0109 contains a single cysteine residue at position 77, which is perfectly conserved in all the sequences analyzed (Fig. 1B); hence, TTC0109 could be an arsenite methyltransferase with a distinct reaction mechanism.

Reasoning that the TTC0109 structure could provide more information regarding the function of TTC0109, we generated a structural model of the protein and performed molecular docking with arsenite and SAM (Fig. S1). The obtained model predicts that TTC0109 forms homodimers via its N-terminal moiety; molecular docking highlighted that arsenite could be coordinated by two histidines, H40 and H179, while C77 interacts with the methyl group of SAM (Fig. S1A).

Although H40 and H179 residues are not conserved in characterized ArsM proteins, they are maintained at an identical position in the translated genomes of all *Thermus* species (five of them are shown in Fig. 1B); moreover, H40 is encompassed in a sequence motif (34-YRVFPTHSE-42) that shares 45% identity (underlined) with a sequence motif (101-YRLADRHVE-109) at the C terminus of the *Tt*SmtB metal binding site (32), strengthening the hypothesis that H40 could be involved in arsenite binding and suggesting an evolutionary connection between the two proteins.

We proceeded with the generation of structural models of mutant proteins in which the amino acids H40 or H179 were replaced with alanine residues producing TTC0109 H40A and TTC0109 H179A and C77 replaced with a serine residue generating TTC0109 C77S (Fig. S1B to D). The predicted models showed that the substitution of either H40 or H179 with an alanine residue altered the three-dimensional structure of TTC0109, whereas the effect of the cysteine-to-serine substitution had only a minimal effect (Fig. S1C and D), supporting the hypothesis that this residue could have a functional role.

TTC0109 is a novel arsenite methyltransferase. A recombinant His-tagged version of TTC0109 was produced and purified to homogeneity from *Escherichia coli* BL21-CodonPlus (DE3)-RIL cells transformed with pET30b(+)/*TtarsM* (predicted mass, 29.1 kDa) (Fig. S2A). Gel filtration chromatography analysis agreed with the *in silico*-predicted dimeric configuration of the protein, showing that the homodimer has a mass of approximately 64.5 kDa (Fig. S2B). To determine whether TTC0109 had arsenite methyltransferase activity, a coupled spectrophotometric enzymatic assay based on the formation of *S*-adenosylhomocysteine (SAH) from SAM after transfer of methyl group(s) on the substrate was employed (33, 34). In this case, the acceptor of methyl groups was As(III). SAH is degraded by SAH nucleosidase into *S*-ribosylhomocysteine

FIG 1 Legend (Continued)
SAM-dependent methyltransferase from five members of the *Thermus* genus (*T. islandicus*, *T. caldilimi*, *T. antranikianii*, *T. oshimai*, and *T. brockianus*), *Anaerolineae bacterium*, and *Mesorhizobium amorphae*; arsenite methyltransferase from *Clostridium* sp. strain BMX (55), *Methanosarcina acetivorans*, *Rhodopseudomonas palustris* (13), *Pseudomonas alcaligenes* (38), *Halobacterium salinarum* (56), and *Cyanidioschyzon merolae* (45); and a methyltransferase domain-containing protein from *Sanguinobacter* sp. (B) Multiple-sequence alignment of hypothetical and functionally characterized arsenite methyltransferases (ArsM) with *T. thermophilus* HB27 *Tt*ArsM. The partial alignment includes 5 members of the *Thermus* genus (*T. islandicus*, *T. caldilimi*, *T. antranikianii*, *T. oshimai*, and *T. brockianus*) (98% identity to *Tt*ArsM, sequence aligned from amino acids 40 to 89), *Clostridium* sp. BMX (55) (28% identity to *Tt*ArsM, from 22 to 193), *R. palustris* (13) (32% identity to *Tt*ArsM, from 23 to 168), *M. acetivorans* (39) (29% identity to *Tt*ArsM, from 19 to 190), *P. alcaligenes* (38) (31% identity to *Tt*ArsM, from 149 to 320), *Cyanidioschyzon merolae* (45) (27.7% identity to *Tt*ArsM, from 29 to 214), and *H. salinarum* (56) (25% identity to *Tt*ArsM, from 58 to 294). Red arrow indicates the catalytic cysteine conserved in characterized ArsM and *Tt*ArsM; green arrows indicate two catalytic cysteines conserved in characterized ArsM, but not in *Tt*ArsM; blue arrow indicates the conserved aspartic acid; and the SAM binding domain, which is part of the typical Rossman fold, is underlined in blue. The two histidines of *Tt*ArsM predicted to interact with arsenite are indicated by black arrows.

and adenine; adenine deaminase acts on adenine producing hypoxanthine, which is converted into urate and hydrogen peroxide (H_2O_2) by xanthine oxidase. The rate of production of H_2O_2 is measured by an increase in absorbance at 510 nm with the help of the colorimetric reagent 3,5-dichloro-2-hydroxybenzensulfonic acid (DHBS). Then, arsenite methyltransferase activity was assayed following the increase in absorbance at 510 nm (35). Preliminary assays were set up to assess the thermal stability of the different components, and, consequently, the optimal assay temperature; afterward, the saturating concentrations of SAM and arsenite were determined. Therefore, the optimal assay conditions resulted as follows: 50°C, 200 μM arsenite, 800 μM SAM, and 3.1 μM TTC0109. Under these conditions, the specific arsenite methyltransferase activity of TTC0109 was 4.5 mU/mg (Fig. 2A). For this reason, here, the TTC0109 protein will be denoted as *Tt*ArsM.

In order to determine which products are formed upon As(III) methylation by *Tt*ArsM, we incubated 10 μM *Tt*ArsM with As(III), GSH, and SAM at 65°C for 24 h; the mixture was then solubilized and analyzed by gas chromatography-mass spectrometry (GC-MS) (36). Two sharp peaks, at 2.92 min and 2.96 min, were present in the chromatogram of the product of the *Tt*ArsM reaction (Fig. 2B) but not in the chromatogram of the product of the control reaction (Fig. 2C). These peaks were tentatively attributed to MMAs and DMAs, respectively. Unambiguous identification of each compound was accomplished by comparison of their retention times and fragmentation spectra with those of MMA and DMA standards. Butterfly plots showing the strong correspondence of experimental versus theoretical fragmentation spectra for MMAs and DMAs are depicted in Fig. 2D and E, respectively. The assignment was further confirmed by matching the experimental fragmentation spectra with those publicly available in the NIST 05 mass spectral library. According to NIST guidelines, score values higher than 700 indicate an effective identification; the scores determined for MMA and DMA were 790 and 820, respectively (37). MMAs and DMAs were manually integrated, and the results are summarized in Fig. S2C.

The results of *in vitro* assays using purified *Tt*ArsM protein confirmed the ability of the protein to methylate As(III)-producing mono- and dimethylated arsenic, the latter being the primary product. The oxidation state of the products could not be determined because the reactions were terminated with H_2O_2, which oxidized all arsenicals to pentavalent states.

Since the *in silico* predictions of *Tt*ArsM led to the hypothesis that C77, H40, and H179 were catalytic amino acids, three mutated versions of *TtarsM*, namely, *TtarsM* C77S, *TtarsM* H40A, and *TtarsM* H179A, were constructed and expressed in *E. coli* BL21-CodonPlus(DE3)-RIL cells, and the corresponding *Tt*ArsM mutants were purified (Fig. S3). Although the expression levels of the three mutant proteins are comparable (Fig. S3), it was not possible to perform *in vitro* characterization of purified *Tt*ArsM H40A and *Tt*ArsM H179A, which precipitated in solution after purification; this phenomenon is probably due to protein instability, thus indicating the importance of these amino acids for *Tt*ArsM structure. On the other hand, the purification of soluble *Tt*ArsM C77S protein was possible, albeit with a lower yield than the wild-type *Tt*ArsM. Nonetheless, using the previously mentioned coupled assay, this mutant enzyme did not show any *in vitro* arsenite methyltransferase activity, confirming the C77 residue plays a role in *Tt*ArsM activity (Fig. 2A). These *in vitro* results demonstrated that *Tt*ArsM has arsenite methyltransferase activity, and its distinct active site suggests a novel reaction mechanism compared to other characterized arsenite methyltransferases (13, 38, 39).

***Tt*SmtB interacts with *Tt*ArsM and binds to its promoter.** This is the first study to report the protein-protein interaction of an ArsR/SmtB transcriptional regulator with a member of the arsenic detoxification system as identified by pulldown and mass spectrometry. For this reason, we decided to confirm the physical interaction between *Tt*SmtB and *Tt*ArsM and to investigate the effect of different metals on the *Tt*SmtB-*Tt*ArsM interaction. A coimmunoprecipitation (co-IP) assay was carried out upon incubation of purified *Tt*ArsM and *Tt*SmtB, either in the presence or in the absence of arsenite, arsenate, cadmium, and antimony. The first three ions are *Tt*SmtB effectors, as their interaction weakened the binding to target promoters, while

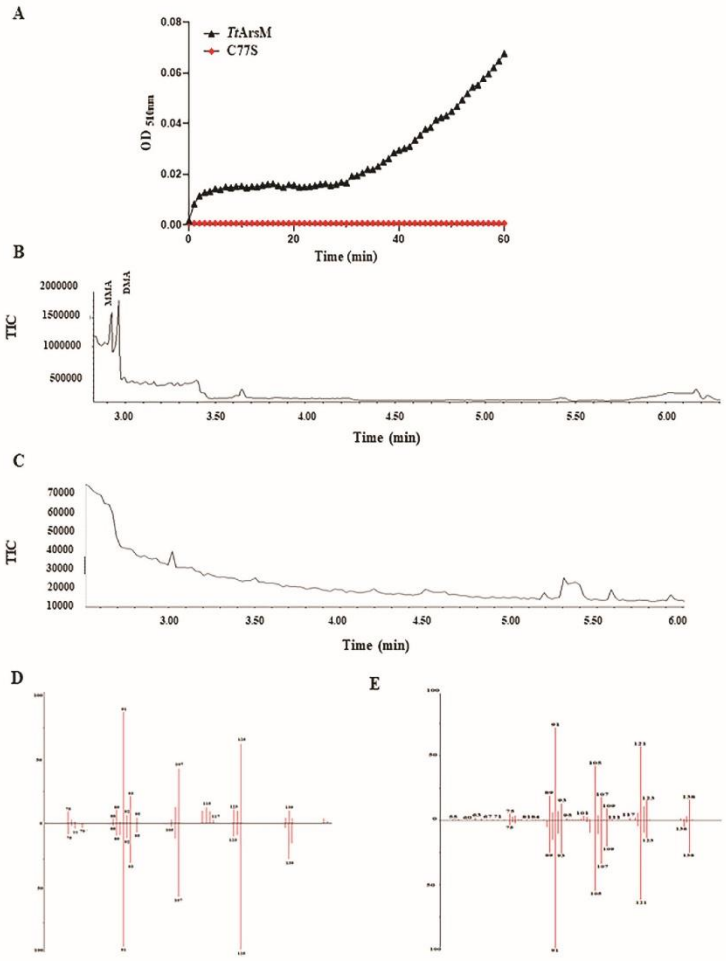

FIG 2 *In vitro* assessment of methyltransferase activity. (A) Arsenite methylation assays. The enzyme-coupled colorimetric assay was carried out in continuous rotation in the presence of 3.1 μM recombinant *Tt*ArsM (black curve) or *Tt*ArsM C77S (red curve), 800 μM SAM, and 200 μM As(III) at 50°C. The absorbance of the reaction mixture was recorded every minute for a total of 1 h. The graph represents the average of three independent experiments, each performed in triplicate. (B) Products of As(III) methylation by purified *Tt*ArsM. Arsenic species were analyzed by GC-MS. Each assay contained 10 μM *Tt*ArsM, 250 μM As(III), 6 mM GSH, and 1 mM SAM, incubated at 65°C for 24 h. (C) Negative control. Both chromatograms were recorded between 2 and 6 min. (D and E) Butterfly plots. The plots show the correspondence between experimental and theoretical fragmentation spectra for MMA and DMA, respectively. The plots confirmed the arsenite methyltransferase activity of *Tt*ArsM.

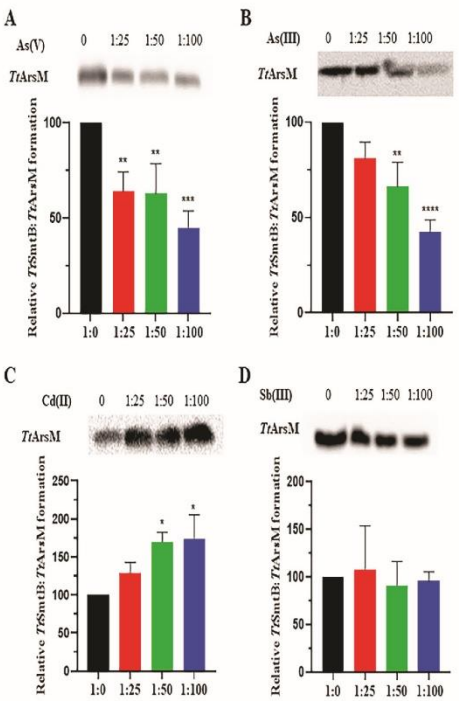

FIG 3 *Tt*SmtB-*Tt*ArsM interaction in the presence of heavy metal ions. Co-IP of *Tt*SmtB-*Tt*ArsM complex with increasing concentrations of arsenate (A), arsenite (B), cadmium (C), and antimony (D). Complexes were immunoprecipitated with anti-*Tt*SmtB antibodies and revealed through Western blotting using anti-His antibodies against the His tag of *Tt*ArsM. (C and D, Bottom) Densitometric analysis of blots of *Tt*SmtB:*Tt*ArsM complex. The intensity of the unchallenged complex was used as a reference. Average values from three biological replicates are shown, with error bars representing standard deviations. Statistical analysis was performed using one-way ANOVA; significant differences are indicated as *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; and ****, $P < 0.0001$.

antimony had no effect on DNA recognition (32). Immunoprecipitation with anti-*Tt*SmtB antibodies, followed by detection of the His-tagged *Tt*ArsM by anti-His-tag antibodies, showed that the two proteins interact and form a complex in the absence of arsenite and arsenate, confirming the existence of physical interaction between them (Fig. 3). No band was detected when the immunoprecipitation was carried out with the unrelated control protein, *Tt*GalA (40). Increasing arsenate and arsenite concentrations negatively affected the stability of *Tt*SmtB-*Tt*ArsM complex; in fact, densitometric analysis of the Western blot revealed up to a 3-fold decrease in the intensity of the band corresponding to the complex (at 1:100 protein/arsenic ratio) in the presence of both arsenate (Fig. 3A) and arsenite (Fig. 3B). Interestingly, the presence of cadmium had the opposite effect, enhancing the band intensity by up to 2-fold (Fig. 3C), suggesting that the interaction of this metal with the complex occurs with a different mechanism. Finally, the presence of antimony had a negligible effect on complex stability, in agreement with previous data showing that this metal ion is not an effector for *Tt*SmtB (32) (Fig. 3D).

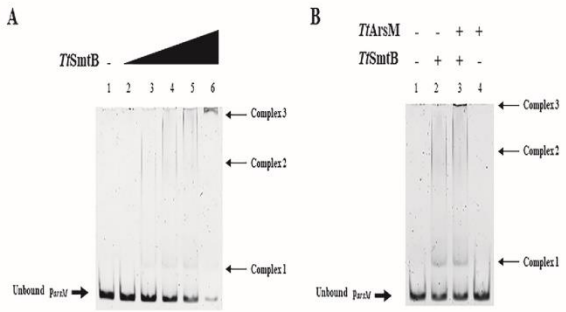

FIG 4 *Tt*SmtB-p_{arsM} interaction. (A) Interaction of p_{arsM} in the presence of increasing concentrations of *Tt*SmtB. Lane 1, negative control; lane 2, 1 μM *Tt*SmtB; lane 3, 2 μM *Tt*SmtB; lane 4, 3 μM *Tt*SmtB; lane 5, 5 μM *Tt*SmtB; and lane 6, 10 μM *Tt*SmtB. (B) EMSA with p_{arsM} in the presence of 3 μM *Tt*SmtB and 3 μM *Tt*ArsM.

Since *Tt*SmtB is the transcriptional repressor of the genes involved in arsenic and cadmium resistance in *T. thermophilus* HB27 (22, 23), we hypothesized that it could also regulate *TtarsM* transcription. Sequence analysis of *TtarsM* promoter (p_{arsM}), a 108-bp-long region upstream of *TtarsM* and encompassing the translation start codon revealed the presence of an inverted repeat region [GAAC(N14)CTTG] between positions −6 and −27 upstream of the start codon. The sequence overlaps −10 and −35 putative basal promoter region, is 100% identical to the *TtarsX* operator recognized by *Tt*SmtB, and matches the consensus binding sites of ArsR/SmtB proteins (41). Hence, we performed electrophoretic mobility shift assay (EMSA) to investigate the capacity of purified *Tt*SmtB to bind to the promoter region of *TtarsM*. *Tt*SmtB binds to p_{arsM} in a concentration-dependent manner, as shown by the gradual formation of lower-mobility complexes and the gradual decrease of residual unbound DNA (Fig. 4A, lanes 2 to 5); at 10 μM protein, the complex hardly enters the gel, suggesting the formation of multiple dimers associated with target DNA (Fig. 4A, lane 6). This observation suggests that by interacting with the regulatory region, *Tt*SmtB controls *TtarsM* transcription in a way comparable to that already reported for other arsenic resistance genes, i.e., the arsenate reductase, the metal ion transporter, and itself (22, 23).

Since the existence of a physical interaction between *Tt*SmtB and *Tt*ArsM was established, we asked whether *Tt*ArsM influenced *Tt*SmtB interaction with p_{arsM}. Therefore, we preincubated 3 μM of both proteins before performing an EMSA under the same conditions described above. Interestingly, when the two proteins are coincubated, shifted bands can be observed (Fig. 4B, lane 3, complex 3) corresponding to complexes of higher molecular weight in comparison to those generated or not by *Tt*SmtB or *Tt*ArsM, respectively (Fig. 4B, lane 2, complex 2, and lane 4); this analysis indicates that *Tt*SmtB-*Tt*ArsM multimeric complexes bind to the promoter and suggests that *Tt*SmtB-*Tt*ArsM protein-protein interaction may function in either transcriptional and posttranscriptional control. Notably, very few studies in bacteria report protein-protein interactions of transcriptional regulators with the product of the genes they regulate (42, 43).

***In vivo* activity of *Tt*ArsM and its mutants in *E. coli*.** Aiming to explore the role of *Tt*ArsM in arsenite resistance *in vivo*, we challenged *E. coli* BL21-CodonPlus (DE3)-RIL strains transformed with plasmids expressing *Tt*ArsM and its catalytic mutants (*Tt*ArsM C77S, *Tt*ArsM H40A, and *Tt*ArsM H179A) to grow in the presence of arsenite. Each recombinant strain was grown in the presence of different arsenite concentrations for 24 h to determine the MIC toward the metal ion. The *Tt*ArsM-expressing strain appeared to be more resistant to arsenite than the control strain (MICs of 6 mM and

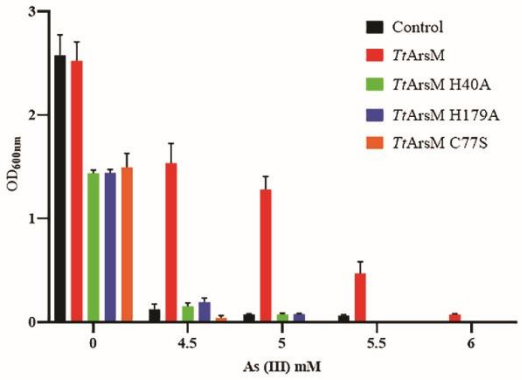

FIG 5 Growth of *E. coli* BL21 strains expressing *Tt*ArsM and its mutants, in the presence of different arsenite concentrations, measured 24 h postinoculation. The strains are *E. coli* BL21/pET30b (black), *E. coli* BL21/pET30/*TtarsM* (red), *E. coli* BL21/pET30/*TtarsM* H40A (green), *E. coli* BL21/pET30/*TtarsM* H179A (blue), and *E. coli* BL21/pET30/*TtarsM* C77S (orange). Average values from three biological replicates are shown, with error bars representing standard deviations.

4.5 mM, respectively). Additionally, the strains expressing mutated *Tt*ArsM were inhibited by the presence of arsenite to the same extent as the control strain (Fig. 5). This shows that the heterologous expression of *Tt*ArsM in *E. coli* increases arsenite resistance even at mesophilic temperatures, indicating the role of *Tt*ArsM in arsenite detoxification. Moreover, the result obtained with the mutant strains demonstrates the role of C77, H40, and H179 in the catalytic function of *Tt*ArsM.

Developing a hyperthermoactive-Cas9 editing tool. We further aimed to investigate *in vivo* the contribution of *Tt*ArsM to the arsenite detoxification mechanism via the deletion of the *TtarsM* gene from the *T. thermophilus* HB27 genome. Nonetheless, the currently available genome editing tool for *T. thermophilus* is time-consuming, not marker-free, and not always efficient (44). For this purpose, we reasoned to develop a marker-free, plasmid-based, homologous recombination (HR) Cas9 counterselection (CS) genome-editing tool for *T. thermophilus* employing ThermoCas9, a thermotolerant and thermoactive Cas9 orthologue (45).

We initially evaluated the targeting efficiency of ThermoCas9 in *T. thermophilus* HB27. Therefore, a set of 3 vectors was constructed, namely, pMK-ThermoCas9-NT, pMK-ThermoCas9-sp1, and pMK-ThermoCas9-sp2 (Fig. 6A) by cloning into the pMK18 vector (46) (i) the codon-harmonized version of the *thermocas9* gene under the transcriptional control of the constitutive *nqo* promoter (44), and (ii) the single guide RNA (sgRNA)-expressing module under the transcriptional control of the constitutive 16S rRNA promoter, either with a nontargeting/control spacer (NT, 5′-CTAGATCCGCAGTAACCCCATGG-3′) or with spacers that target the *TtarsM* gene (sp1, 5′-GGGCGTTGGTGATGTGGGCCCTC-3′, and sp2, 5′-CCACCTCCTCCTCCCGGTAAGGC-3′). The 3 vectors were used to transform *T. thermophilus* HB27, along with pMK-Pnqo-syfp vector (27), as transformation control. The cells were allowed to recover at 70°C before being plated on selective agar plates and incubated overnight at 60°C due to the sensitivity of pMK18 at temperatures above 65°C. The transformation efficiencies of pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2 targeting vectors were significantly reduced compared to the transformation efficiency with the pMK-ThermoCas9-NT nontargeting vector (Fig. 6B). Moreover, the transformation efficiency with the pMK-Pnqo-syfp vector was only slightly higher than the transformation efficiency with the pMK-ThermoCas9-NT vector (Fig. 6B), which could be attributed to the significant

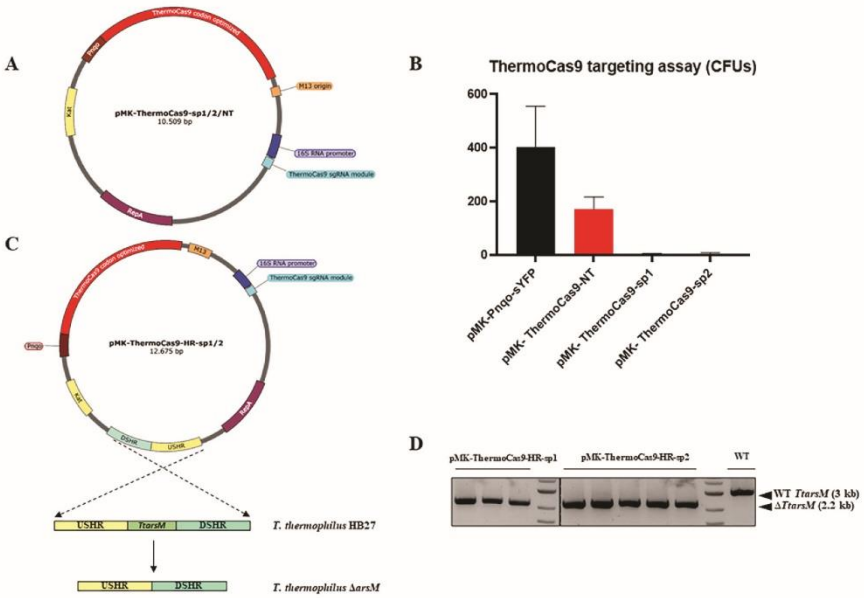

FIG 6 ThermoCas9-based genome engineering in *T. thermophilus* HB27. (A) pMK-ThermoCas9-sp1/2/NT targeting vectors. (B) Graphical representation of the ThermoCas9 targeting assay results (CFU) for assessing the ThermoCas9 toxicity and targeting efficiency in *T. thermophilus* HB27. Average values from three biological replicates are shown, with error bars representing standard deviations. (C) pMK-ThermoCas9-HR-sp1/2 editing vectors, employed for the genomic deletion of the *TtarsM* gene. (D) Agarose gel electrophoresis showing the resulting products from genome-specific colony PCRs on *T. thermophilus* colonies formed from the ThermoCas9-based *TtarsM* deletion process. A wild-type colony was subjected to the same PCR, and the related product is shown here as negative control for *TtarsM* deletion. The expected sizes of the PCR amplification products that correspond to the wild-type and Δ*TtarsM* genotypes are indicated with black arrows.

size difference between the two vectors (7,552 bp and 13,554 bp, respectively). This result indicates that ThermoCas9 is expressed in *T. thermophilus* HB27 cells in an active and not-toxic form, motivating the development of a ThermoCas9-based genome-editing tool.

We set out to develop and test the efficiency of an HR ThermoCas9-based CS genome-editing tool in *T. thermophilus* HB27. For this purpose, we introduced an HR template for the deletion of the *TtarsM* gene into the 3 previously described ThermoCas9 vectors. The HR template was composed of the fused 1-kb upstream and downstream flanking regions of the *TtarsM* gene (Fig. 6C). The three resulting editing vectors, namely, pMK-ThermoCas9-HR-NT, pMK-ThermoCas9-HR-sp1, and pMK-ThermoCas9-HR-sp2, were transformed into *T. thermophilus* HB27 cells, recovered at 70°C, and grown on selective agar plates overnight at 60°C. Colony PCR with genome-specific primers was subsequently employed to screen several colonies for each transformation (Fig. S4; Table S2). None of the colonies from the pMK-ThermoCas9-NT transformation were clean Δ*TtarsM* mutants (0/10 colonies), and only a small number of colonies were mixed wild-type/Δ*TtarsM* mutants (2/10 colonies) (Fig. S4A; Table S2). On the other hand, almost all the screened colonies from the pMK-ThermoCas9-sp1 transformation were clean Δ*TtarsM* mutants (19/19 colonies) (Fig. S4B; Table S2); most of the screened colonies from the pMK-ThermoCas9-sp2 transformation were clean Δ*TtarsM* mutants (13/18 colonies), and the remaining were mixed wild-type/Δ*TtarsM* mutants (5/18 colonies) (Fig. S4C; Table S2); the latter result suggests that less efficient ThermoCas9 targeting is

obtained when employing spacer 2. Subsequently, DNA sequencing on randomly selected clean Δ*TtarsM* mutant colonies was performed to verify the correctness of the genome editing (Fig. S4D; Table S2).

Aiming to test the temperature limit of the developed ThermoCas9-based genome-editing tool, we repeated the editing experiment, increasing the plating temperature to 65°C, corresponding to the temperature limit of the pMK18 backbone for propagation. Under these conditions, the numbers of colonies formed upon transformation with the pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2 were 3 and 5 vectors, respectively, much lower than the corresponding numbers when the plating temperature was 60°C (Table S2). This can be ascribed to the high ThermoCas9 targeting activity at 65°C (45) and decreased vector stability at 65°C. Nonetheless, the DNA sequence of all the screened colonies confirmed that they were clean Δ*TtarsM* mutants, demonstrating the high efficiency of the developed tool at 65°C (Fig. 6D).

Finally, assuming that the curing of the editing plasmid from a Δ*TtarsM* mutant strain would facilitate additional editing steps, we randomly selected a Δ*TtarsM* mutant colony for inoculation in liquid antibiotic-free TM medium for two culturing rounds at 65°C and then plated the cultures on TM agar plates with and without antibiotic. Multiple colonies were found on the antibiotic-free plate and no colonies on the plate supplemented with the antibiotic, demonstrating that the cells were cured from the edited vector. Seven of these colonies were randomly selected and the absence of the plasmid confirmed by colony PCR using *thermoCas9*-specific primers (Fig. S4E).

Therefore, a markerless HR-ThermoCas9-based CS genome-editing tool was developed for *T. thermophilus* HB27, highly efficient at temperatures up to 65°C. Using this tool, a *T. thermophilus* Δ*TtarsM* strain was constructed in less than 10 days (including the plasmid-curing process), expanding the repertoire of available genetic tools for this microorganism and considerably accelerating the required time for editing its genome.

***Tt*ArsM mutant is more sensitive to arsenite.** To compare the arsenic resistance of Δ*TtarsM* to that of wild-type *T. thermophilus* HB27, both strains were grown in TM liquid medium with different arsenite and arsenate concentrations for 24 h (Fig. 7). As expected, the arsenite resistance of Δ*TtarsM* was significantly lower than that of the wild-type strain, with the corresponding MIC values being 18 mM and 40 mM, respectively (Fig. 7A). Moreover, the resistance of the Δ*TtarsM* strain to arsenate was comparable to the wild-type strain (42 mM and 44 mM, respectively) (Fig. 7B), in agreement with its role in arsenite resistance. This result confirmed that the thermoactive arsenite methyltransferase *Tt*ArsM is involved in arsenite detoxification and is a novel component of the arsenic resistance machinery.

Developing a sensitive arsenic bioreporter. In a previous study, we demonstrated that *Tt*ArsX is the arsenic efflux membrane protein of *T. thermophilus* HB27 and reported that a Δ*TtarsX* mutant strain is more sensitive to arsenate and arsenite (23). In this study, we wanted to ascertain whether a strain lacking both *TtarsM* and *TtarsX* would be even more sensitive to arsenic ions than the single-mutant Δ*TtarsX* and Δ*TtarsM* strains and therefore could represent an even better bioreporter strain for arsenic detection.

For this purpose, the HR ThermoCas9-based CS-editing tool was employed to exchange the *TtarsX* gene on the genome of the Δ*TtarsM* strain with the *syfp* reporter gene (27), setting the expression of the encoded thermotolerant yellow fluorescence protein (sYFP) under the control of the arsenic-responsive *TtarsX* promoter (p_{arsX}). The employed editing vector, denoted as pMK-ThermoCas9-HR-*syfp*, contained a spacer that targets the *TtarsX* gene (5′-TTTCGACGGAGGAGGCCTTGGCC-3′) and an HR template composed of the 1-kb upstream flanking genomic region of *TtarsX* followed by *syfp* and the 1-kb downstream flanking genomic region of *TtarsX*. Ten colonies grown after transformation of pMK-ThermoCas9-HR-*syfp* vector into *T. thermophilus* Δ*TtarsM* cells were screened by colony PCR with genome-specific primers and sequenced; eight of them were clean *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) knock-in mutants (Fig. S5), also proving that the developed tool was highly efficient for gene insertions and substitutions.

The double-mutant strain was challenged with different arsenite and arsenate concentrations in TM liquid medium. As shown in Fig. 7A and B, arsenite resistance is strikingly

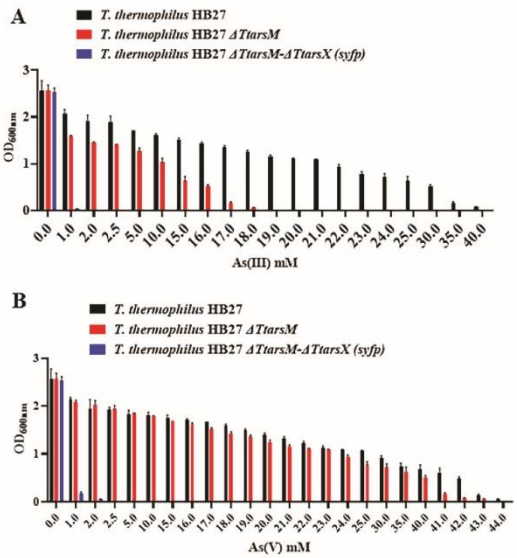

FIG 7 Growth of *T. thermophilus* HB27 (black), *T. thermophilus* HB27 Δ*TtarsM* (red), and *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) (blue) in TM medium in the presence of different concentrations of arsenite (A) and arsenate (B) measured 24 h after inoculation. Average values from three biological replicates are shown, with error bars representing standard deviations.

lower (0.5 mM) than that of the wild-type (40 mM), Δ*TtarsM* (18 mM), and Δ*TtarsX* strains (3 mM) (23). Interestingly, the Δ*TtarsM*-Δ*TtarsX* (*syfp*) strain showed also lower resistance to arsenate than the single-mutant Δ*TtarsX* (2 mM and 3 mM, respectively) (22).

To evaluate the sensitivity to arsenate and arsenite of the whole-cell bioreporter system, exponentially growing cultures of *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) were treated with increasing concentrations of arsenite and arsenate, and the intensity of the emitted fluorescence was compared (Fig. 8). The background fluorescence of the Δ*TtarsM*-Δ*TtarsX* (*syfp*) strain was low, indicating that the system is repressed in the absence of metal ions. Moreover, the developed bioreporter system was able to detect arsenite and arsenate concentrations as low as 0.5 μM (Fig. 8). This performance substantially overtakes the detection limit of the previously developed arsenite and arsenate bioreporter system, which was based on the *T. thermophilus* Δ*TtarsX* strain and plasmid-based expression of the β-galactosidase (23).

DISCUSSION

In this study, we aimed to identify novel proteins involved in the arsenic resistance system of *T. thermophilus* HB27 and employed *Tt*SmtB as a starting point looking beyond its transcriptional regulation activity. As *Tt*SmtB contains a protein interaction domain (32), we set out to identify putative *Tt*SmtB-interacting proteins with a role in arsenic metabolism/detoxification, following an immunoprecipitation and comparative proteomics approach. This strategy led to the discovery of *Tt*ArsM, the first prokaryotic thermoactive arsenite SAM-dependent methyltransferase, evolutionarily distant from other known arsenite methyltransferases.

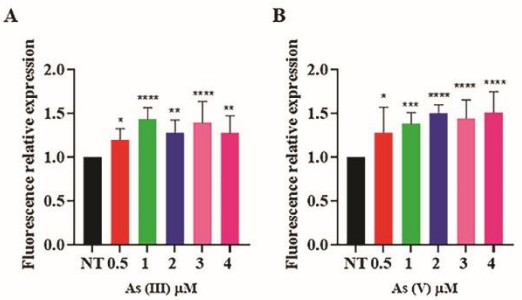

FIG 8 Bioreporter activity. *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) bioreporter strain challenged with increasing concentrations of arsenite (A) and arsenate (B). Average values from three biological replicates are shown, with error bars representing standard deviations. Statistical analysis was performed using one-way ANOVA; significant differences are indicated as *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; and ****, $P < 0.0001$.

The original structure and activity mechanism of *Tt*ArsM were explored *in silico*. Like other arsenite methyltransferases known to date, *Tt*ArsM contains a C-terminal SAM-dependent methyltransferase domain and an N-terminal domain comprised of only one (instead of the usual three) conserved cysteine; the other two catalytic cysteines required for arsenite coordination might be replaced by two histidines identified by docking analysis. Interestingly, H40 is part of an arsenite binding domain typical of many ArsR metalloregulatory proteins (47), and H179 is located close to the cysteine (C177) of the arsenite methyltransferase of the extremophilic alga *Cyanidioschyzon merolae* (48). Moreover, the conservation of these three amino acids among the putative arsenite methyltransferases in the *Thermus* genus supports their possible role in catalysis and suggests an adaptation in this group of microorganisms.

Indeed, the role of the cysteine and histidine residues in the structure-function of *Tt*ArsM predicted *in silico* was demonstrated by site-directed mutagenesis, as the heterologous expression in *E. coli* of *Tt*ArsM mutants conferred lower arsenite resistance than *Tt*ArsM.

To date, only for a few arsenite methyltransferases, the reaction mechanism has been reported (13, 20, 38, 39, 49). Those described possess three, or at least two, cysteine residues present in their catalytic site that can methylate the arsenic sequentially in its trivalent form through alternating reduction and oxidative methylation reactions; of note, different enzymes produce mono-, di-, and trimethylated forms of arsenic in diverse amounts, highlighting a biochemical diversity in the arsenite methylation mechanism (38). The newly discovered *Tt*ArsM, highly conserved within the *Thermus* genus, possesses only one cysteine and can methylate As(III) mainly into DMAs and a smaller amount of MMAs as determined by GC-MS analysis of the products of the *in vitro* assay. To the best of our knowledge, this is the first arsenite methyltransferase functioning with a single cysteine in the active site.

The discovered transcriptional and posttranslational interaction of *Tt*ArsM with the transcriptional regulator *Tt*SmtB was investigated in more detail. It was demonstrated that *Tt*SmtB binds to the promoter region of *TtarsM* and that this binding is stabilized by the *Tt*SmtB-*Tt*ArsM complex. Moreover, co-IP experiments confirmed the interaction of *Tt*SmtB with *Tt*ArsM and showed a reverse correlation between the stability of the complex and arsenic concentration. Presumably, the complex enhances the repression of *TtarsM* transcription in the absence of arsenic ions through a novel mechanism. Hence, through this analysis, we shed light on a novel kind of interaction, rarely described for bacteria, in which the transcriptional repressor of a gene interacts with the protein product of the gene that it regulates (43, 50).

An example of an enzyme that can modulate the transcriptional activity of regulators by protein-protein interaction is reported in the cysteine metabolism of *Bacillus subtilis* where the stable complex formed by CymR (the master regulator of the system) and CysK (*O*-acetyl-L-serine-thiol-lyase) represses the transcription of the genes involved in the cysteine pathway (including the *cysK* gene itself) when cysteine concentration is low. The advantage of this regulatory mechanism is that it employs enzymes that can specifically recognize their substrates or allosteric effectors; thus, enzymes and/or transcriptional regulators can act simultaneously as intracellular molecular sensors and participate in their own transcriptional regulation (42).

The role of *Tt*ArsM in *T. thermophilus* HB27 arsenic detoxification was also demonstrated via the construction and characterization of a Δ*TtarsM* mutant strain. For this purpose, a marker-free homologous recombination and ThermoCas9-based counterselection genome-editing tool was developed which was highly efficient and active at temperatures up to 65°C. Our tool equals the highest reported temperature for a Cas9-based editing tool to date (51, 52). The characterization of the *T. thermophilus* Δ*TtarsM* strain confirmed its expected higher sensitivity to arsenite, but not arsenate, than the wild-type strain. To better define the role of *Tt*ArsM in the context of the already characterized components of the arsenic resistance system, a double mutant was constructed upon exchanging the *TtarsX* efflux pump gene in the *T. thermophilus* Δ*TtarsM* genome with the gene encoding the yellow fluorescent protein; the double mutant resulted as much more sensitive to arsenite and arsenate treatment. Hence, it was demonstrated that *Tt*ArsM and *Tt*ArsX are critical players of the arsenite detoxification system. This is the first example of a successful insertion of a heterologous gene on the *T. thermophilus* genome by genome editing. The double-mutant strain was also considered a sensitive bioreporter for the development of a whole-cell biosensor system. Indeed, it was able to detect arsenite and arsenate concentrations as low as 0.5 μM, showing 40 times higher sensitivity than the previously developed *T. thermophilus* HB27 Δ*TtarsX*- plasmid-based biosensor (23).

In conclusion, this study explores a unique strategy to identify novel enzymes and/or regulative networks in nonmodel bacteria and expands the repertoire of genetic systems for hyperthermophiles. In addition, this work has resulted in a gain of insight into the arsenite/arsenate detoxification mechanism, particularly that of *T. thermophilus*. On top of that, this has allowed us to develop a highly robust and sensitive biosensor.

MATERIALS AND METHODS

***T. thermophilus* HB27 cell extract preparation.** *T. thermophilus* HB27 cultures were grown aerobically at 70°C in TM medium, as previously described (24). Once the cultures reached an optical density at 600 nm (OD_{600}) of 0.5, they were treated either with 8 mM $NaAsO_2$ or with 12 mM NaH_2AsO_4 (Sigma) (the used concentrations were below the previously reported MIC values for arsenate and arsenite) (22, 23) or they remained untreated. Samples were harvested from each culture, either immediately after treatment or 60 min posttreatment. The samples were centrifuged, the precipitates were resuspended in phosphate buffer (20 mM Na_3PO_4, pH 7.5) supplemented with protease inhibitor cocktail (Thermo Scientific), and the resuspended cells were lysed by sonication (10 cycles of 30 s on/30 s off, 40% power; Misonix Sonicator Ultrasonic Processor XL). The lysates were centrifuged and the cell extracts (CFE) used for pulldown assays.

Purification of recombinant *Tt*SmtB, immobilized metal affinity chromatography, and pulldown. C-terminal His-tagged *Tt*SmtB was purified from *E. coli* BL21-CodonPlus (DE3)-RIL cells transformed with the pET28/*Ttsmt*B vector, as previously described (22). Purified C-terminal His-tagged *Tt*SmtB (2 mg) was incubated with 200 μl of Ni^{2+}-NTA resin (Sigma-Aldrich) equilibrated in 20 mM Na_3PO_4, 0.5 M NaCl, and 20 mM imidazole, pH 7.5, for 16 h at 4°C and then washed three times with the same buffer to remove unbound proteins. *T. thermophilus* HB27 CFE, treated with arsenite, treated with arsenate, or not treated were incubated with the functionalized resin (Ni^{2+}-NTA/*Tt*SmtB) for 16 h at 4°C under stirring conditions. Subsequently, the resin was extensively washed, and the interacting proteins were eluted with 20 mM Na_3PO_4, 0.5 M NaCl, and 0.5 M imidazole, pH 7.5. As negative controls, samples of Ni^{2+}-NTA resin not functionalized with *Tt*SmtB were incubated with the same *T. thermophilus* HB27 CFE.

***In situ* hydrolysis and LC-MS/MS analysis.** The fractions eluted from the pulldown process were analyzed by 15% SDS-PAGE and *in situ* hydrolyzed for mass spectrometry analysis. Specifically, monodimensional SDS-PAGE gel was colored with Coomassie brilliant blue; the revealed bands were cut and destained with 100 μl of 0.1 M ammonium bicarbonate (AMBIC) and 130 μl of acetonitrile (ACN). Each band was hydrolyzed *in situ* with 0.1 μg/μl trypsin in 10 mM AMBIC and incubated at first for 1.5 h at 4°C and then for an additional 16 h at 37°C. The hydrolysis reactions were stopped by adding acetonitrile and 0.1% formic acid; then, the samples were filtered and dried in a Savant vacuum centrifuge before

being analyzed by LC-MS/MS mass spectrometry. In detail, before analysis, the samples were dissolved in 10 μl of 0.1% formic acid, and 5 μl were directly loaded into the instrument. Reverse-phase capillary liquid chromatography (HPLC 1200 system experiments), followed by MS analysis, was performed using a binary pump system connected to a nanospray source of the mass spectrometer (28, 53). The latter is represented by a hybrid quadrupole time of flight (Q-TOF) spectrometer (MS Chip 6520 QTOF) equipped with a chip (Agilent Technologies).

***In silico* analysis.** Analysis of the LC-MS/MS data, using Mascot software (http://www.matrixscience.com/search_form_select.html), allowed the identification of putative *Tt*SmtB-interacting proteins. Among these proteins, TTC0109 (UniProt code Q72LF0), here named *Tt*ArsM, was further analyzed using the UniProt database (http://www.uniprot.org); homologous proteins and conserved domains were identified by performing a BLAST analysis (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

A phylogenetic tree of archaeal and bacterial arsenite methyltransferases, SAM-dependent methyltransferases, and methyltransferase domain-containing proteins, including *Tt*ArsM from *T. thermophilus* HB27, was conducted in MEGA X (54). The amino acidic sequences used for the construction of the phylogenetic tree are *Tt*ArsM from *T. thermophilus* HB27; arsenite methyltransferase from *Rhodopseudomonas palustris* (13), *Methanosarcina acetivorans*, *Clostridium* sp. strain BMX (55), *Halobacterium salinarum* (56), *Pseudomonas alcaligenes* (38), and *Cyanidioschyzon merolae* (48); a SAM-dependent methyltransferase from 5 members of the *Thermus* genus (*T. islandicus*, *T. caldilimi*, *T. antranikianii*, *T. oshimai*, and *T. brockianus*); a SAM-dependent methyltransferase from *Mesorhizobium amorphae* and *Anaerolineae bacterium*; and a methyltransferase domain-containing protein from *Sanguinobacter* sp. Phylogenetic reconstruction was accomplished using the maximum-likelihood statistical method.

The alignment of *Tt*ArsM to its templates was based on a multiple-sequence alignment, performed with the program Clustal Omega (57); the amino acidic sequences used for the construction of the alignment of functionally characterized archaeal and bacterial arsenite methyltransferases are *Tt*ArsM from *T. thermophilus* HB27, 5 members of the *Thermus* genus (*T. islandicus*, *T. caldilimi*, *T. antranikianii*, *T. oshimai*, and *T. brockianus*), *Clostridium* sp. BMX, *R. palustris*, *M. acetivorans*, *H. salinarum*, *P. alcaligenes*, and *Cyanidioschyzon merolae* (48).

Models of *Tt*ArsM were generated through I-TASSER (58) (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) using as input the complete sequence of *Tt*ArsM (C score, −2.5). The dimeric structure was predicted using the GalaxyWEB tool (http://galaxy.seoklab.org/index.html) (59). The molecular dockings of *Tt*ArsM with arsenite and SAM were generated using the Hex protein docking server (60). One hundred rigid-body docking solutions were generated per case, and the best 10 were refined by energy minimization. The proposed model for the metal ion docked into *Tt*ArsM is the structure with the smallest distance between arsenite-histidine and cysteine-SAM (4.33 Å from H40 and 5.77 Å from H179 in *Tt*ArsM model and 4.40 Å from *C77*).

Cloning, expression, and purification of recombinant *Tt*ArsM and *Tt*ArsM mutants. The pET30b(+)/*TtarsM* vector was constructed for the expression and subsequent purification of the C-terminal His-tagged version of *Tt*ArsM. For the construction of the pET30b(+)/*TtarsM*, *TtarsM* gene was PCR amplified from *T. thermophilus* HB27 genome, using *Taq* DNA polymerase (Thermo Fisher Scientific) and primers containing the NdeI (*arsMfw*; Table S3 in the supplemental material) and HindIII (*arsMrv*; Table S3) restriction sites at their 5′ ends. The PCR product was purified, digested with the NdeI and HindIII restriction enzymes (NEB), and ligated (T4 ligase; NEB) into NdeI/HindIII-digested pET30b(+) vector (Novagen). The ligase mixture was transformed into *E. coli* TOP10F′ cells, which were plated on LB agar plates supplemented with 50 μg/ml kanamycin (Sigma-Aldrich). Single colonies were selected and inoculated in LB liquid medium supplemented with 50 μg/ml kanamycin. Plasmid isolation and sequencing were subsequently performed before transforming *E. coli* BL21-CodonPlus (DE3)-RIL cells with pET30b(+)/*TtarsM* vector.

To obtain mutation of *TtarsM* gene sequence at specific sites, the QuikChange II-E site-directed mutagenesis kit (Agilent Technologies) was employed; pET30b(+)/*TtarsM* was used as a template and amplified with three different mutagenic primer pairs (Table S3) to get pET30b(+)/*TtarsM* C77S, pET30b(+)/*TtarsM* H40A, and pET30b(+)/*TtarsM* H179A vectors. The reaction mixtures were transformed into *E. coli* TOP10F′ cells were plated on LB agar plates supplemented with kanamycin (50 μg/ml). Single colonies were randomly selected and inoculated in LB liquid medium supplemented with kanamycin (50 μg/ml). Plasmid isolation was subsequently performed, and *E. coli* BL21-CodonPlus (DE3)-RIL cells were transformed with sequence-verified pET30b(+)/*TtarsM* C77S, pET30b(+)/*TtarsM* H40A, and pET30b(+)/*TtarsM* H179A vectors.

For protein expression of the His-tagged versions of *Tt*ArsM C77S, *Tt*ArsM H40A, and *Tt*ArsM H179A catalytic mutants, the recombinant *E. coli* BL21-CodonPlus (DE3)-RIL strains were cultured in LB medium supplemented with kanamycin (50 μg/ml) and chloramphenicol (33 μg/ml). Protein expression was induced via the addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) when the cultures reached OD_{600} of 0.7. The cultures were further incubated with vigorous shaking at 37°C for 16 h, then centrifuged, resuspended in lysis buffer (20 mM NaP, pH 7.4, 50 mM NaCl, and 20 mM imidazole) supplemented with protease inhibitor cocktail (Thermo Scientific), and lysed by sonication (10 cycles of 30 s on/30 s off, 40% power; Misonix Sonicator Ultrasonic Processor XL). The lysates were centrifuged and the supernatants used for the purification on HisTrap HP columns (1 ml; GE Healthcare) connected to an AKTA Explorer system (GE Healthcare). The fractions containing His-tagged *Tt*ArsM proteins were eluted from the columns using a linear gradient of the elution buffer (20 mM NaP, pH 7.4, 50 mM NaCl, and 500 mM imidazole). The eluted protein fractions were subjected to SDS-PAGE analysis, and the fractions containing purified *Tt*ArsM were pooled and dialyzed for 16 h at 4°C in 20 mM NaP, pH 7.4, buffer supplemented with protease inhibitor cocktail (Thermo Scientific). The identity of the purified His-tagged *Tt*ArsM protein was confirmed by mass spectrometry, and protein aliquots were stored at −20°C.

***Tt*ArsM quaternary structure assessment.** The native molecular mass of *Tt*ArsM was determined by loading 500 μg of the purified protein onto an analytical Superdex PC75 column (3.2 by 30 cm) connected to an AKTA Pure system in 50-mM Tris-HCl, pH 7.5, and 0.2-M KCl buffer. The column was calibrated using a set of gel

filtration markers (low range; GE Healthcare), including ovalbumin (43.0 kDa), carbonic anhydrase (29.0 kDa), RNase A (13.7 kDa), and aprotinin (6.5 kDa) as previously described (24).

Methyltransferase activity assay. According to the manufacturer's protocol, the *Tt*ArsM arsenite methyltransferase activity was measured using the SAM510 SAM methyltransferase assay kit (G-Biosciences) with modifications regarding the temperature, SAM concentration, and reaction time (33–35). The assay relies on the degradation of *S*-adenosylhomocysteine (SAH) into urate and hydrogen peroxide by a mixture of enzymes (adenosylhomocysteine nucleosidase, adenine deaminase, and xanthine oxidase). Then, the reaction of hydrogen peroxide with 4-aminoantipyrine produces 5-dichloro-2-hydroxybenzene sulfonic acid (DHBS) with ε_{mM} of 15.0 at 510 nm. A typical reaction mixture containing 200 μM As(III), 800 μM SAM, 3.1 μM the enzyme, SAM enzyme mixture, and SAM colorimetric mix in a final reaction volume of 115 μl was incubated for 1 h at 50°C in a Synergy HTX multimode microplate reader (BioTek). The same reaction mixture was tested with 10 μg of *Tt*ArsC or 10 μg of *Tt*SmtB as negative controls. One unit of arsenite methyltransferase produces 1.0 μmol of DHBS per minute at 50°C under the conditions described above. Preliminary assays were performed to define substrate saturating concentrations, varying the As(III) and SAM concentrations from 50 μM to 300 μM and from 200 μM to 1.2 mM, respectively.

***In vitro* arsenite methylation.** As(III) methylation by *Tt*ArsM was determined in an assay solution containing 10 μM *Tt*ArsM, 250 μM As(III), 6 mM glutathione (GSH), and 1 mM SAM, in Na-phosphate 50 mM, pH 7.4, at 65°C for 24 h; the same reaction without *Tt*ArsM was used as the negative control (control sample). The reactions were terminated by the addition of 10% (vol/vol) H_2O_2. Samples were then filtered through 0.22-μm mixed cellulose ester (MCE) syringe filters and used for GC-MS analysis performed in an Agilent GC 6890, coupled with a 5973 MS detector. A modification of the Huang protocol (36) was employed for the detection of the methylated products, and the arsenic compounds were analyzed by using hyphenated mass spectrometry techniques without derivatization. Unambiguous identification of arsenic compounds was accomplished using authentic standards, including disodium methyl arsonate hexahydrate (MMAs) (ChemService) and dimethylarsinic acids (DMAs) (Merck Life Science S.r.l.).

In more detail, 200 μl of each sample was treated with 800 μl of methanol (Sigma-Aldrich). The supernatants were recovered, vacuum-dried, and solubilized in 10 μl of methanol. We analyzed 1 μl of each sample by GC-MS, using the HP5 capillary column (30 m $\times$ 0.25 mm, 0.25 μm; Agilent). Helium was employed as carrier gas at a rate of 1.0 ml min^{-1}. The temperature of the GC injector was maintained at 230°C; while the oven temperature was initially set at 40°C for 5 min, it was subsequently increased to 280°C for 5 min with an increment of 20°C/min, resulting in a total separation time of 20 min. The temperature of the analyzer was kept at 250°C. The collision energy was set to 70 eV, and the generated fragment ions were analyzed in a mass range of 20 to 450 *m/z*. The presence of arsenic compounds in the analyzed samples was confirmed by comparing the fragmentation spectra of the samples and the corresponding retention times with those of MMA and DMA standard solutions. Finally, the presence of MMAs and DMAs in the samples was confirmed by matching the experimental fragmentation spectra with those published in the NIST 05 mass spectral library. According to the NIST guidelines, the identification was reliable when the matching values were higher than 700 (37). The analyses of the samples were performed in technical triplicates.

Co-IP assay. The protein-protein interaction between *Tt*SmtB and *Tt*ArsM was *in vitro* verified via coimmunoprecipitation assays that employed recombinant *Tt*SmtB, recombinant His-tagged-*Tt*ArsM, anti-*Tt*SmtB antibodies (GeneCust), and His tag antibodies (Sigma-Aldrich). His tag removal from recombinant *Tt*SmtB was performed as previously described (22).

A typical co-IP mixture contained 1 ml of co-IP buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol, and 0.1% Triton X-100), 5 μg of *Tt*SmtB, and 5 μg of His-tagged *Tt*ArsM and was incubated at 4°C for 2 h in continuous rotation. In some cases, arsenite, arsenate, cadmium, and antimony at 1:0, 1:25, 1:50, 1:100 molar ratios, respectively, preincubated with *Tt*SmtB for 10 min at 60°C, were added. As controls, *Tt*ArsM (5 μg), *Tt*SmtB (5 μg), and *Tt*GalA (40) were also separately incubated under the same conditions. All the samples were subjected to immunoprecipitation using 2 μl of purified anti-*Tt*SmtB antibodies (2 μg/μl) (GeneCust) for 3 h at 4°C in continuous rotation before adding 15 μl of protein A-Sepharose beads (Sigma-Aldrich) and allowing the incubation to continue for a further 16 h at 4°C. The formed immunocomplexes were washed with co-IP buffer and analyzed by Western blot on 15% SDS-PAGE, using polyvinylidene difluoride (PVDF) membranes (Millipore) and anti-polyhistidine-peroxidase antibodies (Sigma-Aldrich) diluted 1:10,000, as previously described (61).

ImageJ software (https://imagej.net/) was used for densitometric analysis of the formed bands, setting as maximum value the intensity of the band of the sample containing *Tt*ArsM and *Tt*SmtB in the absence of any metal. Each experiment was performed in three technical and two biological replicates; statistical analysis was performed using one-way analysis of variance (ANOVA), and significant differences are indicated as *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$.

EMSAs. Electrophoretic mobility shift assays (EMSAs) were performed to determine the *in vitro* binding of *Tt*SmtB to the promoter region upstream of *TtarsM*. The 108-bp chromosomal region that encompasses the start codon of the *TtarsM* gene and the 105-bp-long region upstream, denoted as p_{arsM}, was PCR amplified with *Taq* DNA polymerase (Thermo Fisher Scientific) using the p_{arsM} *Fw* and p_{arsM} *Rv* primers (Table S3). EMSA reactions were set up as previously described (23, 62) in the presence of 1 μg of poly(dI-dC), 20 ng p_{arsM}, and increasing concentrations of *Tt*SmtB (1, 2, 3, 5, and 10 μM, considering *Tt*SmtB as a dimer) using SYBR Gold nucleic acid gel stain for band detection. The EMSA reactions that simultaneously employed *Tt*SmtB and *Tt*ArsM were set up in identical conditions using 3 μM of each protein alone or in combination.

Arsenic tolerance of *E. coli* expressing *Tt*ArsM. The following strains were inoculated in 10 ml of LB precultures, supplemented with kanamycin (50 μg/ml) and chloramphenicol (33 μg/ml): (i) *E. coli* BL21-CodonPlus (DE3)-RIL, pET30/*TtarsM*; (ii) *E. coli* BL21-CodonPlus (DE3)-RIL, pET30/*TtarsM C77S*; (iii) *E.*

coli BL21-CodonPlus (DE3)-RIL, pET30/*TtarsM H40A*; (iv) *E. coli* BL21-CodonPlus (DE3)-RIL, pET30/*TtarsM H179A*; and (v) *E. coli* BL21-CodonPlus (DE3)-RIL, pET30 (control). The precultures were incubated at 37°C for 16 h at 180 rpm. Subsequently, 50-ml LB cultures, supplemented with antibiotics, were inoculated with the precultures to initial OD_{600} of 0.08 and incubated at 37°C and 180 rpm until an OD_{600} of 0.6 was reached (exponential growth). At that point, protein expression was induced with 1 mM IPTG, and the cultures were incubated at 37°C and 180 rpm for 3 additional hours. From these growing cells, fresh LB cultures were inoculated to OD_{600} of 0.05 and distributed to 24-well plates (1 ml per well) containing LB medium supplemented with 1 mM IPTG, kanamycin (50 μg/ml), chloramphenicol (33 μg/ml), and increasing concentrations of arsenate and arsenite (from 2.5 mM to 7.0 mM). The MICs endpoint for each strain were determined as the lowest concentration of arsenite at which there was the difference between grown and start culture lower than 0.01 OD_{600} after 16 h of incubation at 37°C (22). All the cells up to the MIC value were able to grow if reinoculated in an arsenic-free medium. The reported values are the average of three biological replicates.

ThermoCas9 editing and targeting constructs. The plasmids used for the ThermoCas9-based targeting and editing experiments are listed in Table S4. The vector pMK18 was used as the template for the construction of the ThermoCas9-based targeting and editing plasmids and the employed primers, the DNA templates, and the DNA fragments, which are listed in Table S4. The *thermoCas9* gene was codon harmonized according to *T. thermophilus* HB27 codon usage using the Galaxy/Codon Harmonizer online tool (63), and it was synthesized (Twist Bioscience) (Table S5). The DNA fragments were designed with appropriate overhangs for NEBuilder HiFi DNA assembly (NEB), and they were obtained through PCR with Q5 polymerase (NEB). The PCR products were subjected to 1% agarose gel electrophoresis, and they were purified using a Zymogen gel DNA recovery kit (Zymo Research). The assembly reactions were transformed to chemically competent *E. coli* DH5α cells (NEB), and the cells were plated on LB agar plates supplemented with kanamycin (50 μg/ml). Single colonies were inoculated in LB medium supplemented with kanamycin (50 μg/ml) for overnight incubation at 37°C. Plasmid material was isolated using the GeneJet plasmid miniprep kit (Thermo Fisher Scientific) and sequence verified (GATC Biotech), and 300 ng of each plasmid (pMK-ThermoCas9-NT/sp1/sp2 and pMK-ThermoCas9-HR-NT/sp1/sp2) was transformed to either *T. thermophilus* HB27 cells (22), as indicated per experimental process.

The obtained plasmid, pMK-ThermoCas9-NT, was used as the backbone to construct a new plasmid to obtain the deletion of the *TtarsX* gene (Table S4) and the insertion of the gene coding sYFP (27). The obtained plasmid (pMK-ThermoCas9-HR-syfp) was used to transform *T. thermophilus* HB27 Δ*TtarsM* cells as already described to obtain the strain denoted Δ*TtarsM*-Δ*TtarsX* (*syfp*).

Arsenic tolerance of *T. thermophilus* HB27 wild-type and mutant strains. Exponentially growing precultures of *T. thermophilus* HB27 (control), *T. thermophilus* HB27 Δ*TtarsM*, and *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) were diluted to an OD_{600} of 0.08 in 10-ml TM cultures containing increasing concentrations of arsenite and arsenate (from 0.1 mM to 50 mM). The cultures were incubated aerobically at 70°C for 18 h, and the MIC values were determined as the lowest concentrations of arsenite and arsenate that completely inhibited the growth of a strain (22). The reported values are the average of three biological replicates.

Bioreporter activity measurement. Overnight cultures of *T. thermophilus* HB27 Δ*TtarsM*-Δ*TtarsX* (*syfp*) strain were diluted to an OD_{600} of 0.08 in TM medium and then grown aerobically at 70°C until an OD_{600} of 0.5 was reached. The cultures were divided into samples of 5 ml each and subsequently supplemented with increasing concentrations of arsenite and arsenate (0.5 μM to 4 μM). After 1 h of incubation at 70°C, 200 μl of each cell sample were removed and centrifuged for 5 min at 6,000 rpm. The pellets were washed twice with equal volumes of phosphate-buffered saline (PBS) 1× and resuspended with equal volumes of PBS 1× before being distributed into a 96-well plate. sYFP fluorescence intensity of each sample was measured employing a Synergy HTX multimode microplate reader (BioTek), using excitation and emission wavelengths of 458 nm and 540 nm, respectively. The measured fluorescence intensities were normalized for the optical density of each sample at 600 nm. The measured fluorescence was reported as fluorescence relative expression, assuming that the fluorescence value of not-treated cells (control) was 1.

Each experiment was performed in three technical and biological replicates. Statistical analysis was performed using one-way ANOVA; significant differences are indicated as *, $P < 0.05$; **, $P < 0.01$; and ***, $P < 0.001$.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 2.3 MB.
FIG S2, TIF file, 0.4 MB.
FIG S3, TIF file, 0.6 MB.
FIG S4, TIF file, 0.5 MB.
FIG S5, TIF file, 0.3 MB.
TABLE S1, DOCX file, 0.02 MB.
TABLE S2, DOCX file, 0.02 MB.
TABLE S3, DOCX file, 0.02 MB.
TABLE S4, DOCX file, 0.02 MB.
TABLE S5, DOCX file, 0.01 MB.

ACKNOWLEDGMENTS

We kindly acknowledge Josè Berenguer from Universidad Autònoma de Madrid for kindly providing the pMK-Pnqo-syfp vector. Lucia Cavalca and Gian Attilio Sacchi from the University of Milan are also acknowledged for donating disodium methyl arsonate hexahydrate (MMAs) (ChemService) and dimethylarsinic acid (DMAs) (Merck Life Science S.r.l.). We also acknowledge Angela Amoresano (Department of Chemical Sciences, University of Naples Federico II) for helpful discussion.

Acknowledgements go to grant ERA-NET Cofund MarTERA, who funded the project FLAshMoB (Functional Amyloid Chimera for Marine Biosensing). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation program, grant agreement no. 834279, and from the Netherlands Organization for Scientific Research (NWO; Spinoza grant SPI 93-537) to J.v.d.O.

REFERENCES

1. Mukhopadhyay R, Rosen BP, Phung LT, Silver S. 2002. Microbial arsenic: from geocycles to genes and enzymes. FEMS Microbiol Rev 26:311–325. https://doi.org/10.1111/j.1574-6976.2002.tb00617.x.
2. Chen Z, Wang Y, Jiang X, Fu D, Xia D, Wang H, Dong G, Li Q. 2017. Dual roles of AQDS as electron shuttles for microbes and dissolved organic matter involved in arsenic and iron mobilization in the arsenic-rich sediment. Sci Total Environ 574:1684–1694. https://doi.org/10.1016/j.scitotenv.2016.09.006.
3. Yang H-C, Rosen BP. 2016. New mechanisms of bacterial arsenic resistance. Biomed J 39:5–13. https://doi.org/10.1016/j.bj.2015.08.003.
4. Gallo G, Puopolo R, Carbonaro M, Maresca E, Fiorentino G. 2021. Extremophiles, a nifty tool to face environmental pollution: from exploitation of metabolism to genome engineering. Int J Environ Res Public Health 18:5228. https://doi.org/10.3390/ijerph18105228.
5. Politi J, Spadavecchia J, Fiorentino G, Antonucci I, Casale S, De Stefano L. 2015. Interaction of Thermus thermophilus ArsC enzyme and gold nanoparticles naked-eye assays speciation between As(III) and As(V). Nanotechnology 26:435703. https://doi.org/10.1088/0957-4484/26/43/435703.
6. Politi J, Spadavecchia J, Fiorentino G, Antonucci I, De Stefano L. 2016. Arsenate reductase from *Thermus thermophilus* conjugated to polyethylene glycol-stabilized gold nanospheres allow trace sensing and speciation of arsenic ions. J R Soc Interface 13:20160629. https://doi.org/10.1098/rsif.2016.0629.
7. Rosen BP. 2002. Biochemistry of arsenic detoxification. FEBS Lett 529:86–92. https://doi.org/10.1016/s0014-5793(02)03186-1.
8. Liu Z, Rensing C, Rosen BP. 2011. Toxicity: resistance pathways for metalloids and toxic metals, p 1–13. *In* Scott RA (ed), Encyclopedia of inorganic and bioinorganic chemistry. John Wiley & Sons, Ltd., Hoboken, NJ.
9. Aulitto M, Gallo G, Puopolo R, Mormone A, Limauro D, Contursi P, Piochi M, Bartolucci S, Fiorentino G. 2021. Genomic insight of Alicyclobacillus mali FL18 isolated from an arsenic-rich hot spring. Front Microbiol 12:639697. https://doi.org/10.3389/fmicb.2021.639697.
10. Mukhopadhyay R, Rosen BP. 2002. Arsenate reductases in prokaryotes and eukaryotes. Environ Health Perspect 110:745–748. https://doi.org/10.1289/ehp.02110s5745.
11. Chen C-M, Misra TK, Silver S, Rosen BP. 1986. Nucleotide sequence of the structural genes for an anion pump. The plasmid-encoded arsenical resistance operon. J Biol Chem 261:15030–15038. https://doi.org/10.1016/S0021-9258(18)66824-3.
12. Wu J, Rosen BP. 1991. The ArsR protein is a trans-acting regulatory protein. Mol Microbiol 5:1331–1336. https://doi.org/10.1111/j.1365-2958.1991.tb00779.x.
13. Qin J, Rosen BP, Zhang Y, Wang G, Franke S, Rensing C. 2006. Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. Proc Natl Acad Sci U S A 103:2075–2080. https://doi.org/10.1073/pnas.0506836103.
14. Chen J, Madegowda M, Bhattacharjee H, Rosen BP. 2015. ArsP: a methylarsenite efflux permease. Mol Microbiol 98:625–635. https://doi.org/10.1111/mmi.13145.
15. Chen S-C, Sun G-X, Yan Y, Konstantinidis KT, Zhang S-Y, Deng Y, Li X-M, Cui H-L, Musat F, Popp D, Rosen BP, Zhu Y-G. 2020. The great oxidation event expanded the genetic repertoire of arsenic metabolism and cycling. Proc Natl Acad Sci U S A 117:10414–10421. https://doi.org/10.1073/pnas.2001063117.
16. Chen S-C, Sun G-X, Rosen BP, Zhang S-Y, Deng Y, Zhu B-K, Rensing C, Zhu Y-G. 2017. Recurrent horizontal transfer of arsenite methyltransferase genes facilitated adaptation of life to arsenic. Sci Rep 7:1–11. https://doi.org/10.1038/s41598-017-08313-2.
17. Chen J, Rosen BP. 2020. The arsenic methylation cycle: how microbial communities adapted methylarsenicals for use as weapons in the continuing war for dominance. Front Environ Sci 8. https://doi.org/10.3389/fenvs.2020.00043.
18. Gihring TM, Druschel GK, McCleskey RB, Hamers RJ, Banfield JF. 2001. Rapid arsenite oxidation by Thermus aquaticus and Thermus thermophilus: field and laboratory investigations. Environ Sci Technol 35:3857–3862. https://doi.org/10.1021/es010816f.
19. Saltikov CW, Olson BH. 2002. Homology of Escherichia coli R773 arsA, arsB, and arsC genes in arsenic-resistant bacteria isolated from raw sewage and arsenic-enriched creek waters. Appl Environ Microbiol 68:280–288. https://doi.org/10.1128/AEM.68.1.280-288.2002.
20. Qin J, Lehr CR, Yuan C, Le XC, McDermott TR, Rosen BP. 2009. Biotransformation of arsenic by a Yellowstone thermoacidophilic eukaryotic alga. Proc Natl Acad Sci U S A 106:5213–5217. https://doi.org/10.1073/pnas.0900238106.
21. Henne A, Brüggemann H, Raasch C, Wiezer A, Hartsch T, Liesegang H, Johann A, Lienard T, Gohl O, Martinez-Arias R, Jacobi C, Starkuviene V, Schlenczeck S, Dencker S, Huber R, Klenk HP, Kramer W, Merkl R, Gottschalk G, Fritz HJ. 2004. The genome sequence of the extreme thermophile Thermus thermophilus. Nat Biotechnol 22:547–553. https://doi.org/10.1038/nbt956.
22. Antonucci I, Gallo G, Limauro D, Contursi P, Ribeiro AL, Blesa A, Berenguer J, Bartolucci S, Fiorentino G. 2017. An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in Thermus thermophilus HB27. Microb Biotechnol 10:1690–1701. https://doi.org/10.1111/1751-7915.12761.
23. Antonucci I, Gallo G, Limauro D, Contursi P, Ribeiro AL, Blesa A, Berenguer J, Bartolucci S, Fiorentino G. 2018. Characterization of a promiscuous cadmium and arsenic resistance mechanism in Thermus thermophilus HB27 and potential application of a novel bioreporter system. Microb Cell Fact 17:78. https://doi.org/10.1186/s12934-018-0918-7.
24. Del Giudice I, Limauro D, Pedone E, Bartolucci S, Fiorentino G. 2013. A novel arsenate reductase from the bacterium Thermus thermophilus HB27: its role in arsenic detoxification. Biochim Biophys Acta 1834:2071–2079. https://doi.org/10.1016/j.bbapap.2013.06.007.
25. Puopolo R, Sorrentino I, Gallo G, Piscitelli A, Giardina P, Le Goff A, Fiorentino G. 2021. Self-assembling thermostable chimeras as new platform for arsenic biosensing. Sci Rep 11:2991. https://doi.org/10.1038/s41598-021-82648-9.
26. Busenlehner LS, Pennella MA, Giedroc DP. 2003. The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance. FEMS Microbiol Rev 27:131–143. https://doi.org/10.1016/S0168-6445(03)00054-8.
27. Blesa A, Berenguer J. 2015. Contribution of vesicle-protected extracellular DNA to horizontal gene transfer in Thermus spp. Int Microbiol 18:177–187. https://doi.org/10.2436/20.1501.01.248.

28. Carpentieri A, Gamberi T, Modesti A, Amoresano A, Colombini B, Nocella M, Bagni MA, Fiaschi T, Barolo L, Gulisano M, Magherini F. 2016. Profiling carbonylated proteins in heart and skeletal muscle mitochondria from trained and untrained mice. J Proteome Res 15:3666–3678. https://doi.org/10.1021/acs.jproteome.6b00475.
29. Rao ST, Rossmann MG. 1973. Comparison of super-secondary structures in proteins. J Mol Biol 76:241–256. https://doi.org/10.1016/0022-2836(73)90388-4.
30. Schubert HL, Blumenthal RM, Cheng X. 2003. Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci 28:329–335. https://doi.org/10.1016/S0968-0004(03)00090-2.
31. Huang K, Xu Y, Packianathan C, Gao F, Chen C, Zhang J, Shen Q, Rosen BP, Zhao F. 2018. Arsenic methylation by a novel ArsM As (III) S-adenosylmethionine methyltransferase that requires only two conserved cysteine residues. Mol Microbiol 107:265–276. https://doi.org/10.1111/mmi.13882.
32. Gallo G, Antonucci I, Pirone L, Amoresano A, Contursi P, Limauro D, Pedone E, Bartolucci S, Fiorentino G. 2019. A physicochemical investigation on the metal binding properties of TtSmtB, a thermophilic member of the ArsR/SmtB transcription factor family. Int J Biol Macromol 138:1056–1063. https://doi.org/10.1016/j.ijbiomac.2019.07.174.
33. Singh J, Vijay S, Mansuri R, Rawal R, Kadian K, Sahoo GC, Kumar M, Sharma A. 2019. Computational and experimental elucidation of Plasmodium falciparum phosphoethanolamine methyltransferase inhibitors: pivotal drug target. PLoS One 14:e0221032. https://doi.org/10.1371/journal.pone.0221032.
34. Patil NA, Basu B, Deobagkar DD, Apte SK, Deobagkar DN. 2017. Putative DNA modification methylase DR_C0020 of Deinococcus radiodurans is an atypical SAM dependent C-5 cytosine DNA methylase. Biochim Biophys Acta Gen Subj 1861:593–602. https://doi.org/10.1016/j.bbagen.2016.12.025.
35. Dou L, Yan F, Pang J, Zheng D, Li D, Gao L, Wang L, Xu Y, Shi J, Wang Q, Zhou L, Shen N, Singh P, Wang L, Li Y, Gao Y, Liu T, Chen C, Al-Kali A, Litzow MR, Chi Y-I, Bode AM, Liu C, Huang H, Liu D, Marcucci G, Liu S, Yu L. 2019. Protein lysine 43 methylation by EZH1 promotes AML1-ETO transcriptional repression in leukemia. Nat Commun 10:1–15. https://doi.org/10.1038/s41467-019-12960-6.
36. Huang M, Ting Wang Y, Ho PC. 2008. Quantification of arsenic compounds using derivatization, solvent extraction and liquid chromatography electrospray ionization tandem mass spectrometry. J Pharm Biomed Anal 48:1381–1391. https://doi.org/10.1016/j.jpba.2008.09.018.
37. Stein SE. 1994. Estimating probabilities of correct identification from results of mass spectral library searches. J Am Soc Mass Spectrom 5:316–323. https://doi.org/10.1016/1044-0305(94)85022-4.
38. Zhang J, Cao T, Tang Z, Shen Q, Rosen BP, Zhao F-J. 2015. Arsenic methylation and volatilization by arsenite S-adenosylmethionine methyltransferase in Pseudomonas alcaligenes NBRC14159. Appl Environ Microbiol 81:2852–2860. https://doi.org/10.1128/AEM.03804-14.
39. Wang P-P, Sun G-X, Zhu Y-G. 2014. Identification and characterization of arsenite methyltransferase from an archaeon, Methanosarcina acetivorans C2A. Environ Sci Technol 48:12706–12713. https://doi.org/10.1021/es503869k.
40. Aulitto M, Fusco S, Fiorentino G, Limauro D, Pedone E, Bartolucci S, Contursi P. 2017. Thermus thermophilus as source of thermozymes for biotechnological applications: homologous expression and biochemical characterization of an α-galactosidase. Microb Cell Fact 16:1–10. https://doi.org/10.1186/s12934-017-0638-4.
41. Osman D, Cavet JS. 2010. Bacterial metal-sensing proteins exemplified by ArsR–SmtB family repressors. Nat Prod Rep 27:668–680. https://doi.org/10.1039/b906682a.
42. Tanous C, Soutourina O, Raynal B, Hullo M-F, Mervelet P, Gilles A-M, Noirot P, Danchin A, England P, Martin-Verstraete I. 2008. The CymR regulator in complex with the enzyme CysK controls cysteine metabolism in Bacillus subtilis. J Biol Chem 283:35551–35560. https://doi.org/10.1074/jbc.M805951200.
43. Fisher SH, Wray LV. 2008. Bacillus subtilis glutamine synthetase regulates its own synthesis by acting as a chaperone to stabilize GlnR-DNA complexes. Proc Natl Acad Sci U S A 105:1014–1019. https://doi.org/10.1073/pnas.0709949105.
44. Cava F, Laptenko O, Borukhov S, Chahlafi Z, Blas-Galindo E, Gómez-Puertas P, Berenguer J. 2007. Control of the respiratory metabolism of Thermus thermophilus by the nitrate respiration conjugative element NCE. Mol Microbiol 64:630–646. https://doi.org/10.1111/j.1365-2958.2007.05687.x.
45. Mougiakos I, Mohanraju P, Bosma EF, Vrouwe V, Bou MF, Naduthodi MIS, Gussak A, Brinkman RBL, Van Kranenburg R, Van Der Oost J. 2017. Characterizing a thermostable Cas9 for bacterial genome editing and silencing. Nat Commun 8:1–11. https://doi.org/10.1038/s41467-017-01591-4.
46. Blesa A, Baquedano I, Quintáns NG, Mata CP, Castón JR, Berenguer J. 2017. The transjugation machinery of Thermus thermophilus: identification of TdtA, an ATPase involved in DNA donation. PLoS Genet 13:e1006669. https://doi.org/10.1371/journal.pgen.1006669.
47. Roy R, Samanta S, Patra S, Mahato NK, Saha RP. 2018. In silico identification and characterization of sensory motifs in the transcriptional regulators of the ArsR-SmtB family. Metallomics 10:1476–1500. https://doi.org/10.1039/c8mt00082d.
48. Ajees AA, Marapakala K, Packianathan C, Sankaran B, Rosen BP. 2012. Structure of an As (III) S-adenosylmethionine methyltransferase: insights into the mechanism of arsenic biotransformation. Biochemistry 51:5476–5485. https://doi.org/10.1021/bi3004632.
49. Dheeman DS, Packianathan C, Pillai JK, Rosen BP. 2014. Pathway of human AS3MT arsenic methylation. Chem Res Toxicol 27:1979–1989. https://doi.org/10.1021/tx500313k.
50. Fusco S, Aulitto M, Iacobucci I, Crocamo G, Pucci P, Bartolucci S, Monti M, Contursi P. 2020. The interaction between the F55 virus-encoded transcription regulator and the RadA host recombinase reveals a common strategy in Archaea and Bacteria to sense the UV-induced damage to the host DNA. Biochim Biophys Acta Gene Regul Mech 1863:194493. https://doi.org/10.1016/j.bbagrm.2020.194493.
51. Le Y, Fu Y, Sun J. 2020. Genome editing of the anaerobic thermophile Thermoanaerobacter ethanolicus using thermostable Cas9. Appl Environ Microbiol 87:e01773-20. https://doi.org/10.1128/AEM.01773-20.
52. Adalsteinsson BT, Kristjansdottir T, Merre W, Helleux A, Dusaucy J, Tourigny M, Fridjonsson O, Hreggvidsson GO. 2021. Efficient genome editing of an extreme thermophile, Thermus thermophilus, using a thermostable Cas9 variant. Sci Rep 11:1–15. https://doi.org/10.1038/s41598-021-89029-2.
53. Monti M, Orru S, Pagnozzi D, Pucci P. 2005. Functional proteomics. Clin Chim Acta 357:140–150. https://doi.org/10.1016/j.cccn.2005.03.019.
54. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol 35:1547–1549. https://doi.org/10.1093/molbev/msy096.
55. Wang P-P, Bao P, Sun G-X. 2015. Identification and catalytic residues of the arsenite methyltransferase from a sulfate-reducing bacterium, Clostridium sp. BXM. FEMS Microbiol Lett 362:1–8. https://doi.org/10.1093/femsle/fnu003.
56. Wang G, Kennedy SP, Fasiludeen S, Rensing C, DasSarma S. 2004. Arsenic resistance in Halobacterium sp. strain NRC-1 examined by using an improved gene knockout system. J Bacteriol 186:3187–3194. https://doi.org/10.1128/JB.186.10.3187-3194.2004.
57. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539. https://doi.org/10.1038/msb.2011.75.
58. Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9:40. https://doi.org/10.1186/1471-2105-9-40.
59. HeeShin W. 2014. Prediction of protein structure and interaction by GALAXY protein modeling programs. Biodesign 2:1–11.
60. Macindoe G, Mavridis L, Venkatraman V, Devignes M-D, Ritchie DW. 2010. HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res 38:W445–W449. https://doi.org/10.1093/nar/gkq311.
61. Fiorentino G, Ronca R, Cannio R, Rossi M, Bartolucci S. 2007. MarR-like transcriptional regulator involved in detoxification of aromatic compounds in Sulfolobus solfataricus. J Bacteriol 189:7351–7360. https://doi.org/10.1128/JB.00885-07.
62. Fiorentino G, Del Giudice I, Bartolucci S, Durante L, Martino L, Del Vecchio P. 2011. Identification and physicochemical characterization of BldR2 from Sulfolobus solfataricus, a novel archaeal member of the MarR transcription factor family. Biochemistry 50:6607–6621. https://doi.org/10.1021/bi200187j.
63. Claassens NJ, Siliakus MF, Spaans SK, Creutzburg SCA, Nijsse B, Schaap PJ, Quax TEF, Van Der Oost J. 2017. Improving heterologous membrane protein production in Escherichia coli by combining transcriptional tuning and codon usage algorithms. PLoS One 12:e0184355. https://doi.org/10.1371/journal.pone.0184355.

Chapter 4 – Extremophiles as source of enzymatic systems for lignocellulose valorization

This chapter reports the genomic analysis of two bacteria recently isolated from Pisciarelli, a volcanic hot spring located in the Area of Campi Flegrei (Pozzuoli, Italy), already characterised in terms of metals and antibiotics resistance. These analyses revealed the presence in both microbes of genes potentially encoding enzymes critical for the degradation of various components of lignocellulosic biomass. Therefore, this chapter is divided into two sections:

- Paper 2-I: this study investigates the potential of *Alicyclobacillus mali* FL18 in biomass valorization. Genomic analysis revealed a significant presence of genes coding for carbohydrate-active enzyme, with GHs playing an important role. Exploring the GH9 family, the focus was on AmCel9, a protein that was purified from recombinant *E. coli* cells and showed impressive functional properties. AmCel9 exhibited broad pH tolerance, remarkable stability, and high activity under various conditions. Its efficient hydrolysis of cellulose-related substrates, including insoluble ones, positioned AmCel9 as a promising biocatalyst for industrial applications.

- Manuscript 2-II (*in preparation*): this work focuses on the genomic annotation and evaluation of the potential of *Geobacillus stearothermophilus* GF16 in biomass valorization. The genome of this novel strain was analysed, highlighting 69 potential carbohydrate-active enzymes, in particular glycosyl transferases and glycoside hydrolases involved in xylan degradation. Indeed, the secretome of *G. stearothermophilus* GF16 grown on xylan as carbon source, was shown to be highly efficient in hydrolyzing synthetic and natural substrates, and various carbohydrate-rich biomasses, competing with the commercial enzyme Viscozyme® L. This ability to saccharify agri-food residues was also evaluated by HPAEC-PAD analyses, positioning *G. stearothermophilus* GF16 as a promising candidate for sustainable industrial applications in biomass hydrolysis with simultaneous production of xylose as valuable major product.

International Journal of
Molecular Sciences

MDPI

Article

Insight into CAZymes of *Alicyclobacillus mali* FL18: Characterization of a New Multifunctional GH9 Enzyme

Miriam Carbonaro [1], Martina Aulitto [1,2], Giovanni Gallo [1], Patrizia Contursi [1], Danila Limauro [1] and Gabriella Fiorentino [1,*]

[1] Department of Biology, University of Naples Federico II, 80126 Naples, Italy
[2] Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
* Correspondence: fiogabri@unina.it

Abstract: In the bio-based era, cellulolytic and hemicellulolytic enzymes are biocatalysts used in many industrial processes, playing a key role in the conversion of recalcitrant lignocellulosic waste biomasses. In this context, many thermophilic microorganisms are considered as convenient sources of carbohydrate-active enzymes (CAZymes). In this work, a functional genomic annotation of *Alicyclobacillus mali* FL18, a recently discovered thermo-acidophilic microorganism, showed a wide reservoir of putative CAZymes. Among them, a novel enzyme belonging to the family 9 of glycosyl hydrolases (GHs), named AmCel9, was identified; in-depth in silico analyses highlighted that AmCel9 shares general features with other GH9 members. The synthetic gene was expressed in *Escherichia coli* and the recombinant protein was purified and characterized. The monomeric enzyme has an optimal catalytic activity at pH 6.0 and has comparable activity at temperatures ranging from 40 °C to 70 °C. It also has a broad substrate specificity, a typical behavior of multifunctional cellulases; the best activity is displayed on β-1,4 linked glucans. Very interestingly, AmCel9 also hydrolyses filter paper and microcrystalline cellulose. This work gives new insights into the properties of a new thermophilic multifunctional GH9 enzyme, that looks a promising biocatalyst for the deconstruction of lignocellulose.

Keywords: lignocellulose waste; renewable sources; cellulose; *Alicyclobacillus*; thermozymes; CAZymes; multifunctional glycosyl hydrolases

Citation: Carbonaro, M.; Aulitto, M.; Gallo, G.; Contursi, P.; Limauro, D.; Fiorentino, G. Insight into CAZymes of *Alicyclobacillus mali* FL18: Characterization of a New Multifunctional GH9 Enzyme. *Int. J. Mol. Sci.* **2023**, *24*, 243. https://doi.org/10.3390/ijms24010243

Academic Editor: Vitaliy Borisov

Received: 18 November 2022
Revised: 15 December 2022
Accepted: 21 December 2022
Published: 23 December 2022

1. Introduction

The realization of an eco-sustainable economy is one of the most urgent current issues of the last few decades, which directs global efforts to reduce CO_2 emissions and use renewable sources to produce biofuels and bio-products to supplant fossil fuels and petroleum-based products. In this context, lignocellulosic waste is a promising renewable resource owing to its abundance and affordability, and it can be used as a raw material to recover high-value-added products from degradation of its complex structure [1].

The composition of lignocellulose waste is extremely heterogeneous and depends on the waste source (i.e., forestry, agricultural, municipal, industrial) as well as conditions of the plant species such as age, stage of growth and others. The main components of plant biomass are cellulose, hemicellulose and lignin, but mineral residues, proteins, pectins and nitrogenous compounds are also present [2]. Native cellulose is the most abundant homopolymer present in the secondary cell wall of plants; the fundamental repeated unit of the cellulose linear chain is cellobiose, a disaccharide of two D-glucose units 180° rotated to each other and linked through β-1,4-glycosidic bonds. Cellulose chains interact via hydrogen bonds and van der Waals forces and are arranged in microfibrillar superstructures where highly ordered regions (i.e., crystalline) alternate to disordered ones (i.e., amorphous) [3]. As a result of such a structure, cellulose has areas characterized by

high fiber strength and low accessibility, making its total decomposition another crucial step in the deconstruction of recalcitrant lignocellulose. The use of biocatalysts for cellulose degradation in biorefineries or other industrial applications has many advantages over chemical hydrolysis, but the costs related to enzyme production constitute one of the main drawbacks; for this reason, the identification of highly active and multifunctional catalysts, that can work skillfully in harsh conditions, is a goal to reduce costs [4,5].

The complete degradation of cellulose requires the synergistic action of several enzymes that display activities on different regions of cellulose fibers. They are all named cellulases and are grouped in endoglucanases, exoglucanases and β-glucosidases. Endoglucanases act on internal β-1,4-glycosidic bonds; exoglucanases or cellobiohydrolases hydrolyze the reducing or non-reducing ends of cellulose chains releasing cellobiose which is further cleaved into two glucose units by β-glucosidases. Moreover, lytic polysaccharide monooxygenases catalyze the oxidative cleavage of a glycosidic bond in the presence of hydrogen peroxide or dioxygen [6].

In particular, endoglucanases are glycosyl hydrolases (GHs) with representatives in archaea, bacteria, fungi and eukaryotic organisms and belong to 13 different families of GHs, although the majority are included in families 5, 6, 7, 9, 12 and 45 [7,8].

In the last few decades, thermophilic bacteria have received noticeable attention as reservoirs of thermotolerant, acid and alkali-stable GH enzymes suitable for application in harsh industrial conditions [9–13]. For example, the *Alicyclobacillus* genus embraces strictly aerobic, endospore-forming, thermo-acidophilic gram-positive bacilli, usually thriving in the soil, spoiled fruit-based beverages and hot springs; they possess ω-alicyclic fatty acids as the major component of membrane lipids [14]. Recent studies reported that species such as *A. acidocaldarius*, *Alicyclobacillus* sp. A4, *A. vulcanalis*, *A. cellulosilyticus*, etc., are intriguing sources of thermoacidophilic cellulases [15,16]. In this regard, a new strain of *A. mali*, FL18, was recently isolated from a hot mud pool at Pisciarelli, a solfataric field located in the volcanic area of Campi Flegrei near Naples in Italy [17].

In this work, we analyzed the lignocellulose degrading potential of *A. mali* FL18 by performing genomic functional annotation to identify the putative catalysts involved in biomass deconstruction. Among the identified coding DNA sequences (CDSs) that were found to be distributed across different CAZymes, we chose to perform more in-depth studies on a GH9 member, AmCel9. The gene was synthetically produced and overexpressed in *Escherichia coli*. The recombinant enzyme was purified and biochemically characterized, revealing activity on a broad range of substrates, pH values and temperatures as well as long-lasting stability at different conditions.

2. Results and Discussion

2.1. Analysis of CAZymes in A. mali FL18

Extremozymes represent the cornerstone for the development of green, efficient and sustainable technologies [18–23]. These biocatalysts, naturally produced by extremophiles and able to work under harsh conditions, are ideal for various industrial processes [2,24–28]. One of the notable applications of these enzymes is their use in the deconstruction of recalcitrant lignocellulosic biomasses [29,30]. Due to the intrinsic features of these enzymes, their industrial market has constantly increased over the years [31]; in the present work, we gain an insight into the CAZymes repertoire of the recently isolated *A. mali* FL18. The prediction obtained through dbCAN 2 metaserver unveiled the presence of 73 CDSs belonging to CAZymes (Table S1) and they correspond to about 2.5% of protein coding genes that are 2832 in the NCBI annotation [32]; this percentage is within the range of the CAZymes encoding genes estimated in the majority of microbial genomes, considering that the CAZomes of microorganisms typically correspond to 1–5% of the predicted CDS [33]. The analysis also showed that the most abundant family predicted in this genome is GH with 32 enzymes, followed by 28 glycosyltransferases (GTs), 11 carbohydrate esterases (CEs) and 2 auxiliary activities (AAs). All the putative genes encoding CAZymes are reported in Table 1 and sketched in Figure 1.

Table 1. *A. mali* FL18 predicted proteins containing CAZymes modules acting on lignocellulose components identified through dbCAN. Protein accession numbers, potential EC number(s), NCBI annotation and the CAZy family are shown.

Accession Number	E.C. Number	NCBI Annotation	CAZy Family
	Cellulose		
MBF8378818.1	3.2.1.21 | 3.2.1.- | 3.2.1.38 | 3.2.1.23 | 3.2.1.74	β-glucosidase	GH1
MBF8377998.1	3.2.1.4 | 3.2.1.6 | 3.2.1.151	glycoside hydrolase family 9 protein	GH9
MBF8377873.1	3.2.1.4 | 3.2.1.8	α-L-arabinofuranosidase	GH51
MBF8378553.1	2.4.1.333 | 2.4.1.- | -	carbohydrate-binding protein	GH94 + GT84
	Hemicellulose		
MBF8378422.1	3.2.1.37 | 3.2.1.55 | 3.2.1.21	glycoside hydrolase family 3 C-terminal domain-containing protein	GH3
MBF8376310.1	3.2.1.22	α-glucosidase/α-galactosidase	GH4
MBF8379065.1	3.2.1.8	endo-1,4-β-xylanase	GH10
MBF8377308.1	3.2.1.20 | 3.2.1.10	α-glucosidase	GH31
MBF8377285.1	3.2.1.177	α-xylosidase	GH31
MBF8378558.1	3.2.1.22	α-galactosidase	GH36
MBF8377281.1	3.2.1.23	β-galactosidase	GH42
MBF8377873.1	3.2.1.4 | 3.2.1.8	α-L-arabinofuranosidase	GH51
MBF8377278.1	3.2.1.55	α-*N*-arabinofuranosidase	GH51
MBF8377854.1	3.2.1.139 | 3.2.1.131	α-glucuronidase	GH67
MBF8378555.1	3.2.1.78 | 3.2.1.25	1,4-β-xylanase	GH113
MBF8376699.1	3.2.1.185	glycoside hydrolase family 127 protein	GH127
MBF8377082.1	2.4.1.339	glycoside hydrolase family 130 protein	GH130
MBF8376362.1	-	polysaccharide deacetylase family sporulation protein PdaB	CE4
MBF8376532.1	-	polysaccharide deacetylase family protein	CE4
MBF8377192.1	-	polysaccharide deacetylase family protein	CE4
MBF8377262.1	-	polysaccharide deacetylase family protein	CE4
MBF8377649.1	-	polysaccharide deacetylase family protein	CE4
MBF8378334.1	-	polysaccharide deacetylase family protein	CE4
MBF8377707.1	-	N-acetylglucosamine-6-phosphate deacetylase	CE9
MBF8376962.1	-	PIG-L family deacetylase	CE14
MBF8377409.1	3.5.1.-	PIG-L family deacetylase	CE14
MBF8377724.1	-	bacillithiol biosynthesis deacetylase BshB2	CE14
	Pectin		
MBF8377980.1	-	glycoside hydrolase family 28 protein	GH28
MBF8377281.1	3.2.1.23	β-galactosidase	GH42
MBF8377278.1	3.2.1.55	α-*N*-arabinofuranosidase	GH51
MBF8376699.1	3.2.1.185	glycoside hydrolase family 127 protein	GH127
	Lignin		
MBF8377978.1	1.10.3.2	multicopper oxidase domain-containing protein	AA1
MBF8377540.1	-	NAD(P)H:quinone oxidoreductase type IV	AA6

Mining of Enzymes for Cellulose Degradation

The in-depth functional analysis of *A. mali* FL18 genome was performed to identify new cellulose-degrading enzymes for biotechnological purposes (Table 1 and Figure 1). Four members of GHs (GH1, GH9, GH51, GH94) were found; among these, GH1 and GH94 are expected to hydrolyze short oligosaccharides, while GH9 and GH51 are putative endoglucanases. In detail, GH1 (MBF8378818.1) encodes a β-glucosidase [34] and GH94 a putative β-D-glycoside phosphorylases (MBF8378553.1), potentially acting on β-linked substrates such as cellobiose, cellodextrin, laminaribiose, N,N′-diacetylchitobiose and cellobionic acid. This latter enzyme could provide an alternative to cellobiose degradation through the intracellular phosphorolysis [35]. *A. mali* FL-18 genome also encodes two

putative arabinofuranosidases (MBF8377873.1 and MBF8377278.1) belonging to the family GH51. In particular, the gene MBF8377873.1, recognized by the dbCAN tool as either putative E.C. 3.2.1.4 (cellulase) or 3.2.1.8 (xylanase), could have a dual role in both cellulose and xylan deconstruction. Indeed, the close relative *A. acidocaldarius* possesses two endoglucanases, CelA (GH9) and CelB (GH51) [36,37], the latter found to be active on both CMC and xylan. Moreover, a recent study has shown that a rumen bacterial strain, *Fibrobacter succinogenes* S85, actively expresses a GH51 endoglucanase 2 (Eg2, CelF, and EgF) and a GH9 endoglucanase 1 (Cel2 and End1) in the presence of cellulosic material [38].

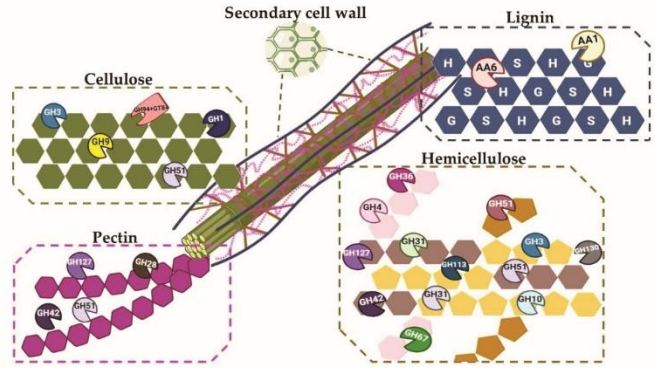

Figure 1. Overview of *A. mali* FL18 GHs on different lignocellulose components. All the GHs revealed by functional annotation are indicated as acting on their preferential lignocellulose substrate.

Interestingly, one putative endoglucanase member of the GH9 family, MBF8377998.1, was also found. The GH9 family comprises some of the most processive and multicatalytic β-1,4 endoglucanases, mixed-linkage endoglucanases, cellobiohydrolases and endoxyloglucanases, primarily from bacteria [39]; they all function by hydrolyzing the glycosidic bond with an inverting mechanism (http://www.cazy.org/GH9.html, accessed on 10 November 2022). GH9 members generally possess a catalytic GH domain (with an α/α 6 barrel-fold containing the active site cleft) and at least one accessory module, such as carbohydrate-binding modules (CBM, that allow binding to insoluble polysaccharides and seem to be related to enzymatic processivity) or Ig-like domains [40,41]. Despite the conservation of aminoacidic sequence, it has not been possible to define a classification of modular structures for this family [39,40]. In fact, a recent analysis of GH9 sequences derived from bacterial metagenomes showed the presence of 19 different modules. Interestingly, only 9% of GH9 members shared five modules, revealing that this family has a more heterogeneous architecture of domains. It has been proposed that such a heterogeneous structure might reflect their ability to work as multifunctional enzymes that can either accommodate various polysaccharides in their active site cleft or exhibit improved activity on the native substrate through increased processivity [4]. The discovery of multifunctional cellulases with different hydrolytic capabilities has an important outcome in the biotechnological application, for the improvement of the enzyme-assisted steps in the lignocellulose biorefinery process. For these reasons, the GH9 protein MBF8377998.1 of 537 aa identified was chosen for in-depth studies; it was named AmCel9.

2.2. *Sequence Analysis of AmCel9*

Using the AmCel9 amino acid sequence as query, BLASTp was used to find sequences with a percentage of identity ranging from 20% to 90%. A multiple alignment was then

obtained with prokaryotic enzymes already characterized, i.e., those from *A. acidocaldarius* (89.3%; CelA; Accession No. WP_012811722.1), *Alicyclobacillus* sp. A4 (48.4%; Alg9; Accession No. ADH59533.1), *Butyrivibrio fibrisolvens* (34.3%; CED1; Accession No. WP_026661505.1), *Paenibacillus curdlanolyticus* (22.5%; PcMulGH9; Accession No. WP_127568011.1), *Thermomonospora fusca* (24.4%; E4; Accession No. WP_011292599.1) and *Clostridium thermocellum* (23.7%; Cel9T; Accession No ABN54011.1) (Figure 2).

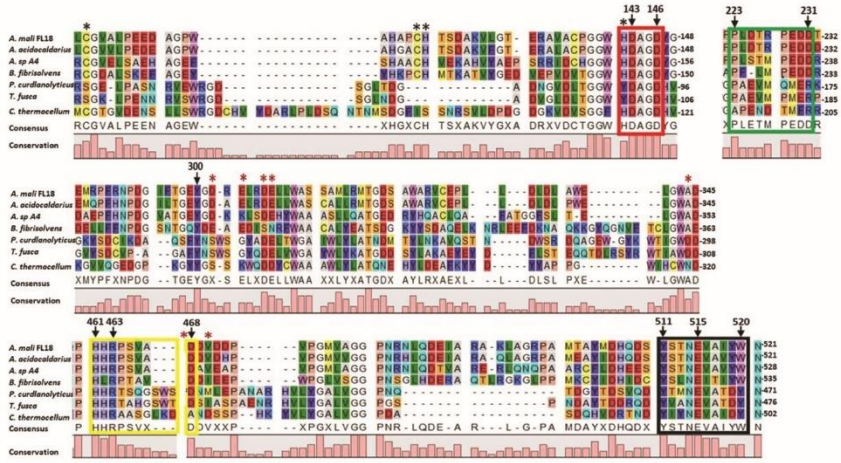

Figure 2. Multiple sequence alignment of AmCel9 aminoacidic sequence (from aa 103 to 521) with other GH9 family members realized with CLC Main Workbench 22.0.2. Red and black boxes highlight conserved catalytic residues; the yellow and green boxes indicate the regions involved in the interaction with Ig-like domain and substrate; the black and red asterisks mark the aa that interact with zinc and calcium ions, respectively. The highly conserved amino acid residues are indicated by arrows.

In the multiple alignment, the red box shows a consensus motif Y/HDAGD typical of GH9 members encompassing the two catalytic aspartic residues (143–146 in AmCel9). Additionally, catalytic glutamate E515 in AmCel9, with a key role as proton donor, is included within a highly conserved motif Y–NEVA-Y/DW/Y (highlighted in the black box, region from aa 511 to 520). The conserved residues Y511, W520 and Y300, as in Cel9A, could be essential to guarantee the right orientation of glutamate to correctly bind the substrate [37]. In the AmCel9 sequence, the residues that in Cel9A are also involved in zinc and calcium binding are conserved; in fact, these ions form coordination bonds with the residues in the catalytic domain labeled with red and black asterisks (Figure 2), contributing to improved thermostability or enzyme substrate affinity [42]. AmCel9 contains a N-terminal Ig-like domain of 85 aa followed by a C-terminus catalytic domain and lacks CBM modules, that are present at the C-terminus of the GH9s of *P. curdlanolyticus* and *T. fusca* [37,43].

To date, the function of the Ig-like domain is not completely clear, but previous studies highlighted its importance in promoting the right folding of the catalytic center or the substrate binding site [44], as for example in CelA of *A. acidocaldarius* where it contributes to enzyme stability during cellulose binding at high temperatures [42]. A recent phylogenetic study has been performed to try to better clarify the role of this module;

since the Ig-like domain is found only in some members of bacterial GH9s, it has been proposed that a common ancestor probably existed, which was lost early during evolution. On the other hand, the presence or absence of other accessory domains (CBM2, CBM3, CBM4/9, dockerin) seems to be independent on the phylogenetic clustering and only the upstream Ig-like module has probably co-evolved with the catalytic domain in the enzymes that maintained it [40]. The alignment shown in Figure 2 supports this hypothesis; in fact, the first four proteins bringing the Ig-like domain at the N-terminus also display the highest similarity in the aminoacidic sequence of the catalytic regions. Moreover, in the characterized Cel9A of *A. acidocaldarius* a dynamic interaction has been evidenced between residues of the Ig-like module and the regions of the catalytic domain corresponding to the "green box" PL—PEDD, and "yellow box", HHRPSVX—and amino acids Y437 and N442 (Figure 2) [42]. It is noteworthy that all the proteins in the alignment contain the yellow box, whereas only the proteins with a N-terminus Ig-like domain (the first four) share the "green box"; for this reason, we also hypothesize that AmCel9 the "green box" of the catalytic domain could interact with the Ig-like domain whereas the "yellow box", could be somewhat implied in substrate binding [37].

Multiple alignment was also used to construct an evolutionary tree to depict the phylogenetic relationship among the proteins investigated (Figure 3).

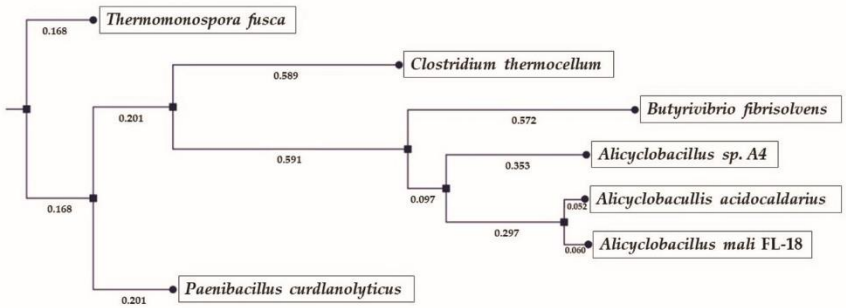

Figure 3. Phylogenetic tree by CLC Main Workbench 22.0.2. The numbers on the nodes correspond to the percentage of bootstrap values for 200 replicates.

The topology of the tree classifies the sequences according to their differences, reflecting their evolutionary distances; in fact, considering the sequence of *A. mali* FL18 as input, the one of *T. fusca* is the farthest from an evolutionary point of view. As expected, the closest is that of *A. acidocaldarius* which shares the highest number of identical aa.

To investigate the structural features of AmCel9, a 3D model of the enzyme was obtained using ColabFold and the obtained structure (Figure 4A) was compared to the crystal structure of Cel9A of *A. acidocaldarius* (Figure 4B) [45].

The two structures resulted in being very similar to each other, both characterized by the Ig-like domain at the N-terminus and the catalytic domain with an $(\alpha/\alpha)_6$ barrel motif at C-terminus. In AmCel9, the Ig-like domain and catalytic domain are linked by a longer loop region near the substrate binding cleft in the three-dimensional structure, probably connected with the endo-activity efficiency [40]. Indeed, Figure 4 shows that the *A. mali* structure lacks two β-sheets connected to the catalytic region that are replaced by a longer loop region; accordingly, E515 is also placed in this region and not in a α-helix [40].

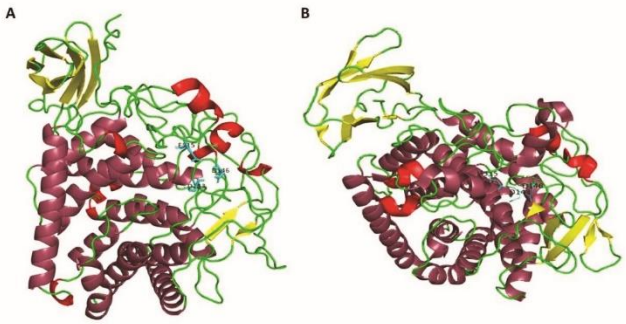

Figure 4. Structural comparison between the predicted 3D structure of AmCel9 (**A**) and the crystal structure of Cel9A of *A. acidocaldarius* (**B**) [45]. For both, the N-terminus Ig-like domain (yellow) is made up of seven β-sheets; the catalytic triad (light blue) is placed in the cleft of the catalytic $(\alpha/\alpha)_6$ barrel domain (purple).

2.3. Recombinant AmCel9: Expression, Purification, and Molecular Weight Analyses

The recombinant N-terminal His-tagged enzyme AmCel9 was expressed in *E. coli* BL21-CodonPlus(DE3) cells transformed with the plasmid pET28b(+)/*AmCel9*. The overexpression was induced in the recombinant cells with 1 mM of isopropyl-1-thio-β-d-galactopyranoside (IPTG), at 37 °C for 16 h, conditions of concentration, temperature and time that resulted in being optimal for the high-level expression of the recombinant protein. Enzyme purification to homogeneity was achieved by a two-step purification procedure (i.e., thermal precipitation at 60 °C and affinity chromatography through a His-trap column), and purified protein was visualized on a SDS-PAGE gel as a band migrating at the expected molecular weight of ~59 kDa (Figure S1 (Supplementary Materials)). The final protein yield was 16 mg/L.

To investigate the quaternary structure of AmCel9, gel-filtration chromatography coupled with a triple-angle light scattering QELS was performed. The results showed that the enzyme in solution has a molecular weight of 58.81 kDa ± 1% (Rh = 7.2 nm ± 3%) indicating that AmCel9 is a monomer resembling the enzyme CelA of the close relative *A. acidocaldarius* (Figure S1) [45].

2.4. pH and Temperature Profile and Stability Properties of AmCel9

To assess the enzymatic activity at different pH values and temperatures, AmCel9 activity was assayed using CMC as the substrate. Although the pH optimum for *A. mali* FL18 growth is pH 4, the purified AmCel9 showed its pH optimum at pH 6.0 and displayed more than 50% activity within the pH range from 5.0 to 7.0 (Figure 5A).

The tolerance within this range of pH is a feature already observed in other characterized processive GH9s and the instability at low pH values could be related to the localization of AmCel9 in the pH-neutral cytoplasm, as also suggested by the lack of a signal sequence for secretion [46,47]. AmCel9 also displayed a significant long-lasting stability at different pH values; indeed, at pH values from 5.0 to 8.0 (Figure 5B) no loss of activity was observed after 4 h of incubation at 4 °C, whereas a residual activity up to 60% was preserved at pH 4.0. Interestingly, after 16 h of incubation at 4 °C, even though the activity drastically decreased at pH 4.0, it was still about 70% at pH values from 5.0 to 8.0. This stability is a peculiar feature since other GH9 members, i.e., Agl9A of *Alicyclobacillus* sp. A4 or CenC of *C. thermocellum*, lost their activity after a few minutes of preincubation at different pH values [48,49].

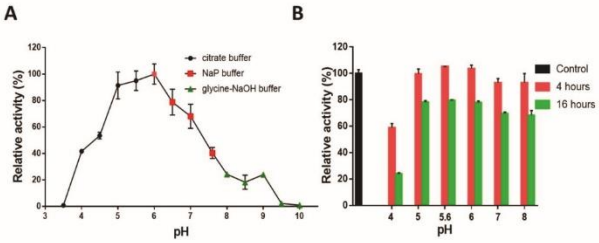

Figure 5. Effect of pH on activity (**A**) and stability (**B**) of AmCel9. (**A**) Black dots indicate the percentage of activity in citrate buffer (pH from 3.5 to 5.5); red squares in Na-P buffer (pH from 6 to 7.5) and green triangles in glycine-NaOH buffer (pH from 8 to 10). (**B**) Relative activities at the optimal pH 6 without incubation (control, black), and after 4 and 16 h of incubation at different pH values (red and green, respectively).

It is of note that the temperature activity profile (Figure 6A) revealed that at pH 6.0 AmCel9 has almost 100% activity at temperatures ranging from 40 °C to 70 °C, and retains more than 50% of activity at 75 °C, a feature also found in Agl9A of *Alicyclobacillus* sp. A4 and in two GH9s of *C. thermocellum* (Cel9T and Cel9D) [46].

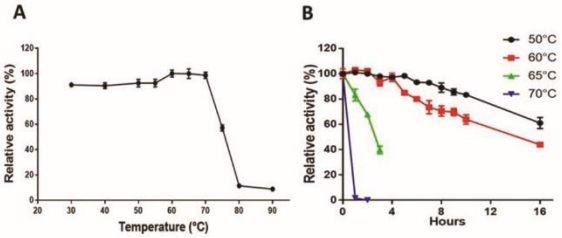

Figure 6. AmCel9 catalytic activity at different temperatures (**A**) and thermo-resistance in the range of 50 °C to 70 °C (**B**).

Regarding the thermo-resistance, the enzyme retained more than 60% and 40% of activity after being incubated at 50 °C and 60 °C for 16 h, respectively, therefore showing good stability (Figure 6B). The comparison of this behavior with the data from the literature is noticeable; for example, the endoglucanases CelA of *A. acidocaldarius* and CenC of *B. licheniformis* despite having their temperature optimum at 70 °C, lost more than 40% of their activity after 2 h of incubation at 60 °C. Moreover, *Bl*Cel9 of *B. licheniformis* also showed long-lasting thermostability at 50 °C, retaining more than 80% of its activity after 24 h; however, at 60 °C, its enzymatic activity dropped significantly to about 40% after only 3 h, whereas AmCel9 was still over 60% active after 10 h [50]. These data suggest that AmCel9 may be considered as an appealing potential biocatalyst in several bioprocesses such as pretreatment of lignocellulosic wastes, requiring high temperatures that are known to favor degradation and substrate solubility; other industrial processes in which pH values and temperature vary, as for example in brewing industry where the use of β-glucanases and β-xylanases (at pH 5.0–6.0 and at 45–65 °C) resulted in an efficient decrease in the viscosity and filtration time [30,51].

2.5. Substrate Specificity and Catalytic Properties

GH9 enzymes generally exhibit heterogeneous hydrolytic capabilities on cellulose; most of them are mainly endoglucanases such as the endo β-1,4-glucanase Cel9A of *A. acidocaldarius* or the endo-β-1,3(4)-glucanase of *Alicyclobacillus* sp. A4 [37,48], but others show exoglucanase activity such as the cellobiohydrolases CbhA and CelK isolated from *C. thermocellum* [52,53]. This family of GHs also includes enzymes that possess multiple catalytic activities, for example, the E4 of *T. fusca* with endo and exo-activities; or bifunctional Cel9X of *C. cellulolyticum* with cellulase/xylanase activities; or the cellulase/mannanase/xylanase activities of *P. curdlanolyticus* B-6 that hydrolyze different polysaccharides [43,54,55].

To assess whether AmCel9 also had a heterogeneous activity towards various substrates, different soluble polysaccharides were tested; the high viscosity of the soluble substrate precluded their preparation at concentrations higher than those utilized (see Section 3). Under these conditions, the best activity displayed by the enzyme was on β-1,4-linked glucans i.e., CMC (512.7 ± 2.0 U/mg) and lichenan (1115 ± 0.01 U/mg); therefore, these substrates were used to evaluate the kinetic parameters of AmCel9. Complete substrate saturation could not be obtained because of the high viscosity of the substrates; hence, the K_M and *V*max values (Table 2) were estimated fitting the data to Michaelis–Menten curve by GraphPad Prism and calculating them from the extrapolated graph (Figure S2).

Table 2. Kinetic parameters of AmCel9 on two substrates. Activity was measured under optimal conditions (pH 6 and temperature of 50 °C) for 1 min.

Substrate	K_M (mg mL^{-1})	k_{cat} (s^{-1})	k_{cat}/K_M (mL mg^{-1} s^{-1})
CMC	10.9	847.2	77.7
Lichenan	7.5	1626.5	216.9

The AmCel9 kinetic parameters indicate a very efficient enzyme on both substrates (Table 2); in fact, results show a high affinity of AmCel9 for CMC and lichenan with a Michaelis constant (K_M) comparable with those determined for other GH9s [8]. The turnover number (k_{cat}) calculated for lichenan was about two-fold higher than that on CMC. Therefore, the k_{cat}/K_M of AmCel9 on this soluble substrate is about three times lower than that on lichenan; the lower enzyme activity could be due to the presence of methyl groups on CMC side chains. Overall, since most other GH9 members have not been characterized to date, the determination of the kinetic parameters of AmCel9 gives new insight into the biochemical features of GH9 family [8].

Unfortunately, an all-embracing comparison of activity values within characterized GH9 members is not feasible because of their heterogeneity as each protein has different specific activity and substrate specificity on several polysaccharides [44]. However, we observed that specific activity of AmCel9 is much higher than that reported for *A. acidocaldarius* (73 U/mg and 109 U/mg on CMC and lichenan, respectively) [37]. In addition, the endoglucanase of *C. thermocellum* (30 U/mg on CMC) [49], GH9b and GH9d enzymes (3.94 U/mg and 1.46 U/mg on CMC, respectively) [40], Alg9 of *Alicyclobacillus* sp. A4 (33.2 U/mg on lichenan) [48] and PcMulGH9 (0.652 U/mg on CMC) have much lower specific activities than AmCel9 [43]. Interestingly, AmCel9 exhibited high activity towards KGM (1330 ±0.012 U/mg), an unusual characteristic for the GH9 members which hydrolyze different types of mannans but with lower specific activity [40,43], and it also showed low hydrolytic activity towards larch arabinogalactan (2.3 ± 0.2 U/mg) and wheat arabinoxylan (1.7 ± 0.6 U/mg). On the other hand, under the same experimental conditions, no activity was detectable on locust-bean gum mannan, guar galactomannan, carob galactomannan, birchwood xylan, beechwood xylan, laminarin, salicin or cellobiose. Overall, the results suggest that AmCel9 is mainly able to hydrolyze β-1,4-glycosidic bonds, and it does not possess significant β-glucosidase activity. Moreover, since AmCel9 possesses β-1,4-endoglucanase activity with three polysaccharide types it can be considered to be a multifunctional cellulase according to a recent definition proposed by Glasgow et al. [4].

Family GH9 also embraces catalysts that are able to degrade insoluble cellulose; to date, the presence of CBM modules is known to be correlated to the enzyme processivity on microcrystalline cellulose [4,40]; for example, PcMulGH9 of *P. curdlanolyticus* and *Bl*Cel9 of *B. licheniformis*, both harboring CBM modules, have specific activities towards Avicel of 28.57 U/g and 40 U/g, respectively [43,50]. On the other hand, enzymes that share the same modular structure of AmCel9 (Ig-like and GH domains) have much lower or a lack of activity on microcrystalline cellulose as in the case of CelA of *A. acidocaldarius*, Agl9A of *Alicyclobacillus* sp. A4 or BP_Cel9 [37,44,48]. Very interestingly, the hydrolytic activity of AmCel9 evaluated on circle paper units of 0.5 mm × 0.5 mm (Whatman® filter paper) and microcrystalline cellulose (Avicel PH101), showed a good enzyme capability to release glucose with values of specific activity of 68.7 ± 0.02 and 62.98 ± 0.05 U/g, respectively, comparable to those reported for PcMulGH9 or *Bl*Cel9 [43,50]. It can be hypothesized that the high activity observed towards heterogeneous solid Avicel could be due to endoglucanase activity on the amorphous cellulose areas of the substrate [56].

The unexpected activity towards insoluble cellulose along with the broad substrate specificity and high catalytic efficiency make AmCel9 a promising multifunctional catalyst, that might more efficiently break down lignocellulosic material without requiring non-eco-friendly pre-treatments.

3. Materials and Methods

3.1. Functional Annotation of A. mali FL18 CAZymes

The genome of *A. mali* FL18 was previously annotated by the Rapid Annotation Subsystem Technology (RAST) and deposited in NCBI (JADPKZ000000000) and IMG-IM (https://img.jgi.doe.gov/cgi-bin/mer/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=2931470914 accessed on 10 November 2022). CAZyme searches were performed using db-CAN metaserver (https://bcb.unl.edu/dbCAN2/ accessed on 10 November 2022) with integrated: (i) HMMER database; (ii) DIAMOND for fast blast hits in the CAZy database and (iii) HMMER for dbCAN-sub [32].

3.2. Reagents and Substrates

Carboxymethyl cellulose sodium salt (CMC), Avicel (PH-101), glucose, cellobiose, salicin, locust bean gum mannan and laminarin were purchased from Sigma-Aldrich (St. Louis, MO, USA). Konjac glucomannan (KGM), birchwood xylan, beechwood xylan, lichenan, guar galactomannan, carob galactomannan, larch arabinogalactan and wheat arabinoxylan were purchased from Megazyme (Bray, Co. Wicklow, Ireland). Whatman® filter paper (75 mm × 100 mm) was cut into circle paper units of 0.5 mm × 0.5 mm. Kanamycin and chloramphenicol were purchased from PanReac AppliChem (Ottoweg, Darmstadt, Germany).

3.3. AmCel9 In Silico Analyses

The aminoacidic sequence of the putative GH9 of *A. mali* FL18, with the Accession No. MBF8377998.1, was compared to sequences of GH9s already characterized available in the GeneBank database by utilizing BLASTp server (https://blast.ncbi.nlm.nih.gov/Blast.cgi accessed on 14 September 2022).

To identify homologous proteins and conserved domains in AmCel9, a multiple alignment and a phylogenetic tree were performed with CLC Main Workbench 22.0.2 software (QIAGEN, Hilden, Germany) by using the amino acid sequences of: Cel9A from *A. acidocaldarius*; Alg9 from *Alicyclobacillus* sp. A4; CED1 from *B. fibrisolvens*; PcMulGH9 from *P. curdlanolyticus*; E4 from *T. fusca* and of Cel9T from *C. thermocellum* [37,43,46,48,54]. A 3D model of AmCel9 was generated using the Colab server (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb accessed on 14 September 2022); this server predicts protein structures from their sequences using a simplified version of AlphaFold v2.0 [57].

3.4. Expression and Purification of Recombinant AmCel9

The plasmid pET28b(+)/*AmCel9*, which contains the gene encoding AmCel9 with a six-His tag at its N-terminus, was purchased from GenScript Biotech (Piscataway, NJ, USA) and used to transform *E. coli* BL21-CodonPlus(DE3) cells (Stratagene, San Diego, CA, USA). These cells were cultured in LB medium supplemented with kanamycin (50 µg/mL) and chloramphenicol (33 µg/mL). Protein expression was induced for 16 h by adding 1 mM of IPTG (PanReac AppliChem, Ottoweg, Darmstadt, Germany) when the cultures reached 0.7 OD_{600}/mL. Cells were harvested by centrifugation at 6750× *g* and resuspended in 20 mL of buffer A (50 mM Tris-HCl pH 8, 300 mM NaCl, 20 mM imidazole, DTT 1 mM) supplemented with lysozyme 0.2 mg/mL (PanReac AppliChem, Ottoweg, Darmstadt, Germany) and with protease inhibitor cocktail 1X (Roche, Basilea, Switzerland). The resuspended cells were lysed with one cycle of freeze-thaw with the following procedure: (i) incubation for 90 min on an orbital shaker at 37 °C; (ii) freezing with ethanol and dry ice for 15 min and (iii) re-incubation for 15 min at 37 °C. After a centrifugation step at 35,000× *g* for 45 min at 4 °C, the purification of AmCel9 was set up through a two-step procedure. A first thermal precipitation at 60 °C for 10 min was followed by centrifugation at 35,000× *g* for 45 min at 4 °C then followed by affinity chromatography on a His-Trap™ column (1 mL, GE Healthcare, Chicago, IL, USA) connected to an AKTA Explorer system. The elution was performed with buffer A supplemented with a linear gradient of imidazole (0–500 mM), and all the peak fractions were pooled and then dialyzed against a storage buffer (50 mM of Tris-HCl pH 8, 300 mM of NaCl and 1 mM of DTT) using an Ultracel® 10 kDa ultrafiltration disc (Sigma-Aldrich, St. Louis, MO, USA) connected to an Amicon® stirred cell. Bradford assay was used to estimate protein concentration using bovine serum albumin as standard. The purity degree of AmCel9 were appraised by SDS-PAGE analyses (12%) after staining the gel with Coomassie brilliant blue R-250 (Biorad, Hercules, CA, USA).

3.5. Determination of AmCel9 Native Molecular Weight

The purified AmCel9 was further analyzed by using a size-exclusion chromatography system connected to Mini DAWN Treos light-scattering system (Wyatt Technology, Santa Barbara, CA, USA) equipped with a QELS module (quasi-elastic light scattering module) and with mass value and hydrodynamic radius (Rh) measurements. One milliliter of protein (0.6 mg/mL) was loaded on a Superdex S200 column (10/300; GE Healthcare, Chicago, IL, USA) equilibrated in buffer 50 mM Tris-HCl pH 8, 200 mM NaCl, 1mM DTT, with a flowrate of 0.5 mL/min. Data were analyzed using Astra 5.3.4.14 software (Wyatt Technology, Santa Barbara, CA, USA) [58].

3.6. Enzyme Assay, Temperature and pH Profile

AmCel9 standard assay was set up at 50 °C and using CMC as substrate; the assay mixture (total volume of 40 µL) contained 1.6% (*w*/*v*) CMC in assay buffer (50 mM Na-P pH 6, 200 mM NaCl and 3 mM DTT). A total of 1 µg of enzyme was added to the substrate mix, pre-incubated for 10 min at 50 °C and the reaction stopped after 1 min at 50 °C in ice for 5 min. Then, 160 µL of 3,5-dinitrosalicylic acid (DNS) was added to the reaction mix, which was transferred in a 96-well microplate and placed in a thermomixer (Eppendorf, Hamburg, Germany) at 100 °C for 20 min [59,60]. After a cooling step at 4 °C of 20 min, the concentration of reducing sugars in solution was determined by measuring A_{540nm} in a 96-well microplate reader (Synergy H4, Biotek, Agilent, Santa Clara, CA, USA) and by interpolating absorbance data with a standard curve that was prepared according to Kim et al. [59]. One unit of enzyme activity (U) was defined as the amount of enzyme required to release 1 µmol of glucose-reducing equivalent in one minute, under the described assay conditions.

The optimal pH was determined by measuring enzyme activity in different buffers from pH 3.0 to 10.0 (each 50 mM), with increments of 0.5 pH units at 50 °C for one minute, using the DNS assay described above. Citrate buffer (pH 3.0–5.5), Na-P buffer (pH 6.0–7.5) and glycine-NaOH buffer (pH 8.0–10) were used. Temperature optimum was measured

by assaying enzyme activity at pH 6.0 from 40 °C to 90 °C with an increase of 10 °C; each experiment was performed in triplicate and results were reported as relative activity (%).

To test the pH stability of recombinant AmCel9, it was preincubated in various buffers from pH 4.0 to 8.0 at 4 °C for 4 and 16 h; residual activity of the enzyme was measured under standard assay conditions. Thermo-resistance was determined by measuring residual activity up to 16 h (at 1-h intervals), pre-incubating the purified enzyme without the substrate at temperatures of 50 °C, 55 °C, 60 °C, 65 °C and 70 °C. All the experiments were performed in triplicate and the 100% was the enzyme activity in the absence of pre-incubation.

3.7. Substrate Specificity and Kinetic Studies

The substrate specificity of AmCel9 was tested by using the standard assay (1 µg of enzyme, at 50 °C and pH 6 for 1 min) on several polymers: cellobiose (1.5%), lichenan (1.4%), laminarin (1.5%), locust bean gum mannan (0.5%), KGM (0.7%), birchwood xylan (0.9%), beechwood xylan (0.9%), guar galactomannan (0.7%), carob galactomannan (0.7%), larch arabinogalactan (0.7%) and wheat arabinoxylan, prepared in assay buffer (50 mM Na-P pH 6, 200 mM NaCl and 3 mM DTT) or in water, according to the specific data-sheet of each substrate. The released reducing monosaccharides were quantified using the DNS method described above, except for cellobiose, for which the amount of free glucose was estimated by using the D-Glucose Assay Kit (GOPOD Format, Megazyme, Bray, Co. Wicklow, Ireland) according to the manufacturer's protocol. The assay on CMC (1.6%) was used as the control; one unit (U) was defined as the amount of enzyme necessary to release 1 µmol of reducing sugars per minute at 50 °C and pH 6; each experiment was performed in triplicate.

The kinetic parameters were determined using CMC and lichenan as substrates: 0.4 µg of AmCel9 was tested with CMC (0.4%, 0.6%, 0.8%, 1%, 1.2%, 1.4%, 1.6%) and 0.2 µg with lichenan (0.2%, 0.4%, 0.6%, 0.8%, 1%, 1.2%, 1.4%). The Michaelis–Menten constant K_M was calculated by non-linear regression analysis using GraphPad 7.0 (Prism software) and each experiment was performed in triplicate.

The hydrolytic activity of the enzyme was also investigated on insoluble celluloses; the assays were performed in a total volume of 500 µL, by incubating 2.5 mg of Avicel (PH-101) or 1 circle paper unit of 0.5 mm × 0.5 mm (Whatman® filter paper) with 35 µg of purified enzyme in the assay buffer (50 mM Na-P pH 6, 200 mM NaCl and 3 mM DTT) at 55 °C for 1 h on an orbital shaker at 200 rpm. To measure the amount of glucose released, 500 µL of DNS solution were added to each sample and the mixture boiled for 10 min. After a cooling step at 4 °C for 30 min, the A_{540nm} was measured and data were interpolated with a glucose standard curve to measure the concentration of the reducing sugar in solution [49,61]. For each reaction, a non-enzyme control was set up and each experiment was performed in quadruplicate.

4. Conclusions

In this study, the potential of the thermoacidophilic bacterium *A. mali* FL18 as a reservoir of CAZymes was investigated. Genome-wide analysis and functional annotation were performed to identify the key contributors in the lignocellulose decomposition and to provide insight into the genes correlated with lignocellulolytic activity. In particular, the analysis of CAZome revealed the presence of 32 putative GHs, indicating that *A. mali* FL18 possesses a palette of interesting activities for the biodegradation of the holocellulosic component of the cell wall. In silico analyses suggest that at least four of these GHs might be involved in cellulose degradation; among them, a novel thermophilic GH9, AmCel9, was biochemically characterized. AmCel9 exhibits a peculiar activity in a broad range of temperatures, as well as thermo-resistance and stability in a wide range of pH values. Moreover, AmCel9 hydrolyses β-glycosidic linkages of synthetic and natural celluloses with relevant specific activities, and, to the best of our knowledge, is the first GH9 member without a CBM module that is also able to hydrolyze insoluble celluloses. Overall, the data

presented in this work prove that AmCel9 is a catalyst deserving further exploitation; in fact, it can be employed either as multifunctional enzyme to ameliorate the degradation of lignocellulosic biomass for the conversion of recalcitrant residues into value-added products or to specifically hydrolyze crystalline cellulose.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms24010243/s1.

Author Contributions: M.C. performed the experiments; M.C. and M.A. performed the analysis; M.C., M.A., G.G., D.L., P.C. and G.F. supervised the project; M.C., M.A. and G.F. drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by GoodbyWaste: ObtainGOOD products–exploit BY-products–reduce WASTE, MIUR 2017-JTNK78.006, Italy and by PRIN 2017-PANACEA CUP: E69E19000530001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank Simonetta Bartolucci and Emilia Pedone for helpful scientific discussion. Luciano Pirone is also acknowledged for the determination of the quaternary structure.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Aich, S.; Datta, S. Efficient Utilization of Lignocellulosic Biomass: Hydrolysis Methods for Biorefineries. In *Biorefineries: A Step Towards Renewable and Clean Energy*; Verma, P., Ed.; Springer: Singapore, 2020; pp. 273–295. ISBN 978-981-15-9593-6.
2. Gallo, G.; Puopolo, R.; Carbonaro, M.; Maresca, E.; Fiorentino, G. Extremophiles, a Nifty Tool to Face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering. *Int. J. Environ. Res. Public Health* **2021**, *18*, 5228. [CrossRef] [PubMed]
3. Piras, C.C.; Fernández-Prieto, S.; de Borggraeve, W.M. Ball Milling: A Green Technology for the Preparation and Functionalisation of Nanocellulose Derivatives. *Nanoscale Adv.* **2019**, *1*, 937–947. [CrossRef] [PubMed]
4. Glasgow, E.; vander Meulen, K.; Kuch, N.; Fox, B.G. Multifunctional Cellulases Are Potent, Versatile Tools for a Renewable Bioeconomy. *Curr. Opin. Biotechnol.* **2021**, *67*, 141–148. [CrossRef] [PubMed]
5. Salzano, F.; Aulitto, M.; Fiorentino, G.; Pedone, E.; Contursi, P.; Limauro, D. *Alicyclobacillus mali* FL18 as a Novel Source of Glycosyl Hydrolases: Characterization of a New Thermophilic β-Xylosidase Tolerant to Monosaccharides. *Int. J. Mol. Sci.* **2022**, *23*, 14310. [CrossRef] [PubMed]
6. Drula, E.; Garron, M.-L.; Dogan, S.; Lombard, V.; Henrissat, B.; Terrapon, N. The Carbohydrate-Active Enzyme Database: Functions and Literature. *Nucleic Acids Res.* **2022**, *50*, D571–D577. [CrossRef]
7. Andlar, M.; Rezić, T.; Marđetko, N.; Kracher, D.; Ludwig, R.; Šantek, B. Lignocellulose Degradation: An Overview of Fungi and Fungal Enzymes Involved in Lignocellulose Degradation. *Eng. Life Sci.* **2018**, *18*, 768–778. [CrossRef]
8. Konar, A.; Aich, S.; Katakojwala, R.; Datta, S.; Mohan, S.V. A Processive GH9 Family Endoglucanase of *Bacillus licheniformis* and the Role of Its Carbohydrate-Binding Domain. *Appl. Microbiol. Biotechnol.* **2022**, *106*, 6059–6075. [CrossRef]
9. Ajeje, S.B.; Hu, Y.; Song, G.; Peter, S.B.; Afful, R.G.; Sun, F.; Asadollahi, M.A.; Amiri, H.; Abdulkhani, A.; Sun, H. Thermostable Cellulases/Xylanases from Thermophilic and Hyperthermophilic Microorganisms: Current Perspective. *Front. Bioeng. Biotechnol.* **2021**, *9*, 794304. [CrossRef]
10. Botturi, A.; Battista, F.; Andreolli, M.; Faccenda, F.; Fusco, S.; Bolzonella, D.; Lampis, S.; Frison, N. Polyhydroxyalkanoated-Rich Microbial Cells from Bio-Based Volatile Fatty Acids as Potential Ingredient for Aquaculture Feed. *Energies* **2021**, *14*, 38. [CrossRef]
11. Wang, R.; Lorantfy, B.; Fusco, S.; Olsson, L.; Franzén, C.J. Analysis of Methods for Quantifying Yeast Cell Concentration in Complex Lignocellulosic Fermentation Processes. *Sci. Rep.* **2021**, *11*, 11293. [CrossRef]
12. Fusco, F.A.; Ronca, R.; Fiorentino, G.; Pedone, E.; Contursi, P.; Bartolucci, S.; Limauro, D. Biochemical Characterization of a Thermostable Endomannanase/Endoglucanase from Dictyoglomus Turgidum. *Extremophiles* **2018**, *22*, 131–140. [CrossRef] [PubMed]
13. Antonucci, I.; Gallo, G.; Limauro, D.; Contursi, P.; Ribeiro, A.L.; Blesa, A.; Berenguer, J.; Bartolucci, S.; Fiorentino, G. Characterization of a Promiscuous Cadmium and Arsenic Resistance Mechanism in Thermus Thermophilus HB27 and Potential Application of a Novel Bioreporter System. *Microb Cell Fact.* **2018**, *17*, 78. [CrossRef] [PubMed]
14. Koga, Y. Thermal Adaptation of the Archaeal and Bacterial Lipid Membranes. *Archaea* **2012**, *2012*, 789652. [CrossRef] [PubMed]
15. Morana, A.; Esposito, A.; Maurelli, L.; Ruggiero, G.; Ionata, E.; Rossi, M.; la Cara, F. A Novel Thermoacidophilic Cellulase from *Alicyclobacillus acidocaldarius*. *Protein Pept. Lett.* **2008**, *15*, 1017–1021. [CrossRef] [PubMed]

16. Bai, Y.; Wang, J.; Zhang, Z.; Shi, P.; Luo, H.; Huang, H.; Luo, C.; Yao, B. Expression of an Extremely Acidic β-1,4-Glucanase from Thermoacidophilic *Alicyclobacillus* sp. A4 in Pichia Pastoris Is Improved by Truncating the Gene Sequence. *Microb Cell Fact.* **2010**, *9*, 33. [CrossRef] [PubMed]
17. Aulitto, M.; Gallo, G.; Puopolo, R.; Mormone, A.; Limauro, D.; Contursi, P.; Piochi, M.; Bartolucci, S.; Fiorentino, G. Genomic Insight of *Alicyclobacillus mali* FL18 Isolated from an Arsenic-Rich Hot Spring. *Front. Microbiol.* **2021**, *12*, 669. [CrossRef] [PubMed]
18. Aulitto, M.; Fusco, S.; Limauro, D.; Fiorentino, G.; Bartolucci, S.; Contursi, P. Galactomannan Degradation by Thermophilic Enzymes: A Hot Topic for Biotechnological Applications. *World J. Microbiol. Biotechnol.* **2019**, *35*, 32. [CrossRef]
19. Aulitto, M.; Fusco, F.A.; Fiorentino, G.; Bartolucci, S.; Contursi, P.; Limauro, D. A Thermophilic Enzymatic Cocktail for Galactomannans Degradation. *Enzyme Microb Technol.* **2018**, *111*, 7–11. [CrossRef]
20. Ing, N.; Deng, K.; Chen, Y.; Aulitto, M.; Gin, J.W.; Pham, T.L.M.; Petzold, C.J.; Singer, S.W.; Bowen, B.; Sale, K.L.; et al. A Multiplexed Nanostructure-Initiator Mass Spectrometry (NIMS) Assay for Simultaneously Detecting Glycosyl Hydrolase and Lignin Modifying Enzyme Activities. *Sci. Rep.* **2021**, *11*, 11803. [CrossRef]
21. Dadwal, A.; Sharma, S.; Satyanarayana, T. Thermostable Cellulose Saccharifying Microbial Enzymes: Characteristics, Recent Advances and Biotechnological Applications. *Int. J. Biol. Macromol.* **2021**, *188*, 226–244. [CrossRef]
22. Turner, P.; Mamo, G.; Karlsson, E.N. Potential and Utilization of Thermophiles and Thermostable Enzymes in Biorefining. *Microb Cell Fact.* **2007**, *6*, 9. [CrossRef] [PubMed]
23. Politi, J.; Spadavecchia, J.; Fiorentino, G.; Antonucci, I.; Casale, S.; de Stefano, L. Interaction of Thermus Thermophilus ArsC Enzyme and Gold Nanoparticles Naked-Eye Assays Speciation between As(III) and As(V). *Nanotechnology* **2015**, *26*, 435703. [CrossRef] [PubMed]
24. Aulitto, M.; Tom, L.M.; Ceja-Navarro, J.A.; Simmons, B.A.; Singer, S.W. Whole-Genome Sequence of *Brevibacillus borstelensis* SDM, Isolated from a Sorghum-Adapted Microbial Community. *Microbiol. Resour. Announc.* **2020**, *9*, 15–16. [CrossRef] [PubMed]
25. Tom, L.M.; Aulitto, M.; Wu, Y.W.; Deng, K.; Gao, Y.; Xiao, N.; Rodriguez, B.G.; Louime, C.; Northen, T.R.; Eudes, A.; et al. Low—Abundance Populations Distinguish Microbiome Performance in Plant Cell Wall Deconstruction. *Microbiome* **2022**, *10*, 183. [CrossRef] [PubMed]
26. Aulitto, M.; Strazzulli, A.; Sansone, F.; Cozzolino, F.; Monti, M.; Moracci, M.; Fiorentino, G.; Limauro, D.; Bartolucci, S.; Contursi, P. Prebiotic Properties of *Bacillus coagulans* MA-13: Production of Galactoside Hydrolyzing Enzymes and Characterization of the Transglycosylation Properties of a GH42 β-Galactosidase. *Microb Cell Fact.* **2021**, *20*, 71. [CrossRef]
27. Aulitto, M.; Fusco, S.; Franzén, C.J.; Strazzulli, A.; Moracci, M.; Bartolucci, S.; Contursi, P. Draft Genome Sequence of *Bacillus coagulans* Ma-13, a Thermophilic Lactic Acid Producer from Lignocellulose. *Microbiol. Resour. Announc.* **2019**, *8*, e00341-19. [CrossRef]
28. Fiorentino, G.; Cannio, R.; Rossi, M.; Bartolucci, S. Decreasing the Stability and Changing the Substrate Specificity of the *Bacillus stearothermophilus* Alcohol Dehydrogenase by Single Amino Acid Replacements. *Protein Eng. Des. Sel.* **1998**, *11*, 925–930. [CrossRef]
29. Si, M.; Liu, D.; Liu, M.; Yan, X.; Gao, C.; Chai, L.; Shi, Y. Complementary Effect of Combined Bacterial-Chemical Pretreatment to Promote Enzymatic Digestibility of Lignocellulose Biomass. *Bioresour. Technol.* **2019**, *272*, 275–280. [CrossRef]
30. Viikari, L.; Alapuranen, M.; Puranen, T.; Vehmaanperä, J.; Siika-Aho, M. Thermostable Enzymes in Lignocellulose Hydrolysis. *Adv. Biochem. Eng. Biotechnol.* **2007**, *108*, 121–145. [CrossRef]
31. Zuliani, L.; Serpico, A.; de Simone, M.; Frison, N.; Fusco, S. Biorefinery Gets Hot: Thermophilic Enzymes and Microorganisms for Second-Generation Bioethanol Production. *Processes* **2021**, *9*, 1583. [CrossRef]
32. Huang, L.; Zhang, H.; Wu, P.; Entwistle, S.; Li, X.; Yohe, T.; Yi, H.; Yang, Z.; Yin, Y. DbCAN-Seq: A Database of Carbohydrate-Active Enzyme (CAZyme) Sequence and Annotation. *Nucleic Acids Res.* **2018**, *46*, D516–D521. [CrossRef] [PubMed]
33. Aulitto, M.; Martinez-Alvarez, L.; Fiorentino, G.; Limauro, D.; Peng, X.; Contursi, P. A Comparative Analysis of *Weizmannia coagulans* Genomes Unravels the Genetic Potential for Biotechnological Applications. *Int. J. Mol. Sci.* **2022**, *23*, 3135. [CrossRef] [PubMed]
34. Tong, X.; Qi, Z.; Zheng, D.; Pei, J.; Li, Q.; Zhao, L. High-Level Expression of a Novel Multifunctional GH3 Family β-Xylosidase/α-Arabinosidase/β-Glucosidase from Dictyoglomus Turgidum in *Escherichia coli*. *Bioorg. Chem.* **2021**, *111*. [CrossRef] [PubMed]
35. Wu, Y.; Mao, G.; Fan, H.; Song, A.; Zhang, Y.H.P.; Chen, H. Biochemical Properties of GH94 Cellodextrin Phosphorylase THA_1941 from a Thermophilic Eubacterium Thermosipho Africanus TCF52B with Cellobiose Phosphorylase Activity. *Sci. Rep.* **2017**, *7*, 4849. [CrossRef] [PubMed]
36. Eckert, K.; Schneider, E. A Thermoacidophilic Endoglucanase (CelB) from *Alicyclobacillus acidocaldarius* Displays High Sequence Similarity to Arabinofuranosidases Belonging to Family 51 of Glycoside Hydrolases. *Eur. J. Biochem.* **2003**, *270*, 3593–3602. [CrossRef] [PubMed]
37. Eckert, K.; Zielinski, F.; lo Leggio, L.; Schneider, E. Gene Cloning, Sequencing, and Characterization of a Family 9 Endoglucanase (CelA) with an Unusual Pattern of Activity from the Thermoacidophile *Alicyclobacillus acidocaldarius* ATCC27009. *Appl. Microbiol. Biotechnol.* **2002**, *60*, 428–436. [CrossRef]
38. Yeoman, C.J.; Fields, C.J.; Lepercq, P.; Ruiz, P.; Forano, E.; White, B.A.; Mosoni, P. In Vivo Competitions between *Fibrobacter succinogenes*, *Ruminococcus flavefaciens*, and *Ruminoccus albus* in a Gnotobiotic Sheep Model Revealed by Multi-Omic Analyses. *mBio* **2021**, *12*, e03533-20. [CrossRef]

39. Gilad, R.; Rabinovich, L.; Yaron, S.; Bayer, E.A.; Lamed, R.; Gilbert, H.J.; Shoham, Y. CelI, a Noncellulosomal Family 9 Enzyme from *Clostridium thermocellum*, Is a Processive Endoglucanase That Degrades Crystalline Cellulose. *J. Bacteriol.* **2003**, *185*, 391–398. [CrossRef]
40. Aymé, L.; Hébert, A.; Henrissat, B.; Lombard, V.; Franche, N.; Perret, S.; Jourdier, E.; Heiss-Blanquet, S. Characterization of Three Bacterial Glycoside Hydrolase Family 9 Endoglucanases with Different Modular Architectures Isolated from a Compost Metagenome. *Biochim. Biophys. Acta (BBA) Gen. Subj.* **2021**, *1865*, 129848. [CrossRef]
41. Wu, L.; Davies, G.J. Structure of the GH9 Glucosidase/Glucosaminidase from Vibrio Cholerae. *Acta Crystallogr. Sect. F* **2018**, *74*, 512–523. [CrossRef]
42. Liu, H.; Pereira, J.H.; Adams, P.D.; Sapra, R.; Simmons, B.A.; Sale, K.L. Molecular Simulations Provide New Insights into the Role of the Accessory Immunoglobulin-like Domain of Cel9A. *FEBS Lett.* **2010**, *584*, 3431–3435. [CrossRef] [PubMed]
43. Phakeenuya, V.; Ratanakhanokchai, K.; Kosugi, A.; Tachaapaikoon, C. A Novel Multifunctional GH9 Enzyme from *Paenibacillus curdlanolyticus* B-6 Exhibiting Endo/Exo Functions of Cellulase, Mannanase and Xylanase Activities. *Appl. Microbiol. Biotechnol.* **2020**, *104*, 2079–2096. [CrossRef] [PubMed]
44. Schröder, C.; Burkhardt, C.; Busch, P.; Schirrmacher, G.; Claren, J.; Antranikian, G. Characterization of a Theme C Glycoside Hydrolase Family 9 Endo-β-Glucanase from a Biogas Reactor Metagenome. *Protein J.* **2018**, *37*, 454–460. [CrossRef] [PubMed]
45. Pereira, J.H.; Sapra, R.; Volponi, J.V.; Kozina, C.L.; Simmons, B.; Adams, P.D. Structure of Endoglucanase Cel9A from the Thermoacidophilic *Alicyclobacillus acidocaldarius*. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *65*, 744–750. [CrossRef]
46. Kesavulu, M.M.; Tsai, J.Y.; Lee, H.L.; Liang, P.H.; Hsiao, C.D. Structure of the Catalytic Domain of the *Clostridium thermocellum* Cellulase CelT. *Acta Crystallogr. D Biol. Crystallogr.* **2012**, *68*, 310–320. [CrossRef]
47. Bakker, E.P. The Role of Alkali-Cation Transport in Energy Coupling of Neutrophilic and Acidophilic Bacteria: An Assessment of Methods and Concepts. *FEMS Microbiol. Rev.* **1990**, *6*, 319–334. [CrossRef]
48. Bai, Y.; Wang, J.; Zhang, Z.; Shi, P.; Luo, H.; Huang, H.; Luo, C.; Yao, B. A Novel Family 9 β-1,3(4)-Glucanase from Thermoacidophilic *Alicyclobacillus* sp. A4 with Potential Applications in the Brewing Industry. *Appl. Microbiol. Biotechnol.* **2010**, *87*, 251–259. [CrossRef]
49. Akram, F.; Khan, M.A.; Hussain, Z.; Nawaz, A.; Iqbal, K.; Shah, A.J. CenC, a Multidomain Thermostable GH9 Processive Endoglucanase from *Clostridium thermocellum*: Cloning, Characterization and Saccharification Studies. *World J. Microbiol. Biotechnol.* **2015**, *31*, 1699–1710. [CrossRef]
50. de Araújo, E.A.; de Oliveira Neto, M.; Polikarpov, I. Biochemical Characterization and Low-Resolution SAXS Structure of Two-Domain Endoglucanase BlCel9 from *Bacillus licheniformis*. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 1275–1287. [CrossRef]
51. Celestino, K.R.S.; Cunha, R.B.; Felix, C.R. Characterization of a β-Glucanase Produced by *Rhizopus microsporus* Var. *microsporus*, and Its Potential for Application in the Brewing Industry. *BMC Biochem.* **2006**, *7*, 23. [CrossRef]
52. Schubot, F.D.; Kataeva, I.A.; Chang, J.; Shah, A.K.; Ljungdahl, L.G.; Rose, J.P.; Wang, B.C. Structural Basis for the Exocellulase Activity of the Cellobiohydrolase CbhA from *Clostridium thermocellum*. *Biochemistry* **2004**, *43*, 1163–1170. [CrossRef] [PubMed]
53. Kataeva, I.; Li, X.-L.; Chen, H.; Choi, S.-K.; Ljungdahl, L.G. Cloning and Sequence Analysis of a New Cellulase Gene Encoding CelK, a Major Cellulosome Component of *Clostridium thermocellum*: Evidence for Gene Duplication and Recombination. *J. Bacteriol.* **1999**, *181*, 5288–5295. [CrossRef] [PubMed]
54. Sakon, J.; Irwin, D.; Wilson, D.B.; Karplus, P.A. Structure and Mechanism of Endo/Exocellulase E4 from *Thermomonospora fusca*. *Nat. Struct. Biol.* **1997**, *4*, 810–818. [CrossRef] [PubMed]
55. Ravachol, J.; Borne, R.; Tardif, C.; de Philip, P.; Fierobe, H.P. Characterization of All Family-9 Glycoside Hydrolases Synthesized by the Cellulosome-Producing Bacterium *Clostridium cellulolyticum*. *J. Biol. Chem.* **2014**, *289*, 7335–7348. [CrossRef] [PubMed]
56. Gao, S.; You, C.; Renneckar, S.; Bao, J.; Zhang, Y.-H.P. New Insights into Enzymatic Hydrolysis of Heterogeneous Cellulose by Using Carbohydrate-Binding Module 3 Containing GFP and Carbohydrate-Binding Module 17 Containing CFP. *Biotechnol. Biofuels* **2014**, *7*, 24. [CrossRef]
57. Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: Making Protein Folding Accessible to All. *Nat. Methods* **2022**, *19*, 679–682. [CrossRef]
58. Pirone, L.; Pitzer, J.E.; D'Abrosca, G.; Fattorusso, R.; Malgieri, G.; Pedone, E.M.; Pedone, P.V.; Roop, R.M.; Baglivo, I. Identifying the Region Responsible for *Brucella abortus* MucR Higher-Order Oligomer Formation and Examining Its Role in Gene Regulation. *Sci. Rep.* **2018**, *8*, 17238. [CrossRef]
59. Kim, J.J.; Kwon, Y.K.; Kim, J.H.; Heo, S.J.; Lee, Y.; Lee, S.J.; Shim, W.B.; Jung, W.K.; Hyun, J.H.; Kwon, K.K.; et al. Effective Microwell Plate-Based Screening Method for Microbes Producing Cellulase and Xylanase and Its Application. *J. Microbiol. Biotechnol.* **2014**, *24*, 1559–1565. [CrossRef]
60. Miller, G.L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. *Anal. Chem.* **1959**, *31*, 426–428. [CrossRef]
61. Ogonda, L.A.; Saumonneau, A.; Dion, M.; Muge, E.K.; Wamalwa, B.M.; Mulaa, F.J.; Tellier, C. Characterization and Engineering of Two New GH9 and GH48 Cellulases from a *Bacillus pumilus* Isolated from Lake Bogoria. *Biotechnol. Lett.* **2021**, *43*, 691–700. [CrossRef]

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Chapter 5 - Insights into the CE15 enzymes from *Dyadobacter fermentans* suggest distinct functional roles *(submitted)*

In this chapter the study focuses on enzymes active towards lignin-carbohydrate complexes (LCCs) tightly associated through covalent bonds formed during lignification. The breakage of these bonds poses a challenge to the complete degradation of lignocellulose into valuable compounds. Cleavage of these bonds is a costly and important step in converting lignocellulosic materials into useful monosaccharides and oligosaccharides. Glucuronoyl esterases (GEs), enzymes of the carbohydrate esterase family 15 (CE15), are known to catalyze the breakdown of ester bonds between lignin and glucuronoxylan LCCs.

Here, three CE15 enzymes from the bacterium *Dyadobacter fermentans* NS114T, which grows endophytically in maize, were studied. These enzymes, namely *Df*CE15A, -B, -C, exhibit different genomic associations and were investigated for their role in lignocellulose degradation. The study involved heterologous expression, purification, and testing of the enzymes on synthetic substrates. The enzymes showed a preference for ester substrates containing glucuronic acid. Structural and amino acid differences between the enzymes were investigated to understand their catalytic mechanisms. Mutagenesis studies revealed unique functionalities, especially in *Df*CE15A.

The overall goal of this research is to contribute to the understanding of a relatively under-characterized enzyme family (CE15), which is critical for the development of technologies aimed at efficient utilization of lignocellulosic wastes. The study has implications for the advancement of lignocellulosic waste valorization technologies.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Obscured due to 12-month embargo request.

Concluding remarks

This thesis explores the potential of extremophilic microorganisms, particularly thermophiles, in various industrial and environmental applications. By exploiting their unique enzymatic capabilities, these microorganisms offer promising approaches for biocatalysis, biotransformation and bioremediation. Key areas of focus include understanding resistance mechanisms to pollutants, such as arsenic, and investigating enzymes involved in lignocellulose degradation.
This study investigates *Tt*ArsM, a methyltransferase involved in arsenic detoxification of ***T. thermophilus*** **HB27** and elucidates its biochemical properties and regulatory mechanisms. Using CRISPR-Cas9 technology, mutants of this bacterium confirm the *Tt*ArsM importance and allow the development of a whole-cell bioreporter for arsenic detection. In addition, functional annotations of the genomes of two bacteria, ***A. mali*** **FL18** and ***G. stearothermophilus*** **GF16**, from extreme environments show potential for lignocellulose hydrolysis, offering sustainable solutions for biomass valorization.
Finally, **glucuronoyl esterases** (GEs) from *Dyadobacter fermentans* NS114^T are investigated for their role in the degradation of lignin-carbohydrate complexes. Through biochemical and structural analyses, insights into substrate preferences and catalytic mechanisms are gained to advance lignocellulosic waste valorization efforts.
In summary, this research sheds light on the diverse applications of extremophilic microorganisms and demonstrates their versatility in addressing environmental challenges and contributing to sustainable industrial processes.

Acknowledgments

I want to thank Professor Johan Larsbrink for giving me the opportunity to work in his research team during my stay at Chalmers University and for supervising my research activities. I also want to express my gratitude to Dr. Scott Mazurkewich for his assistance in supervising my research as a visiting PhD student.

Acknowledgements go also to grants from:

- SIB - Società Italiana di Biochimica e Biologia Molecolare
- GoodbyWaste: Obtain GOOD products – exploit BY-products – reduce WASTE.

Appendix I - Lists of publications

- Gallo G., Puopolo R., **Carbonaro M.**, Maresca E., Fiorentino G.
  Extremophiles, a Nifty Tool to Face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering
  2021, Int J Environ Res Public Health, Vol. 18: 5228
  https://doi.org/10.3390/ijerph18105228

- Gallo G., Mougiakos I., Bianco M., **Carbonaro M.**, Carpentieri A., Illiano A., Pucci P., Bartolucci S., Van der Oost J., Fiorentino G.
  A Hyperthermoactive-Cas9 Editing Tool Reveals the Role of a Unique Arsenite Methyltransferase in the Arsenic Resistance System of *Thermus thermophilus* HB27
  2021, mBio, Vol. 12: e02813-2
  https://doi.org/10.1128/mBio.02813-21

- **Carbonaro M.**, Aulitto M., Gallo G., Contursi P., Limauro D., Fiorentino G.
  Insight into CAZymes of *Alicyclobacillus mali* FL18: Characterization of a New Multifunctional GH9 Enzyme
  2023, Int J Mol Sci, Vol. 24: 243
  https://doi.org/10.3390/ijms24010243

- **Carbonaro M.**, Mazurkewich S., Fiorentino G., Lo Leggio L., Larsbrink J.
  Insights into the CE15 enzymes from *D. fermentans* suggest distinct functional roles
  2024, *(Submitted to Applied Microbiology and Biotechnology)*

- **Carbonaro M.**, Aulitto M., Mazurkewich S., Di Fraia A., Contursi P., Limauro D., Larsbrink J., Fiorentino G.
  Identification of the *Geobacillus stearothermophilus* GF16 genes involved in lignocellulose degradation and utilization of secreted enzymes for efficient degradation of hemicellulose-rich biomasses *(Manuscript in preparation)*

Author's contribution

- Extremophiles, a Nifty Tool to Face Environmental Pollution: From Exploitation of Metabolism to Genome Engineering

Conceptualization, G.G., R.P., **M.C.**, E.M., and G.F.; investigation, G.G., R.P., **M.C.**, and E.M.; data curation, G.G., R.P., **M.C.**, and E.M.; figure preparation, R.P. and **M.C.**; writing—original draft preparation, G.G., R.P., **M.C.**, E.M., and G.F.; writing—review and editing, G.G., R.P., **M.C.**, E.M., and G.F.; visualization, G.G., R.P., **M.C.**, E.M., and G.F.; supervision, G.F.; project administration, G.F.; funding acquisition, G.F. All authors have read and agreed to the published version of the manuscript.

- A Hyperthermoactive-Cas9 Editing Tool Reveals the Role of a Unique Arsenite Methyltransferase in the Arsenic Resistance System of *Thermus thermophilus* HB27

  G.G., I.M., J.V.D.O and G.F. conceived the project. G.G, I.M., M.B., **M.C.**, and A.I. performed the experiments; G.G, and I.M., performed the analysis; G.G, and I.M., wrote the manuscript; G.G, I.M., J.V.D.O., S.B. and G.F. supervised the project; G.G, I.M., **M.C.**, A.C., P.P., J.V.D.O., S.B., and G.F. revised manuscript drafts. All authors have read and agreed to the published version of the manuscript.

- Insight into CAZymes of *Alicyclobacillus mali* FL18: Characterization of a New Multifunctional GH9 Enzyme

  G.F. and **M.C.** conceived the project. **M.C.** performed the experiments; **M.C.** and M.A. performed the analysis; **M.C.**, M.A., G.G., D.L., P.C. and G.F. supervised the project; **M.C.**, M.A. and G.F. drafted the manuscript. All authors have read and agreed to the published version of the manuscript.

- Insights into the carbohydrate esterase family 15 enzymes from *D. fermentans* suggest distinct roles

  S.M. and J.L. conceived the project. **M.C.** and S.M. completed all the experimental work; **M.C.** and S.M. wrote all the manuscript; G.F., L.L.L., and J.L. critically appraised and revised manuscript drafts. All authors read and approved the final manuscript. (*In submission*)

- Identification of the *Geobacillus stearothermophilus* GF16 genes involved in lignocellulose degradation and utilization of secreted enzymes for efficient degradation of hemicellulose-rich biomasses

G.F. and M.A. conceived the project. **M.C**., M.A., A.D.F. and S.M. performed the experiments; **M.C.** and M.A. performed the analysis; **M.C.** and M.A. wrote the manuscript; **M.C.**, M.A., D.L., J.L. and G.F. supervised the project; **M.C.**, M.A., D.L., S.M., J.L. and G.F. revised manuscript drafts. (*In preparation*)

Appendix II - Participation to congresses

List of posters

- G. Gallo, I. Mougiakos, M. Bianco, **M. Carbonaro**, S. Bartolucci, J. van der Oost, G. Fiorentino.
  "Development of a Cas9-based tool for the genome editing of the hyperthermophilic bacterium Thermus thermophilus HB27".
  FEBS2021. The 45th Virtual congress. Online, July 3-8, 2021.

- G. Gallo, M. Aulitto, **M. Carbonaro**, F. Salzano, D. Limauro, P. Contursi, S. Bartolucci and G. Fiorentino.
  "*Alicyclobacillus mali FL18: a new polysaccharide waste-degrading extremophilic bacterium isolated from a hydrothermal hot-spring in Southern Italy".*
  FEBS2021. The 45th Virtual Congress. Online, July 3-8, 2021.

- G. Gallo, I. Mougiakos, M. Bianco, **M. Carbonaro**, S. Bartolucci, J. van der Oost, G. Fiorentino.
  "A hyperthermoactive-Cas9 editing tool for genome manipulation of Thermus thermophilus HB27"
  The 61st SIB Congress 2021. Online, September 23-24, 2021.

- G. Gallo, **M. Carbonaro**, I. Mougiakos, J. van der Oost, J. Berenguer, S. Bartolucci and G. Fiorentino.
  "TtSmtB the playmaker of the arsenic-cadmium resistance system in Thermus thermophilus HB27".
  Regulation and Signal Transduction in Prokaryotes (RSTIP2021). Online, November 15-16, 2021.

- **M. Carbonaro**, M. Aulitto, M. Torre, P. Contursi, D. Limauro and G. Fiorentino.
  "Characterization of a multifunctional cellulase, a new GH9 member from Alicyclobacillus mali FL-18".
  Meeting SIB Campania 2022. Naples, October 26, 2022.

- **M. Carbonaro**, S. Mazurkewich, G. Fiorentino and J. Larsbrink
  "Biochemical characterization of three bacterial glucuronoyl esterases from Dyadobacter fermentans NS114T"
  The 62nd SIB Congress 2023. Florence, September 7-9, 2023.

- **M. Carbonaro**, S. Mazurkewich, G. Fiorentino and J. Larsbrink

"Biochemical characterization of three bacterial glucuronoyl esterases from Dyadobacter fermentans NS114T"
4th Memorial Workshop Maria Ciaramella. Naples, November 16-17, 2023.

List of oral talks

- **M. Carbonaro**, M. Aulitto, M. Torre, P. Contursi, D. Limauro and G. Fiorentino.
  "From hot pool to enzyme discovery: a new multifunctional GH9 enzyme from Alicyclobacillus mali FL-18".
  SIB group. Trends in Biotechnology: the SIB group perspectives 2022. Naples, June 23-24, 2022.

- **M. Carbonaro**, M. Aulitto, M. Torre, P. Contursi, D. Limauro and G. Fiorentino
  "More than just a cellulase: AmCel9 a new multifunctional GH9 enzyme from Alicyclobacillus mali FL-18".
  XVI FISV Congress. 3R: Research, Resilience, Reprise. Naples, September 14-16, 2022.

Congresses organization

- Organization as **Member of the Department of Biology** (Università degli Studi di Napoli Federico II) of the Scientific Manifestation "La Notte Europea Dei Ricercatori".
  Polo Universitario Monte Sant'Angelo (Naples), September 29, 2023.

- Organization as **Member of the Department of Biology** (Università degli Studi di Napoli Federico II) of the Scientific Manifestation "XXXVII Edizione di Futuro Remoto - INTELLIGENZE".
  Organismi termofili: sistemi intelligenti per un mondo ecosostenibile – Città della Scienza (Naples), November 21-26, 2023.

Appendix III - Experience in foreign laboratory

I joined for six months (01/02/2023 to 30/07/2023) the group headed by Johan Larsbrink at the Division of Industrial Biotechnology within the Department of Life Sciences at Chalmers University of Technology, Goteborg, Sweden. I worked on projects regarding the characterization of enzymes involved in the cleavage of covalent bonds between lignin and carbohydrates in plant cell walls, which contribute to the recalcitrance of renewable biomass. The work that I have undertaken regarded the cloning of genes encoding novel carboxyl esterases from *Dyadobacter fermentans*, their protein production, purification, structural analyses, and biochemical characterization on various synthetic model substrates. Moreover, variants of these enzymes were produced and characterised.

CHALMERS Gothenburg, Dec 19th, 2022

Dear Professor Fiorentino,

With this letter I would like to formally invite and welcome your student Miriam Carbonaro to our laboratory at Chalmers University of Technology for up to six months during 2023. She is welcome to join my group from Feb 1st (full duration 01/02/2023 – 31/07/2023) and partake in ongoing projects as well as have a chance to combine her own research with ours, as we have discussed in previous meetings. The main project she will be involved in revolves around enzymatic cleavage of covalent connections between lignin and carbohydrates in plant cell walls which contribute to the recalcitrance of renewable biomass. The work will consist of molecular biology (cloning of novel enzyme-coding genes), protein production and purification, and characterization of the new enzymes on various synthetic model substrates and biomass samples, which are all activities routinely performed in my group.

I hope this exchange can lead to a fruitful collaboration between our laboratories and result in future joint publications and grant proposals.

Sincerely,

Johan Larsbrink, PhD, Associate Professor
Division of Industrial Biotechnology
Department of Biology and Biological Engineering

Contact: johan.larsbrink@chalmers.se

BIOLOGY AND BIOLOGICAL ENGINEERING
Chalmers University of Technology
SE 412 96 Göteborg, Sweden
Visiting address: Kemivägen 10
Phone: +46 (0)31 772 38 39
E-mail: johan.larsbrink@chalmers.se