DEVELOPMENT OF BIOSENSORS FOR MONITORING OF ENVIRONMENTAL POLLUTANTS

Giovanni Gallo

Dottorato in Biotecnologie 32° ciclo

Università di Napoli Federico II

Dottorato in Biotecnologie 32° ciclo

Università di Napoli Federico II

DEVELOPMENT OF BIOSENSORS FOR MONITORING OF ENVIRONMENTAL POLLUTANTS

Giovanni Gallo

Dottorando: Giovanni Gallo

Relatori: Prof. Simonetta Bartolucci
Prof. Gabriella Fiorentino

Coordinatore: Prof. Marco Moracci

Settore Scientifico Disciplinare Biochimica

Bio10

“Never say never, because limits, like fears, are often just an illusion.”

(Michael Jeffrey Jordan)

INDEX

RIASSUNTO

Introduzione

I metalli pesanti sono elementi con un elevato peso atomico, una densità almeno cinque volte maggiore di quella dell'acqua e tossici a basse concentrazioni. Le loro molteplici applicazioni industriali, agricole, mediche e tecnologiche hanno portato ad un'ampia distribuzione nell'ambiente, causando inquinamento ambientale e problemi per la salute umana.

I metalli pesanti sono naturalmente presenti in alcuni ambienti del nostro pianeta, come le sorgenti geotermali (marine e terrestri), habitat considerati estremi poiché, oltre ad elevate concentrazioni di metalli, sono presenti una varietà di condizioni estreme, quali temperatura e/o pH, alta concentrazione di metallo e/o sale. In tali habitat si sono evoluti diversi microrganismi estremofili appartenenti ai domini dei *Bacteria* e *Archaea* che possiedono una straordinaria varietà di meccanismi dinamici e versatili per contrastare stress chimici e fisici e sopravvivere in condizioni ambientali considerate ostili per altre forme di vita. Essi dunque possiedono sistemi metabolici anche per contrastare la tossicità dei metalli pesanti e pertanto partecipano attivamente alla loro mobilizzazione e conversione chimica. Tali caratteristiche hanno attratto gli scienziati anche da un punto di vista applicativo. In questo lavoro di tesi ci siamo occupati di analizzare i meccanismi molecolari responsabili della sopravvivenza e dell'adattamento dei batteri termofili ai metalli pesanti, con particolare enfasi su arseniato As (V), arsenito As (III), cadmio Cd (II) e alla messa a punto di sistemi biotecnologici per il loro rilevamento.

I sistemi di resistenza ai metalli pesanti

I microrganismi si evolvono rapidamente e sono in grado di adattarsi alle variazioni delle condizioni ambientali grazie alla loro propensione al trasferimento genico orizzontale e ai loro alti tassi di crescita. I microrganismi hanno dunque acquisito nel corso dell'evoluzione una vasta gamma di meccanismi di difesa contro i metalli presenti nell'ambiente naturale, tra cui fattori trascrizionali metallo-sensibili che regolano la trascrizione dei geni che codificano per le proteine responsabili della detossificazione, sequestro, efflusso e assorbimento dei metalli.

Generalmente in un microrganismo un sistema di detossificazione è costituito da una proteina sensore, che è un regolatore trascrizionale dei geni per la resistenza ai metalli, enzimi, come reduttasi o metiltransferasi che modificano lo stato di ossidazione dello ione o aggiungono gruppi metilici, e proteine di membrana (trasportatori attivi e/o passivi) per l'efflusso per rimuovere direttamente i metalli dalla cellula. Nei batteri, il sistema di resistenza all'arsenico è uno dei meglio caratterizzati, ed è generalmente mediato dai prodotti genici dell'operone *ars*. Sebbene l'organizzazione dell'operone vari considerevolmente tra le diverse specie, alcuni geni chiave sono sempre presenti. L'organizzazione basale dell'operone è quella dell'operone *arsRBC*, che è presente nel genoma di *E. coli* e nel plasmide di *Staphylococcus aureus* pI2589 e che conferisce moderata resistenza a As(V), As(III) e Sb(III). È stato anche descritto in *E. coli* un operone organizzato in una serie di cinque geni, detto operone *arsRDABC*, che fornisce resistenza a concentrazioni di arsenico più elevate. Presenta infatti il gene *arsC* che codifica per un'arseniato reduttasi citoplasmatica che converte l'arseniato in arsenito, che viene estruso dalla cellula dall'azione di un trasportatore transmembrana codificato dal gene *arsB*. Negli organismi che hanno il gene *arsA*, il prodotto genico, un ATPasi, è accoppiato al trasportatore transmembrana ArsB e aumenta significativamente i livelli di resistenza. Il gene *arsR* codifica per un repressore appartenente alla famiglia ArsR/SmtB, coinvolto nel rilevamento dei metalli e nella regolazione trascrizionale dei geni della resistenza; il prodotto del gene *arsD* funge invece da *metallochaperone* aiutando il trasferimento di arsenito alla subunità ArsA del complesso ArsAB attivandolo. In alcuni batteri è presente un gene accessorio che codifica per una arsenito metiltransferasi, ArsM, questo enzima è in grado di produrre delle forme metilate dell'arsenico che possono diffondere dalla cellula oppure essere estruse tramite trasportatori di membrana.

Thermus thermophilus come organismo modello

Thermus thermophilus viene utilizzato da molti gruppi di ricerca come organismo modello a causa dei suoi alti tassi di crescita, delle alte rese cellulari delle colture, della disponibilità di strumenti genetici e dell'espressione costitutiva di un apparato di competenza naturale straordinariamente efficiente. Il ceppo HB27 di *Thermus thermophilus* (isolato dal prof. Oshima in una pozza termale in Giappone) è un Gram-negativo aerobico che cresce a temperature

comprese tra 50° e 82° C. Le cellule di *T. thermophilus* HB27 appaiono morfologicamente come sottili bacilli, che tendono a formare filamenti in grado di dividersi per divisione binaria. Il genoma di *T. thermophilus* HB27 è composto da un cromosoma (TTC) di 1.894.877 bp e un megaplasmide (TTP) di 232.605 bp, chiamato pTT27. Il contenuto di GC è medialmente del 69,4%.

Caratterizzazione dei sistemi di resistenza all'arsenico di *T. thermophilus* HB27

Nel laboratorio dove ho svolto la mia tesi di dottorato era già stata iniziata la caratterizzazione del sistema di resistenza all'arsenico basandosi sull'evidenza che *T. thermophilus* HB27 è un microrganismo tollerante all'arsenico. L'analisi del genoma ha evidenziato tre geni codificanti i putativi elementi di resistenza: un'arseniato reduttasi *Tt*ArsC (codificata da TTC1502), un regolatore trascrizionale *Tt*SmtB (codificato da TTC0353) e un'ATPasi di membrana (di tipo P1B) *Tt*ArsX (codificata da TTC0354) responsabile dell'efflusso. Il presente lavoro di tesi è stato rivolto alla comprensione delle basi molecolari del sistema di resistenza all'arsenico attraverso la caratterizzazione strutturale e funzionale delle sue componenti, anche allo scopo di utilizzarle per la messa a punto di biosensori per il rilevamento di metalli pesanti. Tra i risultati ottenuti durante questo progetto c'è stata la caratterizzazione di *Tt*SmtB, un componente della famiglia ArsR/SmtB. I membri di questa famiglia sono molto diversi tra loro, ma condividono alcune caratteristiche, tra cui la struttura dimerica, la presenza di un dominio di legame al DNA (basato su un motivo strutturale *helix-turn-helix*) e un dominio di legame al metallo caratterizzato dalla presenza di cisteine (almeno due). Dall'approfondito studio di *Tt*SmtB, del suo ruolo come regolatore e delle sue caratteristiche chimico-fisiche, è emerso che questa proteina fa parte di un sistema di resistenza promiscuo che è utilizzato dalla cellula per detossificare sia l'arsenico che il cadmio. Sia *Tt*SmtB che *Tt*ArsX sono direttamente coinvolte nel processo di resistenza e infatti i ceppi mutanti che presentano delezioni di questi due geni risultano essere più sensibili sia al cadmio che all'arsenico (nelle forme ossidata e ridotta). Utilizzando tecniche di pull down associate ad analisi proteomica abbiamo identificato un'ulteriore componente del sistema di resistenza all'arsenico, cioè una arsenito metiltransferasi (*Tt*ArsM). L'identificazione e caratterizzazione strutturale e funzionale di questa proteina ci ha permesso di capire

ulteriormente come funziona e come è regolato il *pathway* della resistenza all'arsenico. *Tt*SmtB è un repressore trascrizionale, che in assenza di arsenico/cadmio, impedisce l'espressione dei geni coinvolti nella resistenza. Abbiamo dimostrato che è infatti in grado di interagire con il suo stesso promotore e con i promotori di *Tt*ArsC, *Tt*ArsX, *Tt*ArsM e che l'interazione con i DNA bersaglio è impedita in presenza di cadmio o arseniato o arsenito. Inoltre, abbiamo individuato un nuovo livello di regolazione post traduzionale nel meccanismo di tolleranza all'arsenico. L'interazione fisica tra *Tt*SmtB e *Tt*ArsM, dimostrata attraverso esperimenti *in vitro* di CoIP, a basse concentrazioni di arseniato o arsenito mantiene bloccato *Tt*ArsM, invece, ad elevate concentrazioni di As(III) o As(V) il complesso si dissocia rendendo *Tt*ArsM attiva enzimaticamente. In presenza di cadmio, metallo che non viene riconosciuto da *Tt*ArsM, il complesso *Tt*SmtB /*Tt*ArsM viene ulteriormente stabilizzato, mantenendo la totale inattivazione dell'enzima *Tt*ArsM. Questo ulteriore sistema di regolazione è dovuto, probabilmente, ad un'esigenza cellulare di salvaguardare il suo potenziale redox in condizioni di stress ossidativo (causato dalla presenza di cadmio). La caratterizzazione di questo complesso sistema, in tutte le sue parti, è stata fondamentale per la realizzazione di nuovi sistemi di *biosensing* per il monitoraggio ambientale.

Sviluppo di biosensori per il monitoraggio di arsenico e cadmio

I biosensori sono dei dispositivi costituiti da una componente biologica (cellula, proteina, DNA) in grado di interagire con un dato analita e da un trasduttore che trasforma l'interazione in un segnale misurabile. Questi dispositivi consentono una determinazione rapida, diretta e selettiva di una particolare specie chimica anche in presenza di molecole chimicamente simili. La specificità della risposta di questi dispositivi, inoltre, in genere minimizza o elimina la necessità di manipolazioni preliminari del campione e rappresenta un'alternativa all'uso di sistemi analitici costosi e sofisticati, quali, per esempio, la spettrometria di massa.

In questa tesi di dottorato, basandosi sulle conoscenze ottenute sul sistema di resistenza all'arsenico in *T. thermophilus* HB27, sono stati sviluppati due *whole-cells biosensors* per il monitoraggio di arsenico e cadmio. In particolare, il primo *bioreporter* è stato ottenuto a partire dal plasmide pMH*pnqo*, clonando: a) una cassetta contenente il promotore riconosciuto da *Tt*SmtB, a valle del gene reporter codificante una versione termostabile della β-galattosidasi,

e b) una cassetta di espressione per il fattore *Tt*SmtB sotto il controllo del promotore costitutivo *pnqo*. Questo sistema reporter è stato utilizzato per trasformare il ceppo *wilde type* di *T. thermophilus* HB27 e un ceppo mutante (*T. thermophilus* TTC0354::pK18) e, data la capacità di *Tt*SmtB di interagire sia con il cadmio che con l'arsenico (quest'ultimo negli stadi di ossidazione +3 e +5), è stato in grado di rilevare 20 µM arsenico e 10 µM di cadmio. Il secondo biosensore "*whole-cell*" di *T. thermophilus* HB27 è stato messo a punto utilizzando un sistema di "*genome editing*" basato su CRISPR-Cas9, in collaborazione con il Prof. John van der Oost della Wageningen University & Research. La cellula modificata geneticamente attraverso il sistema di *editing* genomico esprime il gene reporter *sYFP*, codificante la proteina fluorescente YFP, a valle del promotore di *Tt*ArsX, utilizzando il sistema endogeno di resistenza all'arsenico e al cadmio, e rileva arseniato, arsenito e cadmio a concentrazioni di circa 0.5 µM. Questo sistema è risultato essere più stabile e riproducibile nel segnale di fluorescenza. Quindi può rappresentare un ottimo candidato per ulteriori caratterizzazioni ai fini di un impiego applicativo.
Infine, è stato identificato e caratterizzato un nuovo ceppo di *Geobacillus stearothermophilus*, che mostra un'elevata resistenza all'arseniato, che rappresenta un ottimo candidato per lo sviluppo di nuovi sistemi di risanamento ambientale.
In conclusione, i risultati ottenuti con questo progetto di dottorato ci hanno permesso di comprendere in maniera più dettagliata il funzionamento dei sistemi cellulari di resistenza all'arsenico e di sviluppare contestualmente dei *bioreporter* che possano fungere da punto di partenza per ulteriori manipolazioni volte alla realizzazione di biosensori da utilizzare nei processi industriali o per il monitoraggio ambientale.

SUMMARY

The high concentrations of heavy metals are the cause of one of the most serious pollution problems of our time. Different chemical or physical methods have been developed to treat polluted sites, but the problem persists given the necessity to dispose the treated waste and any, probably even more toxic, chemical by-product; so, considering advances in biotechnology, there is interest in the development of biosensors for monitoring polluted areas as well as bioremediation techniques.

The aim of this PhD thesis was the development of *whole-cell* biosensors for monitoring of environmental pollutants starting from the characterization of the molecular mechanisms of heavy metal resistance in thermophilic microorganisms.

The first part of this PhD-thesis focuses on the characterization of the *Tt*SmtB, a trans-acting repressor belonging to the ArsR/SmtB family, involved in the metal sensing and in transcriptional regulation.

The characterization of *Tt*SmtB was carried out in order to investigate the molecular determinants of metal binding, DNA binding and to identify target promoter regions.

The second part of this PhD thesis was focused on the development of *whole-cell* biosensors using two different approaches: plasmid-based or genome-based detection systems.

The first whole-cell biosensor was obtained starting from the knowledge acquired on the biological function of *Tt*SmtB, where on a single plasmid we can found the promoter recognized by *Tt*SmtB and the reporter gene.

The second whole-cell biosensor was obtained by engineering *Thermus thermophilus* HB27 through a CRISPR-Cas9 system that allowed us to insert the reporter gene into the genome, under the control of a responsive promoter, allowing us to improve the sensitivity of our reporter system.

The last part of this project was focused on the characterization of a new strain of *Geobacillus stearothermophilus* was isolated from a hot spring located in a solfataric area of Pozzuoli (Napoli).

Chapter I: General Introduction

1. Environmental pollution

Heavy metals are generally defined as metals with relatively high densities, atomic weights, or atomic numbers. Here we refer to heavy metals when they present the following characteristics:

- density exceeding 5.0 g/cm^3
- general behaviour as cations
- low solubility of their hydrates
- aptitude to form complexes
- affinity towards the sulphides

Trace of some heavy metals are required for certain biological processes such as: iron (Fe) and copper (Cu) required for the oxygen and electron transport; or vanadium (V) and manganese (Mn) required for the enzyme regulation[1].

In high concentrations heavy metals are the cause of one of the most serious pollution problems of our time. Therefore, it is important to be informed about the heavy metals and to adopt protective measures against overexposure.

Heavy metal toxicity can result in damaged or reduced mental and central nervous function, lower energy levels, and damage to blood composition, lungs, kidneys, liver, and other vital organs. Long-term exposure may result in slowly progressing physical, muscular, and neurological degenerative processes. Repeated long-term contact with some metals or their compounds may even cause cancer (International Occupational Safety and Health Information Centre, 1999)[2].

Different chemical or physical methods have been developed to treat polluted sites, but the problem persists given the necessity to dispose the treated waste and any, probably even more toxic, chemical byproducts; so, taking into account advances in biotechnology, it's interesting trying to develop biosensors for monitoring heavy metal pollution or to follow the efficacy of bioremediation procedures.

2. Chemistry of Arsenic

Arsenic is an ubiquitous toxic metalloid that contaminates both groundwater sources and soils. More than 100 million people in the

world are at risk from consuming water contaminated by arsenic, and strategies to detect and prevent this global problem are urgently required.

From a chemical point of view, arsenic is a metalloid of the VA group of the periodic table, toxic to the environment and humans[3]. In nature it is present in four different oxidation states: arsenite As(III) and arsenate As(V) are the most common, while the elemental arsenic (0) and arsenide As(-3) are rare[4,5]. The main chemical-physical factors that control the speciation of arsenic are the pH and the redox potential[6]. The metalloid is found in sediments, water and soil both in inorganic form, where in oxidizing conditions prevail the pentavalent arsenic and under reducing conditions the trivalent one[7] and in organic form as the mono-methylarsonic acid (MMA) and di-methylarsinic acid (DMA). Disclosure of organic arsenic is lower than inorganic forms. Since arsenic has different valence states, its treatment and removal from the contaminated sites is not an easy issue.

3. Microbial resistance system to heavy metals

Microorganisms quickly evolve and quickly adapt to environmentally changing conditions, due to their propensity for horizontal gene transfer and their high growth rates. The great genetic diversity of microorganisms accounts for their great metabolic versatility and ability to adapt even to extreme environments that do not allow proliferation of other living organisms[8]. Microorganisms have a wide variety of defence mechanisms against metals present in the natural environment, in fact they have developed metal-responsive transcriptional regulatory proteins that regulate the transcription of genes encoding proteins responsible for metal detoxification, sequestration, efflux and uptake (Fig. 1). Generally, a microbial detoxification system consists of a sensor protein, which is a transcriptional regulator of the genes for resistance to metals, a detoxifying enzyme, such as a reductase, and an efflux pump to remove the metals from the cell directly.

Recent structural studies have revealed five distinct families of metal sensors proteins: MerR, ArsR/SmtB, DtxR, Fur and NikR. The families MerR and ArsR/SmtB regulate the expression of genes needed for metal ion detoxification, efflux and sequestration; here, metal binding results in the activation (MerR) or derepression (ArsR/SmtB) of the resistance operon. In contrast, the families DtxR, Fur and NikR regulate genes coding for proteins involved in metal

ion uptake; in these cases, the metal ions function as co-repressors to switch off uptake genes in defined concentrations of metals[9].

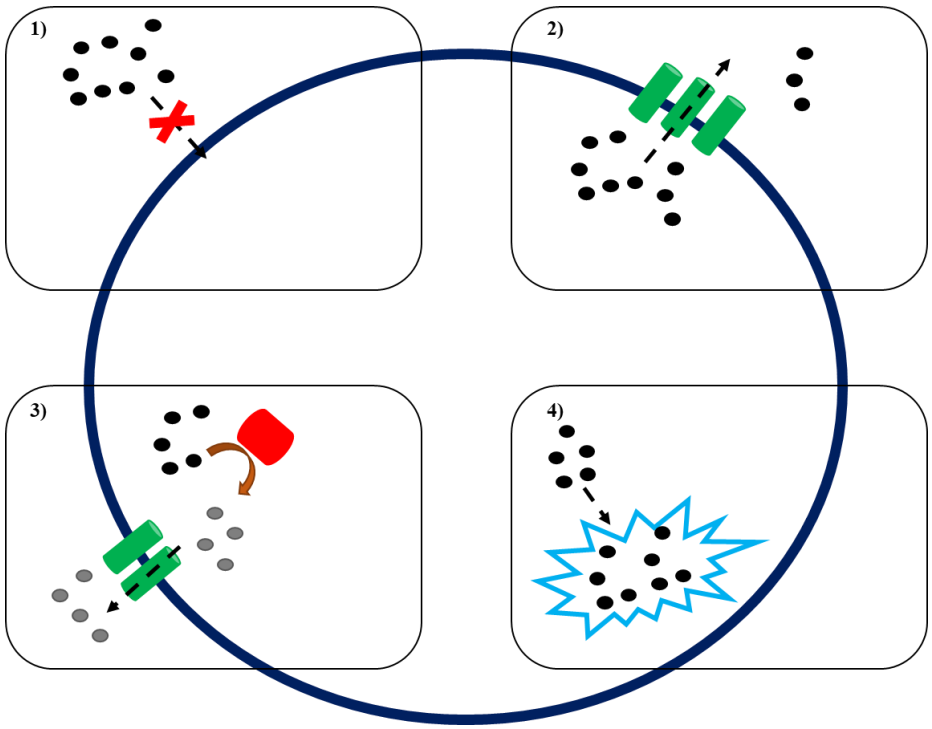


Fig. 1 Generalized illustration of the mechanism of resistance to toxic metals by microorganisms: 1) Extracellular barrier, a selectively permeable system; 2) Efflux of metal ions; 3) Enzymatic reduction of metal ions. 4) Intracellular sequestration by small molecule complexing agents or metal-chelating proteins.

Arsenic tolerance in bacteria is usually mediated by the gene products of *ars* operon[5,10,11]. Although the right operon organization varies considerably among different species, there are some key genes which are always present: the gene set which confers a basal resistance consists in the *arsRBC* operon, present in the *E. coli* genome[11] and in the *Staphylococcus aureus* plasmid pI258[12]. There is also a set of five genes, *arsRDABC* found in the *E. coli* plasmid R773[13], that provides resistance to higher arsenic concentrations. The gene *arsC* encodes a small cytoplasmic arsenate reductase which converts arsenate into arsenite[14], that is extruded from the cell by the action of a transmembrane transporter encoded by the *arsB* gene. In organisms that have the *arsA* gene, its gene product, an

ATPase, is coupled to transmembrane transporter ArsB and significantly increases the levels of resistance[5]. The *arsR* gene encodes a trans-acting repressor belonging to the ArsR/SmtB family, involved in the metal sensing and in transcriptional regulation; the *arsD* product acts as metallochaperon assisting the transfer of arsenite to the ArsA subunit of the ArsAB complex activating it.

Some bacteria possess another gene in the *ars* operon, coding an arsenite methyltransferase. Arsenic methyltransferase is an enzyme that catalyzes the conversion of inorganic arsenic into mono-, di-, and tri-methylated products. Conversion of inorganic arsenic into methylated metabolites affects the environmental transport and the fate of arsenic as well as its metabolism and disposition at cellular and systemic levels [5] in higher organisms.

4. Environmental biomonitoring

There are currently two widely used methods of arsenic detection in drinking water: laboratory-based analytical methods and field-based testing methods. The laboratory-based analytical methods require highly trained personnel and expensive analytical machinery, such as inductively coupled plasma mass spectroscopy (ICP-MS) and atomic absorption spectroscopy. Further, the delay in turnaround time between specimen collection and result availability limits their day-to-day use. So different chemical methods have been developed to treat polluted sites, but the problem persists given the necessity to dispose the treated waste and any, probably even more toxic, chemical bioproducts. Biosensors and bioreporters are safe, alternative methods to detect environmental pollutants such as arsenic[15].

A biosensor is an analytical device, used for the detection of an analyte, that combines a biological component with a physicochemical detector[16]. In a biosensor, the bioreceptor is designed to interact with the specific analyte of interest to produce an effect measurable by the transducer. High selectivity for the analyte among a matrix of other chemical or biological components is a key requirement of the bioreceptor. While the type of biomolecule used can vary widely, biosensors can be classified according to the most common types of bioreceptor interactions involving: antibody/antigen[17], enzymes/ligands, nucleic acids/DNA, cellular structures/cells, or biomimetic materials[18]. A *whole-cell* biosensor can be utilized as a rapid and sensitive method to detect

and quantify such toxic species; these types of biosensors rely on the analysis of the expression of a reporter gene that is controlled by a promoter responsive to a particular toxic compound[19].

5. Thermophiles and biotechnology

Microorganisms can be roughly classified, according to the range of temperature at which they can grow, in: i) psychrophiles (below 20°C), ii) mesophiles (from 20° up to 45°C), iii) thermophiles (from 50° up 70°C) and iv) extreme thermophiles (above 70°C). From biological point of view, the study of these microbes, has an important implication in regard to ecology as well as evolution, for the peculiarity of some microbial metabolisms[20]. The impact of biotechnology on our lives is inescapable and the comparison between the natural habitats of thermophiles and the typical industrial cultivations reveals both the advantages and the limits of the applicability of these microorganisms[21].Today the use of microorganisms has become so prevalent in biotechnology applications, to render their discovery and genetic modification a goal of our economy. Among industrial microorganism, thermophiles are highly promising to achieve this purpose. The lower cooling costs, lower contamination risk, and easier downstream processing of the final product, converted this microorganisms into ideal platforms for the new generation of sustainable bioprocesses[22].
In this thesis, we focused on thermophilic and extremophilic microorganisms with the aim to use them in environmental monitoring or bioremediation.
For this purpose, we followed two strategies:

- A model organism, *T. thermophilus* HB27, was employed to deeply characterize the molecular mechanism of resistance to toxic metals in order to develop tools for environmental biomonitoring;
- New microorganisms were isolated from hot springs in order to exploit their genetic and physiological contents for future bioremediation purposes.

6. *Thermus thermophilus* as a model organism

Many research groups use *Thermus thermophilus* as model organism, because of its high growth rates, cell yields of the cultures,

the availability of genetic tools, and the constitutive expression of an impressively efficient natural competence apparatus[23].
T. thermophilus HB8 and HB27 are aerobic Gram-negative microbes, that grows at temperatures ranging from 50° to 82°C. HB27 strain was originally isolated from a natural thermal environment in Japan by Professor Tairo Oshima[24]; its optimal growth is at 74°C and at pH 7.0. Many thermophilic organisms are strictly anaerobic, as a result of an adaptation to the low solubility of oxygen at these temperatures. However, the *Thermus* gender is an exception: *T. thermophilus* HB8, for example, can grow in anaerobic conditions by complete or partial denitrification, or using heavy metals as final acceptors of the anaerobic respiration. HB27 strain, nevertheless, cannot grow in anaerobic conditions. *T. thermophilus* cells morphologically appear as thin bacilli, which tend to form filaments able to divide by binary division. As all Gram-negative bacteria, it is surrounded by an outer membrane composed primarily of phospholipids and lipopolysaccharides, and a thin layer of peptidoglycan, responsible for the structural rigidity of the cell; this last is itself surrounded by amorphous material, covalently linked to the peptidoglycan[23]. *T. thermophilus* HB27 genome is composed by a chromosome (TTC) of 1.894.877 bp and a megaplasmid (TTP) of 232.605 bp, named pTT27. GC content is medially 69.4%. Regions that show a low GC content represent ribosomal RNA clusters (Fig. 2)[25].

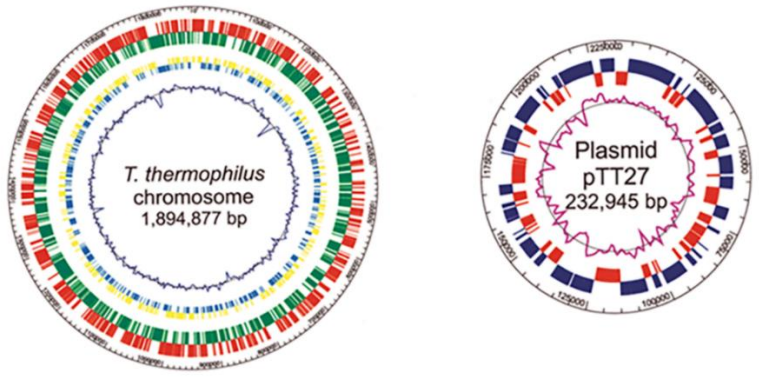


Fig. 2 Maps of *T. thermophilus* HB27 chromosome and plasmid [25].

Some thermophilic bacteria belonging to the genus *Thermus*, isolated from different geothermal vents, could use arsenate for

respiration (*Thermus sp.* HR13) and to rapidly oxidize arsenite (*T. aquaticus* and *T. thermophilus*). *T. thermophilus* HB27 showed high tolerance to concentrations of arsenate and arsenite; investigation on the molecular mechanism underlying this resistance, started from the identification, characterization and regulation of the arsenate reductase (*Tt*ArsC)[26]. Showing that *Tt*ArsC is an important component in the arsenic resistance in *T. thermophilus* and giving the first structural-functional characterization of a thermophilic arsenate reductase[26].

In the past years, our research group has developed an efficient enzyme-based biosensor, immobilizing *Tt*ArsC on gold-nanoparticles, representing a first attempt to develop a biosensor using the components of the arsenic resistance system of *Thermus thermophilus* HB27. This device provides a rapid method to evaluate the presence of arsenic and gives a proof of principle that thermophilic/thermostable metal sensing proteins are effective tools for biomonitoring[27,28].

7. CRISPR-Cas9 technology

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) and CRISPR-associated (*cas*) genes constitute a system of innate immunity of some bacteria and archaea that allow to defend themselves from the intrusion of exogenous genetic materials. A CRISPR locus is composed of two components: a CRISPR array, and an operon where the *cas*-genes are located. The CRISPR array includes 30 to 40 nt partially palindromic repetitive segments, also known as repeats, interspaced by short variable DNA sequences known as spacers, obtained from previous exposures to exogenous DNA/RNA[29,30]. The expression of these components is regulated by a leader sequence that contains regulatory elements including a promoter[30,31]. The *cas*-gene operon encodes several and highly diverse genes that are essential not only for targeting the invading DNA, but also for the crRNA processing, and for the addition of new spacers to the CRISPR array[32]. In recent years, genome editing was exploited for industrial applications, especially for improvement and development of strains with the ability of converting an unsustainable bioprocess into sustainable one such as those involving bioethanol and bioplastic[33,34].

ThermoCas9 consists in a class 2 type-IIC, RNA-guided DNA-endonuclease, which was discovered in *Geobacillus thermodenitrificans* T12 by Daas et al.[35]. It is shown to be active, *in*

vitro, between 20 and 70 °C, and up to 55 °C *in vivo*, becoming a powerful tool for genome editing both in mesophilic and thermophilic organisms[36]. Its PAM (Protospacer Adjacent Motif) sequence is composed by 8 nt, where the last 4 positions are crucial for protospacer recognition (optimal: 5′-NNNN**CCAA**-3′)[36]. Also, protospacers ranging 19 to 23 nt, cleave targets with the highest efficiency[36]. The DBS takes place 3 nucleotides upstream the PAM sequence[36]. ThermoCas9 shows several advantageous characteristics when comparing to the most widely employed Cas9 for genome editing, SpCas9. Firstly, ThermoCas9 is active *in vitro* between 20 and 70 °C and up to 55 °C *in vivo*[36], while SpCas9 is not active, *in vivo,* at or above 42 °C, meaning that can only be applied in mesophilic or facultative thermophilic organisms[37]. Secondly, the size of ThermoCas9 (1082 amino acids) is smaller than its orthologue SpCas9 (1368 amino acids)[37]. ThermoCas9 has stringent PAM-preference at its lower temperature range, providing potentially less off-targets. Also, *in vitro*, showed to tolerate fewer spacer-protospacer mismatches than SpCas9, contributing to a higher specificity[36]. In this thesis, we investigated what are the limits of ThermoCas9 activity *in vivo*, adapting the system to use it in *T. thermophilus* HB27, hence developing a new powerful tool for genome editing.

8. Aim of the work

The aim of this thesis has been focused on the identification of the molecular mechanisms that drive bacterial resistance to heavy metals for the construction of devices for environmental monitoring of arsenic or other heavy metals.

The **chapter II** is an overview of main heavy metal microbial resistance systems, focused on thermophilic microorganisms in view of potential biotechnological applications.

Chapter III describes a thorough characterization of *Tt*SmtB, an ArsR/SmtB family member that is the playmaker of the arsenic/cadmium resistance system in *T. thermophilus* HB27. In this chapter, Paper-I reports the first structural-functional characterization of the protein, identifies the specific sequences recognized by the protein on target DNAs and the realization of mutant strains of *T thermophilus* (Δsmtb::kat and TTC0354::pk18) that allowed us to study the function of these two proteins *in vivo*. The Paper-II focuses on the chemical-physical characterization of the protein to shed light on the molecular mechanisms of the

interaction with the effectors, arsenic and cadmium. Knowledge at molecular level of protein-metal interaction is important to design metal binding domains as scaffolds in metal biorecovery or biosensing in the environment.
Chapter IV describes the characterization of *Tt*ArsX, a P_{1B}-type membrane ATPase responsible for the efflux of heavy metals. Through this characterization, it was possible to obtain a *whole-cell* biosensor, based on a plasmid containing a thermo-adapted version of β-galactosidase as a reporter gene, under the transcriptional control of *Tt*ArsX responsive promoter. Measurements of β-galactosidase activities are linked to the detection of low quantities of arsenate, arsenite and cadmium.
The identification of a new component of the arsenic resistance in *T. thermophilus* HB27 is the main objective of the **chapter V**, which reports the application of pull-down and proteomic techniques for the identification of new molecular partners involved in heavy metal resistance. in this chapter it is also described the development of a genome-editing technique based on CRISPR-Cas9 technology to knock out the gene coding for the arsenite methyltransferase and to realize a plasmide-free *whole-cell* biosensor.
In the **chapter VI** we describe the identification and characterization of a new arsenic resistant strain of *Geobacillus stearothermophilus*, isolated from a hot spring located in a solfataric area of Pozzuoli (Napoli).

References:

1. Emsley, J. Nature's building blocks: an AZ guide to the elements. Oxford University Press, (2011).
2. Kim, H., Geupil J., and Youngdae Y. "Specific heavy metal/metalloid sensors: current state and perspectives." Applied Microbiology and Biotechnology (2019): 1-8.
3. Neubauer, O. "Arsenical cancer: a review." British journal of cancer 1.2 (1947): 192.
4. Cullen W. R. and Reimer K. J. Chemical Reviews 89 (4), (1989): 713-764
5. Rosen, B. P. "Biochemistry of arsenic detoxification." FEBS letters 529.1 (2002): 86-92.

6. Smedley, P. L., and Kinniburgh D. G. "A review of the source, behaviour and distribution of arsenic in natural waters." Applied geochemistry 17.5 (2002): 517-568.
7. Planer-Friedrich, B., London, J., McCleskey, R. B., Nordstrom, D. K., and Wallschläger, D. "Thioarsenates in geothermal waters of Yellowstone National Park: determination, preservation, and geochemical importance." Environmental science & technology 41.15 (2007): 5245-5251.
8. Díaz, E. "Bacterial degradation of aromatic pollutants: a paradigm of metabolic versatility." (2004).
9. Pennella, M. A., and Giedroc D. P. "Structural determinants of metal selectivity in prokaryotic metal-responsive transcriptional regulators." Biometals 18.4 (2005): 413-428.
10. Gallo, G., Puopolo, R., Limauro, D., Bartolucci, S. and Fiorentino, G. "Metal-tolerant thermophiles: from the analysis of resistance mechanisms to their biotechnological exploitation." The Open Biochemistry Journal 12.1 (2018): 149-160.
11. Carlin, A., Shi, W., Dey, S., and Rosen, B. P. "The ars operon of Escherichia coli confers arsenical and antimonial resistance." Journal of bacteriology 177.4 (1995): 981-986.
12. Silver, S. "Genes for all metals—a bacterial view of the periodic table." Journal of Industrial Microbiology and Biotechnology 20.1 (1998): 1-12.
13. Chen, C. M., Misra, T. K., Silver, S., and Rosen, B. P."Nucleotide sequence of the structural genes for an anion pump. The plasmid-encoded arsenical resistance operon." Journal of Biological Chemistry 261.32 (1986): 15030-15038.
14. Martin, P., DeMel, S., Shi, J., Gladysheva, T., Gatti, D. L., Rosen, B. P., and Edwards, B. F. "Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme." Structure 9.11 (2001): 1071-1081.
15. Diesel, E., Schreiber, M., and van der Meer, J. R. "Development of bacteria-based bioassays for arsenic detection in natural waters." Analytical and bioanalytical chemistry 394.3 (2009): 687-693.
16. Turner, A., Karube, I., and Wilson, G. S. Biosensors: fundamentals and applications. Oxford university press. (1987).
17. Juzgado, A., Soldà, A., Ostric, A., Criado, A., Valenti, G., Rapino, S., Conti, G., Fracasso, G., Paolucci, F. and Prato, M. “Highly sensitive electrochemiluminescence detection of a prostate cancer biomarker”. J. Mater. Chem. B, 5 (2017): 6681-6687.

18. Vo-Dinh, T., and Cullum, B. "Biosensors and biochips: advances in biological and medical diagnostics." Fresenius' journal of analytical chemistry 366.6-7 (2000): 540-551.
19. Fiorentino, G., Ronca, R., and Bartolucci, S. "A novel E. coli biosensor for detecting aromatic aldehydes based on a responsive inducible archaeal promoter fused to the green fluorescent protein." Applied microbiology and biotechnology 82.1 (2009): 67-77.
20. Oren, A. "Molecular ecology of extremely halophilic Archaea and Bacteria." FEMS Microbiology Ecology 39.1 (2002): 1-7.
21. Strazzulli, A., Iacono, R., Giglio, R., Moracci, M., and Cobucci-Ponzano, B. "Metagenomics of hyperthermophilic environments: biodiversity and biotechnology." Microbial Ecology of Extreme Environments. Springer, Cham, 2017. 103-135.
22. F Bosma, E., van der Oost, J., M de Vos, W., and van Kranenburg, R. "Sustainable production of bio-based chemicals by extremophiles." Current Biotechnology 2.4 (2013): 360-379.
23. Cava, F., Hidalgo, A., and Berenguer, J. "*Thermus thermophilus* as biological model." Extremophiles 13.2 (2009): 213.
24. Oshima, T., and Imahori, K. "Description of Thermus thermophilus (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa." International Journal of Systematic and Evolutionary Microbiology 24.1 (1974): 102-112.
25. Henne, A., Brüggemann, H., Raasch, C. *et al*. "The genome sequence of the extreme thermophile *Thermus thermophilus*." Nature biotechnology 22.5 (2004): 547-553.
26. Del Giudice, I., Limauro, D., Pedone, E., Bartolucci, S., and Fiorentino, G. "A novel arsenate reductase from the bacterium *Thermus thermophilus* HB27: its role in arsenic detoxification." Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1834.10 (2013): 2071-2079.
27. Politi, J., Spadavecchia, J., Fiorentino, G., Antonucci, I., Casale, S., and De Stefano, L. "Interaction of *Thermus thermophilus* ArsC enzyme and gold nanoparticles naked-eye assays speciation between As (III) and As (V)." Nanotechnology 26.43 (2015): 435703.
28. Politi, J., Spadavecchia, J., Fiorentino, G., Antonucci, I., and De Stefano, L. "Arsenate reductase from *Thermus thermophilus* conjugated to polyethylene glycol-stabilized gold nanospheres

allow trace sensing and speciation of arsenic ions." Journal of The Royal Society Interface 13.123 (2016): 20160629
29. Bolotin, A., Quinquis, B., Sorokin, A., & Ehrlich, S. D. "Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin." Microbiology 151.8 (2005): 2551-2561.
30. Rath, D., Amlinger, L., Rath, A., & Lundgren, M. "The CRISPR-Cas immune system: biology, mechanisms and applications." Biochimie 117 (2015): 119-128.
31. Alkhnbashi, O. S., Shah, S. A., Garrett, R. A., Saunders, S. J., Costa, F., & Backofen, R. "Characterizing leader sequences of CRISPR loci." Bioinformatics 32.17 (2016): i576-i585.
32. Jansen, R., Embden, J. D. V., Gaastra, W., & Schouls, L. M. "Identification of genes that are associated with DNA repeats in prokaryotes." Molecular microbiology 43.6 (2002): 1565-1575.
33. Jagadevan, S., Banerjee, A., Banerjee, C., Guria, C., Tiwari, R., Baweja, M., & Shukla, P. "Recent developments in synthetic biology and metabolic engineering in microalgae towards biofuel production." Biotechnology for biofuels 11.1 (2018): 185.
34. Javed, M. R., Noman, M., Shahid, M., Ahmed, T., Khurshid, M., Rashid, M. H., & Khan, F. "Current situation of biofuel production and its enhancement by CRISPR/Cas9-mediated genome engineering of microbial cells." Microbiological research 219 (2019): 1-11.
35. Daas, M. J., Vriesendorp, B., van de Weijer, A. H., van der Oost, J., & van Kranenburg, R. "Complete genome sequence of *Geobacillus thermodenitrificans* T12, a potential host for biotechnological applications." Current microbiology 75.1 (2018): 49-56.
36. Mougiakos, I., Mohanraju, P., Bosma, E.F. et al. "Characterizing a thermostable Cas9 for bacterial genome editing and silencing." Nature communications 8.1 (2017): 1-11.
37. Mougiakos, I., Bosma, E. F., Weenink, K., Vossen, E., Goijvaerts, K., van der Oost, J., & van Kranenburg, R. "Efficient genome editing of a facultative thermophile using mesophilic spCas9." ACS synthetic biology 6.5 (2017): 849-861.

Chapter II

Metal-Tolerant Thermophiles: From the Analysis of Resistance Mechanisms to their Biotechnological Exploitation.

The extreme environments on our planet are excellent places to find microorganisms capable of tolerating conditions of temperature, pH, salinity, pressure and high elevated concentrations of heavy metals. The study of the genetic and physiology of these microorganisms has highlighted the presence of multiple mechanisms for facing different chemical and physical stresses.

In the last decades, the study of extremophilic microorganisms has had a massive increase, also due to the boost of these microorganisms in biotechnology.

In the context of this PhD thesis, the study of heavy metal resistance systems in thermophilic microorganisms represents the starting point for the development of new systems for environmental monitoring.

This review is focuses on the molecular mechanisms responsible for survival and adaptation of thermophiles to toxic metals, with particular emphasis on arsenate As(V), arsenite As(III), cadmium Cd(II), and on current biotechnologies for their detection, extraction and removal.

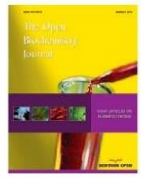
BENTHAM OPEN

CrossMark

The Open Biochemistry Journal

Content list available at: www.benthamopen.com/TOBIOCJ/

DOI: 10.2174/1874091X01812010149, 2018, *12*, 149-160



REVIEW ARTICLE

Metal-Tolerant Thermophiles: From the Analysis of Resistance Mechanisms to their Biotechnological Exploitation

Giovanni Gallo, Rosanna Puopolo, Danila Limauro, Simonetta Bartolucci and Gabriella Fiorentino*

Department of Biology, University of Naples Federico II, Edificio 7, via Cinthia n. 6, 80126 Naples, Italy

Received: August 1, 2018 Revised: October 19, 2018 Accepted: October 22, 2018

Abstract:

Extreme terrestrial and marine hot environments are excellent niches for specialized microorganisms belonging to the domains of Bacteria and Archaea; these microorganisms are considered extreme from an anthropocentric point of view because they are able to populate harsh habitats tolerating a variety of conditions, such as extreme temperature and/or pH, high metal concentration and/or salt; moreover, like all the microorganisms, they are also able to respond to sudden changes in the environmental conditions. Therefore, it is not surprising that they possess an extraordinary variety of dynamic and versatile mechanisms for facing different chemical and physical stresses. Such features have attracted scientists also considering an applicative point of view. In this review we will focus on the molecular mechanisms responsible for survival and adaptation of thermophiles to toxic metals, with particular emphasis on As(V), As(III), Cd(II), and on current biotechnologies for their detection, extraction and removal.

Keywords: Toxic metals, Resistance systems, Metal bioremediation, Thermostable biosensors, Environmental conditions, Physical stresses.

1. HEAVY METALS: TOXICITY AND TRANSFORMATION

Heavy metals are among the most persistent and toxic pollutants in the environment [1]. Even in small concentrations, they can threat human health as well as the environment because they are non-biodegradable. There is no widely agreed criterion for definition of a heavy metal. Depending on the context, this term can acquire different meanings: for example, in metallurgy a heavy metal may be defined by its density, in physics by its atomic number, and in chemistry by its chemical behavior [2]. The *International Union of Pure and Applied Chemistry* (IUPAC) does not consider these definitions correct; in this review, according to IUPAC, every "heavy" metal has the following characteristics: density exceeding 5.0 g/cm^3; general behavior as cations; low solubility of their hydrates; aptitude to form complexes and affinity towards the sulfides.

In 2010, the World Health Organization (WHO) estimated that more than 25% of total diseases were linked to environmental factors including exposure to toxic chemicals [3]. For example, lead, [Pb(II)], one of the most common heavy metals, is thought to be responsible for 3% of cerebrovascular disease worldwide [4]; while cadmium, [Cd(II)], has been classified as carcinogen by the *International Agency for Research on Cancer* (IARC) on the basis of several evidence in both humans and experimental animals [5, 6]. Furthermore, hazards associated with exposure to other metal ions like chromium [Cr(II)], mercury [Hg(I)], and arsenic [As(III) and As(V)], have been well established in the literature [7 - 12]. The risk related to heavy metal exposure depends on the concentration and time [13].

Table **1** reports the concentration limits of the most common heavy metals in drinkable water, suggested by the *World Health Organization* (WHO) and the *Environmental Protection Agency* (EPA). It also shows the possible sources of these contaminants in drinking water and the potential health effects from long-term exposure [12, 14].

* Address correspondence to this author at the Department of Biology, University of Naples Federico II, Edificio 7, *via* Cinthia n. 6, 80126 Naples, Italy; Tel: +39081679167; E-mail: fiogabri@unina.it

1874-091X/18

Table 1. Limit concentrations of common heavy metals in drinkable water, suggested by the *World Health Organization* (WHO) and the *Environmental Protection Agency* (EPA).

Contaminant	WHO (mg/L)	EPA (mg/L)	Potential Health Effects from Long-Term Exposure	Sources of Contaminant in Drinking Water
As	0.010	0.010	Skin damage or problems with circulatory systems, and may have increased risk of getting cancer	Erosion of natural deposits; runoff from orchards, runoff from glass and electronics production wastes
Ba	7	2	Increase in blood pressure	Discharge of drilling wastes; discharge from metal refineries; erosion of natural deposits
Cd	0.003	0.005	Kidney damage	Corrosion of galvanized pipes; erosion of natural deposits; discharge from metal refineries; runoff from waste batteries and paints
Cr (total)	0.050	0.100	Allergic dermatitis	Discharge from steel and pulp mills; erosion of natural deposits
Cu	2	1.300	Liver or kidney damage	Corrosion of household plumbing systems; erosion of natural deposits
Hg (inorganic)	0.006	0.002	Kidney damage	Erosion of natural deposits; discharge from refineries and factories; runoff from landfills and croplands
Pb	0.010	0.015	Kidney problems; high blood pressure	Corrosion of household plumbing systems
Sb	0.020	0.006	Increase in blood cholesterol; decrease in blood sugar	Discharge from petroleum refineries; fire retardants; ceramics; electronics; solder

On the other hand, heavy metals naturally occur in the Earth's crust. They are present in soils, rocks, sediments, air and waters and can be used and modified by local microbial communities, which are actively involved in metal geochemical cycles, affecting their speciation and mobility. Many metals are essential for life because they are actively involved in almost all aspects of metabolism: as examples, iron and copper are involved in the electron transport, manganese and zinc influence enzymatic regulations. However, their excess can disrupt natural biochemical processes and cause toxicity. For these reasons, all the microorganisms have evolved resistance systems to get rid of the cell of toxic metals as well as molecular mechanisms to maintain metal homeostasis. These systems frequently rely on a balance between uptake and efflux processes [15]. Because of microbial adaptation, microorganisms can also contribute to increase toxicity levels [16, 17]. For example, several studies in Bangladesh have demonstrated that microbial processes enhance the arsenic contamination in near- and sub-surface aquifers, because arsenate-respiring bacteria can liberate As(III) from sediments, adsorptive sites of aluminum oxides or ferrihydrite, or from minerals, such as scorodite [18].

Metal biotransformation impacts human health through the food chain: examples include the oxidation of Hg(0) to Hg(II), and the subsequent methylation to methylmercury compounds, which can be accumulated by fish and marine mammals in the aquatic environment [19].

Despite their relevant toxicity, in a report of the European Commission (named "Critical-Metals in the Path towards the Decarbonisation of the EU Energy Sector"), several heavy metals such as cadmium, chromium and lead are included into the classification of critical raw materials. According to the sustainable low-carbon economic policy of EU, these metals are expected to become a bottleneck in a near future to the supply- chain of various low- carbon energy technologies [20]. Therefore, it is very important to detect and recover these heavy metals to achieve both environmental safeguard and sustainable economic strategies. Common sources of heavy metals in this context include mining and industrial wastes, vehicle emissions, lead-acid batteries, fertilizers, paints, treated timber, aging water supply infrastructures, and microplastics floating in the world's oceans [21, 22].

2. METAL RESISTANCE MECHANISMS

Microorganisms able to tolerate high levels of heavy metal ions have evolved in ore deposits, hydrothermal vents, geothermal sites, as well as in different polluted sites [23]. Metal tolerance of thermophilic Bacteria/Archaea is due to several mechanisms, many also found in mesophilic counterparts, such as: extracellular barrier, metal ion transport into and outside the cell, the utilization of toxic metal ions in metabolism or the presence of metal resistance genes with different genomic localization (chromosome, plasmid or transposon) [24].

To date, at least four main mechanisms of heavy metal resistance, schematized in Fig. (**1**), are described which can be even found in the same microorganism [24 - 31]:

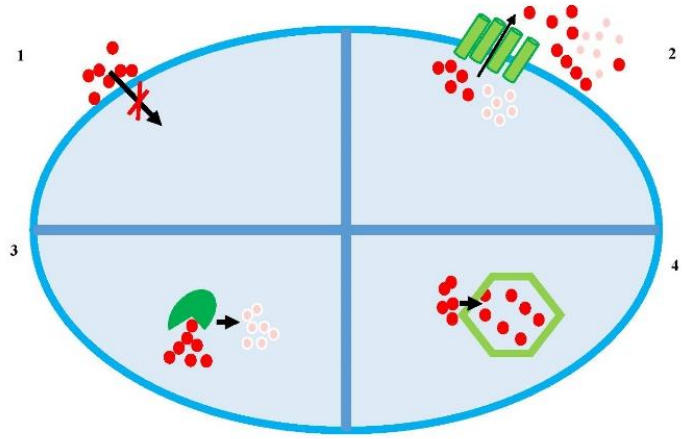


Fig. (1). Generalized illustration of the genetic mechanism of resistance to toxic metals by microorganisms: **1)** Extracellular barrier, a selectively permeable system; **2)** Efflux of metal ions; **3)** Enzymatic reduction of metal ions. **4)** Intracellular sequestration by small molecule complexing agents or metal-chelating proteins.

- Extracellular barrier;
- Active transport of metal ions (efflux);
- Enzymatic reduction of metal ions;
- Intracellular sequestration.

The cell wall or plasma membrane can prevent metal ions from entering the cell. Bacteria belonging to different taxonomical groups can adsorb metal ions by ionisable groups of the cell wall (carboxyl, amino, phosphate and hydroxyl groups) [32]. However, many metal ions enter the cell *via* the systems responsible for the uptake of essential elements: for example, Cr(II) is transported inside the cell *via* sulphate transport system [33], whereas Cd(II), Zn(II), Co(II), Ni(II) and Mn(II), enter the cells using systems of magnesium transport [34]. Moreover, As(V) is taken into cells by phosphate transport systems and As(III) has been shown to be taken up by glucose permeases [35].

Both in Gram-positive and Gram-negative bacteria the arsenic resistance system is composed by operons of 3-5 genes carried on plasmids or chromosomes; the two most common contain either five genes (*arsRDΔBC*), as in the plasmid R773 of the *Escherichia coli*, or three genes (*arsRBC*), as in the plasmid pI258 of *Staphylococcus aureus* [36, 37]. The *arsR* gene encodes a trans-acting repressor of the ArsR/SmtB family involved in transcriptional regulation [30], *arsB* encodes an As(III) efflux transporter [38], and *arsC* encodes a cytoplasmic arsenate reductase that converts As(V) to As(III), which is extruded outside the cell [39]. Where present, ArsD is a metallochaperone that transfers trivalent metalloids to ArsA, the arsenite-stimulated efflux ATPase [31].

2.1. Transport of Metal Ions: Efflux Systems

The majority of thermophilic microorganisms (both belonging to Archaea and Bacteria domains) are resistant to heavy metals thanks to an active transport and/or efflux of metal ions outside the cells. The genetic determinants of efflux systems can be localized on chromosomes [40] and on plasmids [41]. In most cases, the expression of metal efflux genes is transcriptionally controlled by co-transcribed metal sensor proteins [42].

In microorganisms, efflux systems consist of proteins belonging to three families: CDF (cation diffusion facilitator), P-type ATPases and RND (resistance, nodulation, cell division) [43]. CDF proteins and P-type ATPases of Gram-negative bacteria transport specific substrates through the plasma membrane into the periplasm. CDF proteins are metal transporters occurring in all the three domains of life whose primary substrates are mainly ions of divalent metals like Zn(II), Co(II), Ni(II), Cd(II), and Fe(II) and export metals through a chemiosmotic gradient formed by H^+ or K^+ [44, 45]. Most of these proteins have six transmembrane helices containing a zinc-binding site within the transmembrane domains, and a binuclear zinc-sensing and binding site in the cytoplasmic C-terminal region [46]. These proteins exhibit an unusual degree of sequence divergence and size variation (300-750 residues).

Differently from CDF-proteins, P-type ATPases transfer both monovalent or bivalent metal ions with high affinity for sulfhydryl groups (Cu(I)/Ag(I), Zn(II)/Cd(II)/Pb(II)) and use ATP hydrolysis to transport ions across cellular membranes [43]. They are composed of three conserved domains: i) a transmembrane helix bundle, allowing substrate translocation; 2) a soluble ATP binding domain containing a transiently phosphorylated aspartate residue; 3) a soluble actuator domain (AD). Those belonging to P_{1B}-type are capable to drive the efflux out of cells of both essential transition metal ions (*e.g.*, Zn(II), Cu(I), and Co(II)) and toxic metals (*e.g.*, Ag(I), Cd(II), Pb(II)) contributing to their homeostasis maintenance. In a recent study on a huge number of P_{1B}-type ATPase they have been classified into seven distinct subfamilies (1_{B-1} 1_{B-7}) but the molecular basis of metal ion specificity remains unclear [47]. Several thermophilic P_{1B}-type ATPases involved in metal efflux have been characterized. The thermophilic bacterium *Thermus thermophilus* HB27 contains in its genome three genes coding for putative P_{1B}-type ATPases: *TTC1358*, *TTC1371*, and *TTC0354*; these genes are annotated, respectively, as two putative copper transporter (CopA and CopB) and a zinc-cadmium transporter (Zn(II)/Cd(II)-ATPase) [48], involved in heavy metal resistance. *Archaeoglobus fulgidus* possesses a CopA protein driving the outward movement of Cu(I) or Ag(I) characterised by a conserved CPC metal binding site and a cytoplasmic metal binding sequence (also containing cysteine residues) at its N- and C- terminus [49].

Members of the RND family are efflux pumps, especially identified in Gram-negative bacteria that can be divided in subfamilies depending on the substrate transported; they actively export heavy metals, hydrophobic compounds, nodulation factors [50]. The heavy metal efflux (HME) RND sub-family functions for metal ion efflux powered by a proton-substrate antiport. The prototype family member from *E. coli* is CusA; it works in conjunction with the membrane fusion protein CusB and the outer-membrane channel CusC forming a tripartite complex spanning the entire cell envelope to export Cu(I) and Ag(I) [51]. The crystal structures of CusA in the absence and presence of bound Cu(I) or Ag(I) has been recently solved providing structural information [52].

To the best of our knowledge, thermophilic microbial genomes do not contain genes encoding proteins of the HME-RND family. A Blast analysis of CusA against thermophilic genomes revealed homology with integral membrane proteins of the ACR (Activity Regulated Cytoskeleton Associated Protein) family involved in drug and/or heavy metal resistance [53].

2.2. Enzymatic Reduction of Metal Ions: Metal Reductases

Many thermophilic microorganisms employ intracellular enzymatic conversions combined with efflux systems to obtain heavy metal resistance. Enzymatic reduction of metal ions can result in the formation of less toxic forms like Hg(II) reduced to Hg(0), Cr(V) converted into Cr(III) [23] or, as in the case of As(V), reduction in the more toxic As(III) which is the only form extruded by the cell.

Several thermophilic metal reductases have been described so far; for example, *Tt*ArsC from *Thermus thermophilus* HB27 is an arsenate reductase which enzymatically converts As(V) in As(III) [54] using electrons provided from the thioredoxin-thioredoxin reductase system and employing a catalytic mechanism in which the thiol group of a N-terminal cystein performs a nucleophilic attack on the arsenate [54]. As told before the arsenite is then extruded by a dedicated efflux protein.

MerA from *Sulfolobus solfataricus* is flavoprotein that catalyzes the reduction of Hg(II) to volatile Hg(0), converting toxic mercury ions into relatively inert elemental mercury [55].

The thermophilic bacteria isolated from various ecological niches can also reduce a broad spectrum of other heavy metal ions such as Cr(V), Mo(VI) and V(V) [56] that serve as terminal acceptors of electrons during their anaerobic respiration [57].

These systems are generally finely regulated by specific transcription factors. As an example, the transcription of the arsenic resistance system of *T. thermophilus* HB27 is regulated by *Tt*SmtB, a protein belonging to the ArsR/SmtB family which acts as the As(V) and As(III) intracellular sensor [30]. In the absence of metal ions, the protein binds to regulatory regions upstream of *TTC1502,* encoding *Tt*ArsC, and *TTC0354*, encoding the efflux membrane protein *Tt*ArsX, (a P_{1B}- type ATPase, see above). In a recent study from our group it was demonstrated that *Tt*ArsX and *Tt*SmtB are also responsible for Cd(II) tolerance [58].

2.3. Metal Intracellular Sequestration

Another common mechanism to inactivate toxic metal ions is the intracellular sequestration or the complexation of metal ions by various compounds in the cytoplasm. The metallothioneins and phytochelatins are two classes of peptides

rich in cysteine residues which bind metal ions through the sulfhydrylic groups [59].

Metallothioneins constitute a superfamily of ubiquitous cytosolic small (25–82 amino acids), cysteine-rich (7–21 conserved Cys residues) proteins able to bind metal ions, mainly Cd(II), Zn(II) and Cu(I), *via* metal-thiolate clusters in the absence of aromatic amino acids and histidine residues [60]. They are multifunctional proteins whose synthesis is stimulated by heavy metals and other environmental stressors; based on this latter evidence the metallothionein promoter of *Tetrahymena thermophila* has been employed in the development of a *whole-cell* biosensor for the detection of heavy metals [61].

Among prokaryotes, the ability to synthesize metallothionein has been demonstrated in the cyanobacterium *Synechococcus* sp. PCC 7942 which contains two genes *smtA* and *smtB* inducible by Cd(II) and Zn(II). The peptide contained fewer cysteine residues than the eukaryotic metallothionein [62].

3. APPLICATIONS IN BIOTECHNOLOGY

A detailed understanding of the molecular mechanisms responsible for resistance to toxic metals in metal-tolerant microorganisms is also crucial for a potential use in the environmental monitoring of metal contamination and to set up bioremediation processes, the most promising being biosorption and removal as insoluble complexes Fig. (**2**).

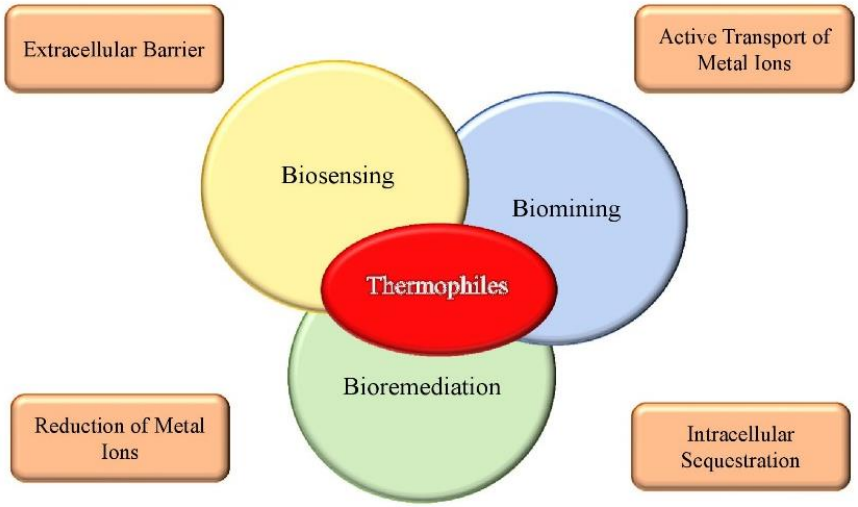


Fig. (2). Schematic representation of potential applications in biotechnology of metal-tolerant thermophiles.

The traditional approach for monitoring the environmental pollution is based on chemical or physical analysis and allows highly accurate and sensitive determination of the exact composition of any sample. These analyses require specialized and expensive instrumentations, as the ICP-MS (*Inductively Coupled Plasma - Mass Spectrometry*) which, to date, is the most adopted technique for detecting heavy metals [63].

The need for accurate, not expensive, on-site and real-time measurements has led to the development of sensors based on biomolecules and nanomaterials [63 - 66]. Biosensors are analytical devices which integrate a biological recognition element with a physical transducer to generate a measurable signal proportional to the concentration of the analyte [67 - 69]. The use of biological molecules is a considerable advantage in the sensor field, because of their high specificity: these sensors are based on the specific interaction between enzymes and their substrates, antibodies and antigens, target molecules to their receptors, or the high affinity of nucleic acid strands to their complementary sequences [67]. Nevertheless, some biomolecules can be too labile for the exploitation on the marketplace. In this context, the biomolecules of the thermophiles are more stable at high temperatures than the mesophilic counterparts. In fact they have already been exploited in biotechnology, as demonstrated by the development of the polymerase chain reaction (PCR) and the use of thermozymes in many industrial applications [70 - 73].

The recent developments in nanotechnology have also opened new horizons for biosensing: nanomaterials are attractive because of their unique electrical, chemical and physical properties (*i.e.* size, composition, conductivity, magnetism, mechanical strength, and light-absorbing and emitting). The most studied of them, carbon nanotubes (CNTs), graphene, metal nanoparticles (MNPs), and quantum dots (QDs), have been especially targeted for developing novel biosensors [64, 73 - 76]. An example of a successful match between thermophilic biomolecules and nanomaterial for the development of a heavy metal biosensor is reported by Politi *et al.* [77, 78]. In this work, *Tt*ArsC, the arsenate reductase from *T. thermophilus* HB27 [54] was conjugated to polyethylene glycol-stabilized gold nanospheres. The new nanobiosensor revealed high sensitivities and limits of detection equal to 10 ± 3 M^{-12} and 7.7 ± 0.3 M^{-12} for As(III) and As(V), respectively [77, 78].

A detailed understanding of the molecular mechanisms responsible for resistance to toxic metals is crucial to develop whole cell biosensors for the detection of chemicals in the environment so that not only thermophilic biomolecules, but also thermophilic microorganisms can constitute heavy metal biosensors. Many whole-cell biosensors for metal ions detection have been already described in the literature based on the realization of reporter systems containing regulatory cis-acting sequences interacting with the cognate transcriptional metal sensor repressor; however, to date, there are few reports on thermophilic whole cell biosensors [79, 80].

In the work by Poli *et al.* [57], *Anoxybacillus amylolyticus*, an acidothermophilic bacterium isolated from geothermal soil samples in Antarctica, was observed to be resistant to metals like Ni(II), Zn(II), Co(II), Hg(II), Mn(II), Cr(VI), Cu(II) and Fe(III). A decrease in α-amylase activity, correlated with a decrease in α-amylase production, was observed in the presence of heavy metals: it is speculated to use it as a toxicological indicator of heavy metals in a potential microbial bioassay employing whole cells [57].

Biomining and bioremediation represent a new branch of biotechnology, named biometallurgy, addressed at heavy metal recovery including the processes that involve interactions between microorganisms and metals or metal-bearing minerals [81]. Biomining refers to the exploitation of microorganisms to extract and recover metals from ores and waste concentrates (the term is often used synonymously with *bioleaching* when the metals are solubilized during the process); on the other hand, bioremediation focuses on the transformation of a toxic substance into a harmless or less toxic one from contaminated sites [82, 83]. In this contest, among the heavy metals resistance systems developed by microorganisms, and in particular by thermophiles, biosorption and bioaccumulation are emerging as promising low-cost methodologies for bioremediation [84, 85]. Biosorption and bioaccumulation are two processes that consist into the ability of microorganisms to accumulate heavy metals from wastewater through metabolic pathways or physical-chemical uptake; but while biosorption is a passive process depending on the composition of the cellular surface and following a kinetic equilibrium, bioaccumulation is an energy driven process and requires an active metabolism [59].

Several examples of biosorption and bioaccumulation are provided by microorganisms belonging to *Geobacillus sp.*, which are highly tolerant to Cd(II), Cu(II) and Zn(II) [86 - 88]. *Özdemir, S. et al,* indicated a *G. toebii subsp. decanicus* as an efficient viable biosorbent for heavy metals. This study clearly shows that thermophiles can be used for removal and recovery of heavy metals from industrial wastewater [89].

The interest into the potential applications of heavy metal resistant thermophiles has led to the development of tools and assays for screening them on lab scale [61, 90 - 92], with the final goal to design and set up bacterial bioprocesses on the industrial scale [93 - 95]. Table **2** summarizes some examples of bioprocesses employing heavy metals performed by thermophiles.

Table 2. Examples of heavy metals bioprocesses performed by thermophiles.

Heavy Metals	Microrganisms	Bio Processes	References
$CuFeS_2$	*Leptospirillum ferriphilum* *Acidithiobacillus caldus*	Bioleaching	*Zhou H.B. et al, 2009* [94]
Fe(II)	*Sulfobacillus sp.*	Bioleaching	*Hawkes R. et al, 2006* [92]
As(III), As(V)	*Thermus thermophilus*	Biosensing	*Politi J. et al, 2015* [77]
Ni(II), Zn(II), Co(II), Hg(II), Mn(II), Cr(VI), Cu(II) and Fe(III)	*Anoxybacillus amylolyticus*	Biosensing	*Poli et al, 2008* [57]
Cd(II)	*Tetrahymena thermophila*	Biosensing	*Amaro F. et al, 2011* [61]
Cd(II), Cu(II), Ni(II), Mn(II), Zn(II)	*Geobacillus toebii subsp. decanicus* *Geobacillus thermoleovorans subsp. stromboliensis*	Biosorption	*Özdemir, S. et al, 2012* [86]

(Table 2) contd.....

Heavy Metals	Microrganisms	Bio Processes	References
Cd(II)	*Geobacillus stearothermophilus* *Geobacillus thermocatenulatus*	Biosorption	*Hetzer, A. et al, 2006* [87]
Fe(III), Cr(III), Cd(II), Pb(II), Cu(II), Co(II), Zn(II), Ag(I)	*Geobacillus thermodenitrificans*	Biosorption	*Chattereji S.K. et al, 2010* [88]

4. RESULT

Metal-tolerant thermophiles exhibit metabolic and physiological features that distinguish them from other major life groups due to their adaptation to extreme environments.

CONCLUSION

Most of the knowledge regarding mechanisms of adaptation/resistance to toxic metals has been discovered using traditional microbiological/biochemical techniques and thanks to the use of genomics, transcriptomics, proteomics as well as to the recent development of genetic tools for many of these organisms. With the advent of next-generation sequencing technologies, comparative genomics and metagenomics projects, it appears that even novel metabolic features can be discovered, further expanding our understanding of environmental microbiology. Such an integrated view opens to new opportunities for biotechnological applications in commercially relevant processes such as the monitoring of metal concentrations in the environment, the recovery of precious and strategic metals and the setup of microbial-based remediation strategies.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

This work was supported by Università degli Studi di Napoli Federico II that financed the original research project: "Immobilization of Enzymes on hydrophobin-functionalized Nanomaterials".

REFERENCES

[1] Mathew, B.B.; Singh, H.; Biju, V.G.; Krishnamurthy, N.B. Classification, source and effect of environmental pollutants and their biodegradation. *J. Environ. Pathol. Toxicol. Oncol.*, **2017**, *36*(1), 55-71. [http://dx.doi.org/10.1615/JEnvironPatholToxicolOncol.2017015804] [PMID: 28605331]

[2] Nieboer, E.; Richardson, D.H. The replacement of the nondescript term 'heavy metals' by a biologically and chemically significant classification of metal ions. *Environ. Pollut. B*, **1980**, *1*(1), 3-26. [http://dx.doi.org/10.1016/0143-148X(80)90017-8]

[3] Organization, W.H. *WHO human health risk assessment toolkit: Chemical hazards.*, **2010**.

[4] Mamtani, R.; Stern, P.; Dawood, I.; Cheema, S. Metals and disease: A global primary health care perspective. *J. Toxicol.*, **2011**, *2011*, 319136. [http://dx.doi.org/10.1155/2011/319136] [PMID: 22007209]

[5] Järup, L. Hazards of heavy metal contamination. *Br. Med. Bull.*, **2003**, *68*, 167-182. [http://dx.doi.org/10.1093/bmb/ldg032] [PMID: 14757716]

[6] Loprieno, N. International Agency for Research on Cancer (IARC) monographs on the evaluation of carcinogenic risk of chemicals to man:"Relevance of data on mutagenicity". *Mutation Research/Environmental Mutagenesis and Related Subjects*, **1975**, *31*(3), 201.

[7] Boening, D.W. Ecological effects, transport, and fate of mercury: A general review. *Chemosphere*, **2000**, *40*(12), 1335-1351. [http://dx.doi.org/10.1016/S0045-6535(99)00283-0] [PMID: 10789973]

[8] Gomez-Caminero, A. Environmental health criteria 224 arsenic and arsenic compounds. In: *United Nations Environment Programme*; , **2001**.

[9] Brown, T.J. World mineral production 2010-2014. In: *British geological survey*; , **2014**.

[10] Butiuc-Keul, A.; Momeu, L.; Craciunas, C.; Dobrota, C.; Cuna, S.; Balas, G. Physico-chemical and biological studies on water from Aries River (Romania). *J. Environ. Manage.*, **2012**, *95*(Suppl.), S3-S8. [http://dx.doi.org/10.1016/j.jenvman.2011.04.017] [PMID: 21596474]

[11] Kotaś, J.; Stasicka, Z. Chromium occurrence in the environment and methods of its speciation. *Environ. Pollut.*, **2000**, *107*(3), 263-283. [http://dx.doi.org/10.1016/S0269-7491(99)00168-2] [PMID: 15092973]

[12] Gorchev, H.G.; Ozolins, G. WHO guidelines for drinking-water quality. *WHO Chron.*, **1984**, *38*(3), 104-108. [PMID: 6485306]

[13] Thompson, M.R.; Schwartz Barcott, D. The concept of exposure in environmental health for nursing. *J. Adv. Nurs.*, **2017**, *73*(6), 1315-1330. [http://dx.doi.org/10.1111/jan.13246] [PMID: 28000242]

[14] *EPA, National primary drinking water regulations*, **1995**, *1*

[15] Pedone, E.; Bartolucci, S.; Fiorentino, G. Sensing and adapting to environmental stress: The archaeal tactic. *Front. Biosci.*, **2004**, *9*(29097), 2909-2926. [http://dx.doi.org/10.2741/1447] [PMID: 15353325]

[16] Gadd, G.M. Metals and microorganisms: A problem of definition. *FEMS Microbiol. Lett.*, **1992**, *100*(1-3), 197-203. [http://dx.doi.org/10.1111/j.1574-6968.1992.tb05703.x] [PMID: 1478456]

[17] Riordan, J.F. The role of metals in enzyme activity. *Ann. Clin. Lab. Sci.*, **1977**, *7*(2), 119-129. [PMID: 192123]

[18] Stolz, J.F.; Basu, P.; Santini, J.M.; Oremland, R.S. Arsenic and selenium in microbial metabolism. *Annu. Rev. Microbiol.*, **2006**, *60*, 107-130. [http://dx.doi.org/10.1146/annurev.micro.60.080805.142053] [PMID: 16704340]

[19] Vierkötter, A. Environmental pollution and skin aging. *Hautarzt*, **2011**, *62*(8) 577-578, 580-581. [PMID: 21656109]

[20] *Critical Metals in the Path towards the Decarbonisation of the EU Energy Sector. JRC Scientific & Policy Report*; European Commission, **2013**, pp. 1-246.

[21] Kadirvelu, K.; Thamaraiselvi, K.; Namasivayam, C. Removal of heavy metals from industrial wastewaters by adsorption onto activated carbon prepared from an agricultural solid waste. *Bioresour. Technol.*, **2001**, *76*(1), 63-65. [http://dx.doi.org/10.1016/S0960-8524(00)00072-9] [PMID: 11315812]

[22] Cole, M.; Lindeque, P.; Halsband, C.; Galloway, T.S. Microplastics as contaminants in the marine environment: A review. *Mar. Pollut. Bull.*, **2011**, *62*(12), 2588-2597. [http://dx.doi.org/10.1016/j.marpolbul.2011.09.025] [PMID: 22001295]

[23] Ranawat, P.; Rawat, S. Metal-tolerant thermophiles: Metals as electron donors and acceptors, toxicity, tolerance and industrial applications. *Environ. Sci. Pollut. Res. Int.*, **2018**, *25*(5), 4105-4133. [http://dx.doi.org/10.1007/s11356-017-0869-2] [PMID: 29238927]

[24] Bruins, M.R.; Kapil, S.; Oehme, F.W. Microbial resistance to metals in the environment. *Ecotoxicol. Environ. Saf.*, **2000**, *45*(3), 198-207. [http://dx.doi.org/10.1006/eesa.1999.1860] [PMID: 10702338]

[25] 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*(12), 1242-1256. [http://dx.doi.org/10.1111/j.1462-2920.2003.00463.x] [PMID: 14641571]

[26] 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*(8), 909-916. [http://dx.doi.org/10.1094/PHYTO.2002.92.8.909] [PMID: 18942971]

[27] 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*(3), 335-347. [http://dx.doi.org/10.1111/j.1574-6976.2001.tb00581.x] [PMID: 11348688]

[28] Cervantes, C.; Gutierrez-Corona, F. Copper resistance mechanisms in bacteria and fungi. *FEMS Microbiol. Rev.*, **1994**, *14*(2), 121-137. [http://dx.doi.org/10.1111/j.1574-6976.1994.tb00083.x] [PMID: 8049096]

[29] 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*(20), 8915-8919. [http://dx.doi.org/10.1073/pnas.88.20.8915] [PMID: 1924351]

[30] 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**, *10*(6), 1690-1701. [http://dx.doi.org/10.1111/1751-7915.12761] [PMID: 28696001]

[31] Lin, Y.F.; Walmsley, A.R.; Rosen, B.P. An arsenic metallochaperone for an arsenic detoxification pump. *Proc. Natl. Acad. Sci. USA*, **2006**,

103(42), 15617-15622.
[http://dx.doi.org/10.1073/pnas.0603974103] [PMID: 17030823]

[32] 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*(4), 273-280.
[http://dx.doi.org/10.1023/A:1009291931258] [PMID: 11247032]

[33] Gilotra, U.; Srivastava, S. Plasmid-encoded sequestration of copper by Pseudomonas pickettii strain US321. *Curr. Microbiol.*, **1997**, *34*(6), 378-381.
[http://dx.doi.org/10.1007/s002849900199] [PMID: 9142746]

[34] Nies, D.H.; Silver, S. Metal ion uptake by a plasmid-free metal-sensitive Alcaligenes eutrophus strain. *J. Bacteriol.*, **1989**, *171*(7), 4073-4075.
[http://dx.doi.org/10.1128/jb.171.7.4073-4075.1989] [PMID: 2472382]

[35] Yang, H.C.; Rosen, B.P. New mechanisms of bacterial arsenic resistance. *Biomed. J.*, **2016**, *39*(1), 5-13.
[http://dx.doi.org/10.1016/j.bj.2015.08.003] [PMID: 27105594]

[36] Saltikov, C.W.; Olson, B.H. 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.*, **2002**, *68*(1), 280-288.
[http://dx.doi.org/10.1128/AEM.68.1.280-288.2002] [PMID: 11772637]

[37] Ji, G.; Silver, S. Regulation and expression of the arsenic resistance operon from Staphylococcus aureus plasmid pI258. *J. Bacteriol.*, **1992**, *174*(11), 3684-3694.
[http://dx.doi.org/10.1128/jb.174.11.3684-3694.1992] [PMID: 1534328]

[38] Patel, P.C.; Goulhen, F.; Boothman, C.; Gault, A.G.; Charnock, J.M.; Kalia, K.; Lloyd, J.R. Arsenate detoxification in a Pseudomonad hypertolerant to arsenic. *Arch. Microbiol.*, **2007**, *187*(3), 171-183.
[http://dx.doi.org/10.1007/s00203-006-0182-9] [PMID: 17160678]

[39] Mukhopadhyay, R.; Rosen, B.P.; Phung, L.T.; Silver, S. Microbial arsenic: From geocycles to genes and enzymes. *FEMS Microbiol. Rev.*, **2002**, *26*(3), 311-325.
[http://dx.doi.org/10.1111/j.1574-6976.2002.tb00617.x] [PMID: 12165430]

[40] Franke, S.; Grass, G.; Nies, D.H. The product of the ybdE gene of the Escherichia coli chromosome is involved in detoxification of silver ions. *Microbiology*, **2001**, *147*(Pt 4), 965-972.
[http://dx.doi.org/10.1099/00221287-147-4-965] [PMID: 11283292]

[41] Gupta, A.; Matsui, K.; Lo, J.F.; Silver, S. Molecular basis for resistance to silver cations in Salmonella. *Nat. Med.*, **1999**, *5*(2), 183-188.
[http://dx.doi.org/10.1038/5545] [PMID: 9930866]

[42] Bartolucci, S.; Contursi, P.; Fiorentino, G.; Limauro, D.; Pedone, E. Responding to toxic compounds: a genomic and functional overview of Archaea. *Front. Biosci.*, **2013**, *18*, 165-189.
[http://dx.doi.org/10.2741/4094] [PMID: 23276916]

[43] Nies, D.H. Efflux-mediated heavy metal resistance in prokaryotes. *FEMS Microbiol. Rev.*, **2003**, *27*(2-3), 313-339.
[http://dx.doi.org/10.1016/S0168-6445(03)00048-2] [PMID: 12829273]

[44] Paulsen, I.T.; Saier, M.H., Jr A novel family of ubiquitous heavy metal ion transport proteins. *J. Membr. Biol.*, **1997**, *156*(2), 99-103.
[http://dx.doi.org/10.1007/s002329900192] [PMID: 9075641]

[45] Kolaj-Robin, O.; Russell, D.; Hayes, K.A.; Pembroke, J.T.; Soulimane, T. Cation diffusion facilitator family: Structure and function. *FEBS Lett.*, **2015**, *589*(12), 1283-1295.
[http://dx.doi.org/10.1016/j.febslet.2015.04.007] [PMID: 25896018]

[46] Kambe, T. Molecular architecture and function of ZnT transporters.*Current topics in membranes*; Elsevier, **2012**, pp. 199-220.

[47] Smith, A.T.; Smith, K.P.; Rosenzweig, A.C. Diversity of the metal-transporting P1B-type ATPases. *J. Biol. Inorg. Chem.*, **2014**, *19*(6), 947-960.
[http://dx.doi.org/10.1007/s00775-014-1129-2] [PMID: 24729073]

[48] Schurig-Briccio, L.A.; Gennis, R.B. Characterization of the PIB-Type ATPases present in Thermus thermophilus. *J. Bacteriol.*, **2012**, *194*(15), 4107-4113.
[http://dx.doi.org/10.1128/JB.00849-12] [PMID: 22636781]

[49] Mandal, A.K.; Argüello, J.M. Functional roles of metal binding domains of the Archaeoglobus fulgidus Cu(+)-ATPase CopA. *Biochemistry*, **2003**, *42*(37), 11040-11047.
[http://dx.doi.org/10.1021/bi034806y] [PMID: 12974640]

[50] Tseng, T-T.; Gratwick, K.S.; Kollman, J.; Park, D.; Nies, D.H.; Goffeau, A.; Saier, M.H., Jr The RND permease superfamily: An ancient, ubiquitous and diverse family that includes human disease and development proteins. *J. Mol. Microbiol. Biotechnol.*, **1999**, *1*(1), 107-125.
[PMID: 10941792]

[51] Su, C.C.; Long, F.; Lei, H.T.; Bolla, J.R.; Do, S.V.; Rajashankar, K.R.; Yu, E.W. Charged amino acids (R83, E567, D617, E625, R669, and K678) of CusA are required for metal ion transport in the Cus efflux system. *J. Mol. Biol.*, **2012**, *422*(3), 429-441.
[http://dx.doi.org/10.1016/j.jmb.2012.05.038] [PMID: 22683351]

[52] Long, F.; Su, C.C.; Zimmermann, M.T.; Boyken, S.E.; Rajashankar, K.R.; Jernigan, R.L.; Yu, E.W. Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport. *Nature*, **2010**, *467*(7314), 484-488.

[http://dx.doi.org/10.1038/nature09395] [PMID: 20865003]

[53] Murakami, S.; Nakashima, R.; Yamashita, E.; Yamaguchi, A. Crystal structure of bacterial multidrug efflux transporter AcrB. *Nature*, **2002**, *419*(6907), 587-593.
[http://dx.doi.org/10.1038/nature01050] [PMID: 12374972]

[54] 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*, **2013**, *1834*(10), 2071-2079.
[http://dx.doi.org/10.1016/j.bbapap.2013.06.007] [PMID: 23800470]

[55] 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*(2), 427-437.
[http://dx.doi.org/10.1128/JB.186.2.427-437.2004] [PMID: 14702312]

[56] Smirnova, G.F. [Distribution of bacteria resistant to oxygen-containing anions-xenobiotics]. *Mikrobiol. Z.*, **2005**, *67*(5), 11-18. [Distribution of bacteria resistant to oxygen-containing anions-xenobiotics].
[PMID: 16396107]

[57] 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*(2), 99-106.
[http://dx.doi.org/10.1016/j.resmic.2008.10.012] [PMID: 19070660]

[58] 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*(1), 78.
[http://dx.doi.org/10.1186/s12934-018-0918-7] [PMID: 29776370]

[59] Ranawat, P.; Rawat, S. Metal-tolerant thermophiles: metals as electron donors and acceptors, toxicity, tolerance and industrial applications. *Environ. Sci. Pollut. Res. Int.*, **2017**, 1-29.
[PMID: 29238927]

[60] Coyle, P.; Philcox, J.C.; Carey, L.C.; Rofe, A.M. Metallothionein: The multipurpose protein. *Cell. Mol. Life Sci.*, **2002**, *59*(4), 627-647.
[http://dx.doi.org/10.1007/s000180200845 4-2] [PMID: 12022471]

[61] Amaro, F.; Turkewitz, A.P.; Martín-González, A.; Gutiérrez, J.C. Whole-cell biosensors for detection of heavy metal ions in environmental samples based on metallothionein promoters from Tetrahymena thermophila. *Microb. Biotechnol.*, **2011**, *4*(4), 513-522.
[http://dx.doi.org/10.1111/j.1751-7915.2011.00252.x] [PMID: 21366892]

[62] Ybarra, G.; Webb, R. Effects of divalent metal cations and resistance mechanisms of the cyanobacterium Synechococcus sp. strain PCC 7942. *Journal of Hazardous Substance Research*, **1999**, *2*(1), 1.
[http://dx.doi.org/10.4148/1090-7025.1011]

[63] Verma, N.; Singh, M. Biosensors for heavy metals. *Biometals*, **2005**, *18*(2), 121-129.
[http://dx.doi.org/10.1007/s10534-004-5787-3] [PMID: 15954738]

[64] Zhang, W. Nanomaterial-based biosensors for environmental and biological monitoring of organophosphorus pesticides and nerve agents. *TrAC Trends in Analytical Chemistry*, **2014**, *54*, 1-10.
[http://dx.doi.org/10.1016/j.trac.2013.10.007]

[65] Bontidean, I.; Berggren, C.; Johansson, G.; Csöregi, E.; Mattiasson, B.; Lloyd, J.R.; Jakeman, K.J.; Brown, N.L. Detection of heavy metal ions at femtomolar levels using protein-based biosensors. *Anal. Chem.*, **1998**, *70*(19), 4162-4169.
[http://dx.doi.org/10.1021/ac9803636] [PMID: 9784752]

[66] Aragay, G.; Merkoc, A. Recent trends in macro-, micro-, and nanomaterial-based Tools and strategies for heavy-metal detection. **2011**, 3433-3458.

[67] Belkin, S. Microbial whole-cell sensing systems of environmental pollutants. *Curr. Opin. Microbiol.*, **2003**, *6*(3), 206-212.
[http://dx.doi.org/10.1016/S1369-5274(03)00059-6] [PMID: 12831895]

[68] Su, L.; Jia, W.; Hou, C.; Lei, Y. Microbial biosensors: A review. *Biosens. Bioelectron.*, **2011**, *26*(5), 1788-1799.
[http://dx.doi.org/10.1016/j.bios.2010.09.005] [PMID: 20951023]

[69] Book, G. *Compendium of Chemical Terminology*; International Union of Pure and Applied Chemistry, **2014**, p. 528.

[70] Zeldes, B.M.; Keller, M.W.; Loder, A.J.; Straub, C.T.; Adams, M.W.; Kelly, R.M. Extremely thermophilic microorganisms as metabolic engineering platforms for production of fuels and industrial chemicals. *Front. Microbiol.*, **2015**, *6*, 1209.
[http://dx.doi.org/10.3389/fmicb.2015.01209] [PMID: 26594201]

[71] Zamost, B.L.; Nielsen, H.K.; Starnes, R.L. Thermostable enzymes for industrial applications. *J. Ind. Microbiol.*, **1991**, *8*(2), 71-81.
[http://dx.doi.org/10.1007/BF01578757]

[72] Vieille, C.; Zeikus, G.J. Hyperthermophilic enzymes: sources, uses and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.*, **2001**, *65*(1), 1-43.
[http://dx.doi.org/10.1128/MMBR.65.1.1-43.2001] [PMID: 11238984]

[73] Atomi, H.; Sato, T.; Kanai, T. Application of hyperthermophiles and their enzymes. *Curr. Opin. Biotechnol.*, **2011**, *22*(5), 618-626.
[http://dx.doi.org/10.1016/j.copbio.2011.06.010] [PMID: 21741818]

[74] Du, D.; Huang, X.; Cai, J.; Zhang, A. Comparison of pesticide sensitivity by electrochemical test based on acetylcholinesterase biosensor. *Biosens. Bioelectron.*, **2007**, *23*(2), 285-289.
[http://dx.doi.org/10.1016/j.bios.2007.05.002] [PMID: 17590326]

[75] Liu, Y.; Munteanu, C.R.; Fernandez-Lozano, C.; Pazos, A.; Ran, T.; Tan, Z.; Yu, Y.; Zhou, C.; Tang, S.; González-Díaz, H. Experimental study and ANN dual-time scale perturbation model of electrokinetic properties of microbiota. *Front. Microbiol.*, **2017**, *8*, 1216.
[http://dx.doi.org/10.3389/fmicb.2017.01216] [PMID: 28713345]

[76] Joshi, K.A. A disposable biosensor for organophosphorus nerve agents based on carbon nanotubes modified thick film strip electrode. *Electroanalysis*, **2005**, *17*(1), 54-58.
[http://dx.doi.org/10.1002/elan.200403118]

[77] 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*(43), 435703.
[http://dx.doi.org/10.1088/0957-4484/26/43/435703] [PMID: 26436536]

[78] 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*(123), 20160629.
[http://dx.doi.org/10.1098/rsif.2016.0629] [PMID: 27707908]

[79] Fiorentino, G.; Ronca, R.; Bartolucci, S. A novel E. coli biosensor for detecting aromatic aldehydes based on a responsive inducible archaeal promoter fused to the green fluorescent protein. *Appl. Microbiol. Biotechnol.*, **2009**, *82*(1), 67-77.
[http://dx.doi.org/10.1007/s00253-008-1771-0] [PMID: 18998120]

[80] Fernández, M.; Morel, B.; Ramos, J.L.; Krell, T. Paralogous Regulators ArsR1 and ArsR2 of Pseudomonas putida KT2440 as a Basis for Arsenic Biosensor Development. *Appl. Environ. Microbiol.*, **2016**, *82*(14), 4133-4144.
[http://dx.doi.org/10.1128/AEM.00606-16] [PMID: 27208139]

[81] 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. *N. Biotechnol.*, **2015**, *32*(1), 121-127.
[http://dx.doi.org/10.1016/j.nbt.2013.08.004] [PMID: 23994422]

[82] Johnson, D.B. Biomining - Biotechnologies for waste extracting and recovering metals from ores and waste materials. *Curr. Opin. Biotechnol.*, **2014**, 24-31.
[http://dx.doi.org/10.1016/j.copbio.2014.04.008]

[83] Gadd, G.M. Bioremedial potential of microbial mechanisms of metal mobilization and immobilization. *Curr. Opin. Biotechnol.*, **2000**, *11*(3), 271-279.
[http://dx.doi.org/10.1016/S0958-1669(00)00095-1] [PMID: 10851150]

[84] Dunbar, W.S. Biotechnology and the mine of tomorrow. *Trends Biotechnol.*, **2017**, *35*(1), 79-89.
[http://dx.doi.org/10.1016/j.tibtech.2016.07.004] [PMID: 27612568]

[85] Donati, E.R.; Castro, C.; Urbieta, M.S. Thermophilic microorganisms in biomining. *World J. Microbiol. Biotechnol.*, **2016**, *32*(11), 179.
[http://dx.doi.org/10.1007/s11274-016-2140-2] [PMID: 27628339]

[86] 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*(1), 155-163.
[http://dx.doi.org/10.1007/s11274-011-0804-5] [PMID: 22806791]

[87] 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*(6), 4020-4027.
[http://dx.doi.org/10.1128/AEM.00295-06] [PMID: 16751511]

[88] Chatterjee, S.K.; Bhattacharjee, I.; Chandra, G. Biosorption of heavy metals from industrial waste water by Geobacillus thermodenitrificans. *J. Hazard. Mater.*, **2010**, *175*(1-3), 117-125.
[http://dx.doi.org/10.1016/j.jhazmat.2009.09.136] [PMID: 19864059]

[89] Özdemir, S. Biosorption of Cd, Cu, Ni, Mn and Zn from aqueous solutions by thermophilic bacteria, Geobacillus toebii sub. sp. decanicus and Geobacillus thermoleovorans sub. sp. stromboliensis: Equilibrium, kinetic and thermodynamic studies. *Chem. Eng. J.*, **2009**, *152*(1), 195-206.
[http://dx.doi.org/10.1016/j.cej.2009.04.041]

[90] Valenzuela, L.; Chi, A.; Beard, S.; Orell, A.; Guiliani, N.; Shabanowitz, J.; Hunt, D.F.; Jerez, C.A. Genomics, metagenomics and proteomics in biomining microorganisms. *Biotechnol. Adv.*, **2006**, *24*(2), 197-211.
[http://dx.doi.org/10.1016/j.biotechadv.2005.09.004] [PMID: 16288845]

[91] Zammit, C. Evaluation of quantitative real-time polymerase chain reaction for enumeration of biomining microorganisms in culture. *Hydrometallurgy*, **2008**, *94*(1-4), 185-189.
[http://dx.doi.org/10.1016/j.hydromet.2008.05.034]

[92] 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*(1-4), 229-236.
[http://dx.doi.org/10.1016/j.hydromet.2006.03.027]

[93] Miller, P.; Jiao, F.; Wang, J. The bacterial oxidation (BACOX) plant at Laizhou, Shandong Province, China–the first three years of operation. *Proceedings of the Bac-Min 2004 Conference*, **2004**.

[94] Zhou, H-B.; Zeng, W.M.; Yang, Z.F.; Xie, Y.J.; Qiu, G.Z. Bioleaching of chalcopyrite concentrate by a moderately thermophilic culture in a stirred tank reactor. *Bioresour. Technol.*, **2009**, *100*(2), 515-520. [http://dx.doi.org/10.1016/j.biortech.2008.06.033] [PMID: 18657418]

[95] Brierley, C.L.; Brierley, J.A. Progress in bioleaching: part B: Applications of microbial processes by the minerals industries. *Appl. Microbiol. Biotechnol.*, **2013**, *97*(17), 7543-7552. [http://dx.doi.org/10.1007/s00253-013-5095-3] [PMID: 23877580]

Chapter III

The biological, chemical-physical characterization of *Tt*SmtB, a transcriptional factor belonging to the ArsR/SmtB family

In order to develop a *whole-cell* biosensor for environmental monitoring of heavy metal ions, a detailed understanding of the molecular mechanisms responsible for resistance to toxic metals is a crucial.

In this chapter we focus on *Tt*SmtB, a transcription factor belonging to the ArsR/SmtB family.

This chapter describes the biological role of the metal sensor *Tt*SmtB transcriptional factor putting emphasis on the identification of the specific sequence recognised on the DNA and the chemical-physical characterization; the results obtained are organized in: Paper I) An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in *Thermus thermophilus* HB27; and Paper II) A physicochemical investigation on the metal binding properties of *Tt*SmtB, a thermophilic member of the ArsR/SmtB transcription factor family.

The Paper I report a detailed *in vivo* and *in vitro* characterization of *Tt*SmtB with identification of target DNA binding sites, regulated genes and response to arsenic. This study allowed us to understand the biological role of this protein that is transcriptional regulation of the expression of arsenic resistance genes in *Thermus thermophilus* HB27 upon arsenic challenge.

The Paper II describes the complete characterization of the chemical-physical abilities of *Tt*SmtB. This study allowed us to better understand the mechanism of interaction with metals, the specificity as well as resistance to extreme conditions of this protein. Furthermore, this study has provided us fundamental information for the development of a new type of *whole-cell* biosensor for cadmium and arsenic.

microbial biotechnology

Open Access

An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in *Thermus thermophilus* HB27

Immacolata Antonucci,[1] Giovanni Gallo,[1] Danila Limauro,[1] Patrizia Contursi,[1] Ana Luisa Ribeiro,[2] Alba Blesa,[2] José Berenguer,[2] Simonetta Bartolucci[1] and Gabriella Fiorentino[1,*]

[1] *Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Complesso Universitario Monte S. Angelo, 80126, Naples, Italy.*

[2] *Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Campus Universidad Autónoma de Madrid 28049, Madrid, Spain.*

Summary

Arsenic resistance is commonly clustered in *ars* operons in bacteria; main *ars* operon components encode an arsenate reductase, a membrane extrusion protein, and an As-sensitive transcription factor. In the As-resistant thermophile *Thermus thermophilus* HB27, genes encoding homologues of these proteins are interspersed in the chromosome. In this article, we show that two adjacent genes, *TtsmtB,* encoding an ArsR/SmtB transcriptional repressor and, *TTC0354,* encoding a Zn^{2+}/Cd^{2+}-dependent membrane ATPase are involved in As resistance; differently from characterized *ars* operons, the two genes are transcribed from dedicated promoters upstream of their respective genes, whose expression is differentially regulated at transcriptional level. Mutants defective in *TtsmtB* or *TTC0354* are more sensitive to As than the wild type, proving their role in arsenic resistance. Recombinant dimeric *Tt*SmtB binds *in vitro* to both promoters, but its binding capability decreases upon interaction with arsenate and, less efficiently, with arsenite. *In vivo* and *in vitro* experiments also demonstrate that the arsenate reductase (*Tt*ArsC) is subjected to regulation by *Tt*SmtB. We propose a model for the regulation of As resistance in *T. thermophilus* in which *Tt*SmtB is the arsenate sensor responsible for the induction of *Tt*ArsC which generates arsenite exported by TTC0354 efflux protein to detoxify cells.

Received 7 September, 2016; revised 5 June, 2017; accepted 6 June, 2017. *For correspondence. E-mail fiogabri@unina.it; Tel. +39081679167.

Microbial Biotechnology (2017) **10**(6), 1690–1701
doi:10.1111/1751-7915.12761

Funding Information

This study was supported by grants from the Regione Campania, legge 5 (Italy, CUP number E69D15000210002) to SB, and from the project BIO2016-77031-R of the Spanish Ministry of Economy and Competitiveness to JB.

Introduction

Arsenic (As) is an ubiquitous metalloid naturally present in soil, water and air that adversely affects human and animal health. Because of its abundance and toxicity, monitoring arsenic concentration in the environment and in several foodstuffs used for human consumption is very important.

Under reducing conditions, the highly toxic trivalent arsenite, As(III), is the more abundant form, whereas in oxygenated environments, the less-toxic and more stable pentavalent arsenate, As(V), dominates. Arsenite enters the cell through aquaglyceroporins; as it has a high affinity for sulfur, it exerts its toxicity through binding to dithiols in proteins, in glutathione (GSH) and in lipoic acid contributing to protein/enzyme inactivation (Liu *et al.*, 2004; Meng *et al.*, 2004). On the other hand, arsenate enters the cells through phosphate transporters and its toxicity is mediated by replacing phosphate in essential biochemical reactions (Tawfik and Viola, 2011; Kamerlin *et al.*, 2013).

The abundance of arsenic in the environment has guided the evolution of multiple defence strategies in almost all microorganisms (Contursi *et al.*, 2013); for instance, despite being toxic, some microorganisms also use arsenic as electron acceptor in anaerobic respiratory chains or as electron donor for chemolythotrophic growth and even for anoxigenic photosynthesis (Kulp *et al.*, 2008; van Lis *et al.*, 2013); other microorganisms are able to methylate inorganic arsenic or de-methylate the organic forms (Qin *et al.*, 2006).

In general, in many prokaryotes, arsenic resistance is linked to the presence of plasmid- or chromosomally encoded *ars* operons with a variable number of genes. The simplest resistance system involves cytoplasmic reduction of As(V) to As(III) by arsenate reductase and further extrusion of As(III) by a membrane protein whose expression is regulated by ArsR, a trans-acting repressor of the ArsR/

SmtB family (Wysocki *et al*., 2003; Jacobson *et al*., 2012). Arsenate reductases use thioredoxin, glutaredoxin or mycoredoxin as electron donors (Rosen, 2002), whereas arsenite extrusion is mediated by two families of proteins: ArsB proteins, that have been found only in bacteria (Rosen, 1999), and Acr3 proteins, with representatives in bacteria, fungi and plants (Indriolo *et al*., 2010). Two additional genes encoding an ATPase component (ArsA) of the arsenite transporter (ArsB) and a metallochaperon (ArsD) can increase the efficiency of the arsenite efflux system (Lin *et al*., 2006) in prokaryotes. Genomic analysis on thousands of microorganisms revealed new genes in *ars* operons with unknown functions. As an example, parallel pathways for organic arsenicals have been recently identified (Yang and Rosen, 2016).

Regarding transcriptional regulators, different families of metal-sensing proteins (identified as family HTH_5 in the Pfam database) have been described in bacteria, with ArsR/SmtB being the most extensively studied and named after its founding members, *Escherichia coli* ArsR and *Synechococcus* PCC 7942 SmtB (Wu and Rosen, 1991; Morby *et al*., 1993). The members of the ArsR/SmtB family have many common features, but also display a great diversity in metal-sensing motifs and metal-binding mechanisms. They share a dimeric structure and contain a helix–turn–helix (HTH) or winged HTH DNA-binding domain, and their sequence includes an ELCV (C/G)D motif, defined as the metal-binding box, located within the HTH region (Shi *et al*., 1994; Cook *et al*., 1998). Binding of a metal to this motif interferes with DNA binding. The SmtB protein binds to imperfect 12-2-12 inverted repeats (or a half of this site) located within the operator–promoter region or overlapping the transcriptional start site of the regulated promoters. As binding to the metal alleviates transcriptional repression, these factors act as metal-sensitive transcription repressors (Osman and Cavet, 2010). Despite conservation of the metal-binding motif, the selectivity for the metal and its binding mode differ among SmtB homologues due to the capability of the conserved Cys residues to form metal-thiolate bonds of different geometry and coordination (Guerra and Giedroc, 2012). This is consistent with the hypothesis that metal-binding sites in DNA-binding proteins have evolved convergently in response to environmental pressures (Ordonez *et al*., 2008).

With all of the activities above described, microbes actively participate in the geochemical cycling of arsenic in the environment, promoting or inhibiting arsenic release (Fernandez *et al*., 2014). Specifically, thermophilic microorganisms influence the biogeochemistry of arsenic compounds in different geothermal springs (Donahoe-Christiansen *et al*., 2004). However, to date, information regarding the molecular mechanisms of arsenic resistance in these habitats is still preliminary although it is an essential prerequisite to develop effective systems for arsenic sensing and monitoring in the environment.

The thermophilic Gram-negative bacterium *Thermus thermophilus* HB27 is capable of growing in the presence of arsenate and arsenite in concentrations that are lethal for other microorganisms. The putative resistance genes have not been found in a single resistance operon but scattered and associated with chromosomal genes apparently not functionally related. In a recent work, we demonstrated the involvement of a thioredoxin-coupled arsenate reductase (*Tt*ArsC) in the As-resistance mechanism and hypothesized that arsenic-dependent induction of *TtarsC* could be mediated by factors such as ArsR/SmtB transcriptional regulators (Del Giudice *et al*., 2013).

In this study, we employ a genetic and biochemical approach to demonstrate that two adjacent genes, *TTC0353* (*TtsmtB*), encoding a putative ArsR/SmtB transcriptional repressor, and *TTC0354*, encoding a recently described membrane Zn^{2+}/Cd^{2+} ATPase (Schurig-Briccio and Gennis, 2012), play a key role in the arsenic resistance.

Results

TtsmtB and TTC0354 are arsenic-regulated genes

The *TTC0353* gene of *T. thermophilus* HB27 encodes a putative protein annotated as a member of the ArsR/SmtB family of transcriptional regulators herein named *Tt*SmtB. *Tt*SmtB is a 123-amino-acid-long protein (predicted molecular weight of 13 508.79 Da and a pI of 8.54) with a HTH DNA-binding motif and a conserved ELCVCD metal-binding box located in the α-3/α-4 helices and in the α-4 helix respectively (Fig. S1A). Sequence alignment showed 50% of identity with the structurally characterized SmtB transcriptional repressor from *Synechococcus* PCC 7942 (VanZile *et al*., 2000). Secondary and tertiary structure predictions (using the software I-TASSER at http://zhanglab.ccmb.med.umich.edu/I-TASSER/) revealed an organization in six α-helices and one β-sheet comparable to that found in ArsR/SmtB regulators (Cook *et al*., 1998; Fig. S1B). Moreover, the presence of the conserved metal-binding box and of a cysteine residue (Cys 10) at the N-terminus putatively involved in metal binding strongly suggests a key role for this protein in metal sensing.

TTC0354 is separated from *TtsmtB* by 32 bp and encodes a putative membrane metal transporter (predicted molecular weight of 71 814.95 Da and a pI of 8.21) with a heavy-metal-associated (HMA) motif displaying ATPase activity stimulated by cations, whose role in metal tolerance is not clear (Schurig-Briccio and Gennis, 2012; Fig. S1C).

To analyse whether the expression *of TtsmtB* and *TTC0354* was regulated by arsenic, qRT-PCR assays were carried out on RNAs isolated from cells treated with

arsenate or arsenite at subinhibitory concentrations. As shown in Fig. 1, both ions determine significant increase in the transcription of *TtsmtB* (15-fold and 9-fold) and *TTC0354* (4-fold and 3-fold).

Identification of the promoters of the arsenic-related genes

To better understand the function and the regulation of *TtsmtB* and *TTC0354*, we first verified whether the two genes were transcribed from independent promoters (Fig. 2A). For this, we identified transcription start site(s) (TSS) by primer extension. A TSS was detected for *TTC0354* on total RNA isolated from cells grown without As and upstream from this TSS, −10 and −35 boxes could be identified. Interestingly, this region contained a palindromic sequence (6-2-6) matching the consensus of the ArsR/SmtB-binding site (Osman and Cavet, 2010); a putative ATG start codon 38 bp downstream and a conserved ribosome-binding site (RBS) (−7 from this ATG) were also identified (Fig. 2B).

On the other hand, a TSS could be also identified for *TtsmtB* at position − 3 from the first putative translated nucleotide (TTG) when cells were treated with arsenic, suggesting low expression levels under basal conditions. Furthermore, consensus-like promoter boxes were not found at the appropriate upstream distance from this TSS probably due to the fact that this promoter is intragenic (Fig. 2C); in fact, *TtsmtB* gene overlaps with an upstream gene (*TTC0352*) encoding the putative small subunit of pyridoxal 5′-phosphate (PLP) synthase, an enzyme for *de novo* biosynthesis of PLP coenzyme, with no obvious functional relation to As resistance.

Role of TtSmtB and TTC0354 in arsenic resistance

As the aim of our study was to establish a role for *Tt*SmtB and TTC0354 in the arsenic resistance, the genes were inactivated by insertion of a gene cassette encoding a thermostable resistance to kanamycin in *TtsmtB* (Fig. S2) or the suicide vector pK18 in *TTC0354* (Fig. S3). The screening of kanamycin-resistant recombinants was carried out by PCR, and the deletion/insertion was confirmed by sequence analysis.

In the absence of arsenic, mutant and wild-type strains grew at similar rates, showing that the genes were not relevant for *T. thermophilus* viability under such conditions. Nevertheless, when they were phenotypically analysed by MIC assays, a decrease in arsenate and arsenite resistance in both mutants was observed (Table 1). Inactivation of *TtsmtB* resulted in 2.5-fold and 1.5-fold reductions in the MIC of arsenate and arsenite, respectively, while mutation of *TTC0354* leads to 15-fold and 13-fold reduction respectively. These results indicate that the two genes are relevant for arsenic resistance and are consistent with the predicted functions as an arsenic transcriptional regulator and an efflux transporter respectively.

Purification and structural characterization of TtSmtB

To better characterize the regulatory role of *Tt*SmtB, the corresponding gene was cloned in pET28b (+) plasmid to generate a His-tagged fusion protein that was expressed in *E. coli* BL21-CodonPlus(DE3)-RIL and purified to homogeneity.

Far-UV circular dichroism spectra showed a typical circular dichrogram of a helical protein with negative maxima at 208 and 222 nm and one positive peak at 195 nm (Fig. 5A), indicative of a predominantly folded structure with an α-β content. Data obtained from deconvolution of the CD spectra established that the recombinant protein was properly folded with approximately 32.6% α-helix and 37.1% turn sheet, according to the structural model (Fig. S1). Furthermore, gel filtration experiments showed that the protein is a homodimer of about 27 kDa in solution.

As *Tt*SmtB is a dimer with three cysteine residues (Cys 10, Cys 62 and Cys 64) per subunit, we analysed whether they were involved in intramolecular disulfide

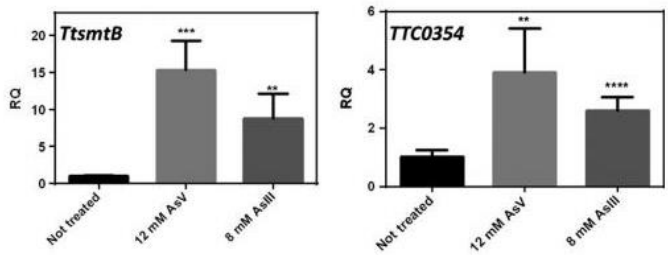


Fig. 1. Differential induction of *TtsmtB* and *TTC0354*. qRT-PCR expression analysis of *TtsmtB* (*TTC0353*) and *TTC0354*, from *T. thermophilus* HB27 exponential cultures grown for 45 min with arsenate (12 mM) or arsenite (8 mM). Error bars indicate the standard deviation of the average values in two independent experiments in triplicate samples (**$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$).

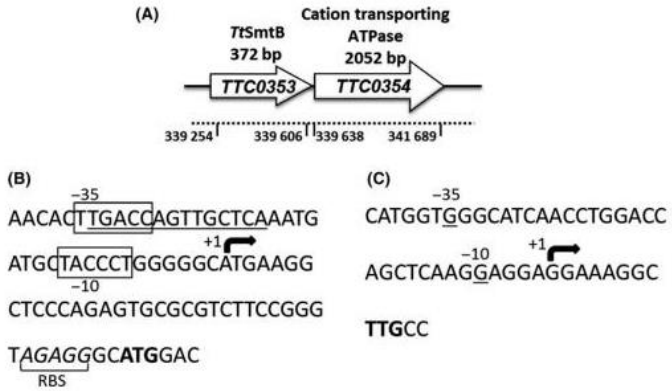


Fig. 2. Transcriptional start site determination of *TTC0354* and *TtsmtB*. (A). Genomic context of *TtsmtB* and *TTC0354*. (B). *TTC0354* promoter sequence analysis. The mapped transcription start site (+1) is marked with an arrow. −10 and −35 regions from the transcription start site are boxed. Predicted SmtB-binding site is underlined. The translation start codon is indicated in bold. RBS is marked. (C). Sequence of *TtsmtB* promoter region. The mapped transcription start site (+1) is marked with an arrow. −10 and −35 nt from the transcription start site are underlined. The translation start codon is indicated in bold.

Table 1. Bacterial resistance to arsenic.

	MIC	
Strain	As(V)	As(III)
Thermus thermophilus HB27	44 mM	40 mM
Thermus thermophilus ΔsmtB::kat	18 mM	32 mM
Thermus thermophilus TTC0354::pK18	3 mM	3 mM

bridges or not by treating the protein with iodoacetamide, followed by trypsin (and chymotrypsin) digestion, and MALDI-TOF mass spectrometry peptide analysis (70.5% sequence coverage). The mass of peptides containing the cysteines was increased of 57 Da indicating that such residues are not involved in disulfide bridges and suggesting that they are likely involved in metal coordination.

Binding analysis of TtSmtB to sequences upstream of TtsmtB, TTC0354 and TtarsC

To verify whether *Tt*SmtB recognizes the identified regulatory regions, we performed EMSAs. We tested the following: (i) *TtsmtB* promoter to verify autoregulation; (ii) *TTC0354* promoter, the only containing a palindromic sequence matching the consensus of the ArsR/SmtB-binding site in other organisms; (iii) the region upstream of *TtarsC* (*TTC1502*), to examine whether arsenate reductase could be a regulatory target of *Tt*SmtB (Del Giudice *et al.*, 2013). All the DNA fragments, obtained through PCR amplifying 149, 143 and 78 bp regions upstream of *TtsmtB*, *TTC0354* and *TtarsC*, respectively, were incubated with 2.5 and 7.5 μM of purified recombinant protein. The results show that *Tt*SmtB can form complexes with different mobility and intensity in the three promoters in a concentration-dependent manner, suggesting that the protein binds to them differentially (Fig. 3).

Further analyses were performed to better characterize the interaction of *Tt*SmtB with the putative promoter controlling *TTC0354* (*0354*p) as well as its own promoter (*TtsmtB*p) (Fig. S4). We first assessed whether *Tt*SmtB binding to the promoters was specific; the *0354*p-*Tt*SmtB complexes dissociated in the presence of an excess of unlabelled specific probe whereas remained bound in the presence of a molar excess (200× and 400×) of unspecific DNA. Titration with increasing concentrations of *Tt*SmtB indicated that the protein binds to this region

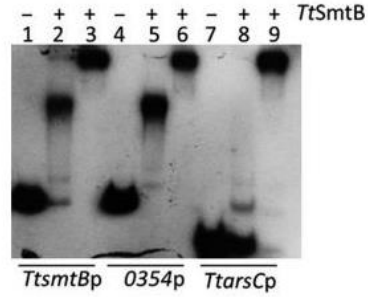
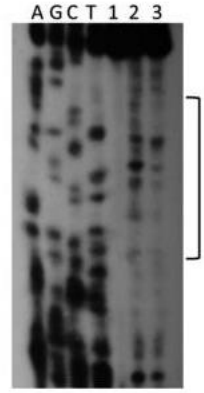


Fig. 3. Activity of *Tt*SmtB. Binding of 2.5 μM (lanes 2, 5, 8) and 7.5 μM (lanes 3, 6, 9) *Tt*SmtB to *TtsmtB*p, *0354*p and *TtarsC*p.

in a concentration-dependent manner; at saturating concentrations, the protein determined a shift with low mobility, suggesting either that other binding sites with different affinities could exist in the DNA sequence analysed or that multiple dimers could associate with the cognate DNA (Kar *et al.*, 2001). The profile obtained by fitting densitometric data to a binding curve with a Hill slope gave an overall apparent equilibrium dissociation constant (K_d) of 1.4 ± 0.07 μM, suggesting that DNA binding is cooperative, as reported for other characterized SmtB family members (Mandal *et al.*, 2007; Chauhan *et al.*, 2009). On the other hand, titration of *Tt*SmtB showed that the protein also forms multiple complexes with its own promoter and binds cooperatively but with lower affinity (K_d of 5.0 ± 0.3 μM) and lower specificity (Fig. S4).

To map the *Tt*SmtB-binding site on the *TTC0354* promoter, DNase I footprinting was performed using a 143 bp DNA including sequences −100 to +42 from the mapped transcription start site. As shown in Fig. 4, a 28 bp protected region was observed containing the TTGCTCAA sequence matching a consensus with binding sites of ArsR/SmtB members. Interestingly, this sequence overlaps to −35 box, strongly suggesting that *Tt*SmtB binding at this promoter could hamper RNA polymerase binding.



−35
5′-AAAACACTTGACCAGTTGCTCAAATGAT
+1
−10
GCTACCCTGGGGGCATGAAGGCTC-3′

Fig. 4. DNase I footprinting analysis of *Tt*SmtB on *TTC0354* promoter. Footprinting was performed at the non-template strand using 0.0 μg (lanes 1–2) or 4 μg (~3 μM, lane 3) of purified *Tt*SmtB, in the absence (lane 1) or in the presence of 3 U of DNase I (lanes 2–3). DNA fragments were analysed in parallel with a sequencing reaction by denaturing gel electrophoresis. Position of the footprint (sequence underlined) relative to the *TTC0354* promoter is shown.

TtSmtB in vitro *interaction with arsenic*

The subsequent task was to determine whether *Tt*SmtB binds arsenic *in vitro* and whether the interaction modifies its DNA-binding ability.

Circular dichroism was employed to investigate the secondary structure of *Tt*SmtB in the presence or absence of increasing concentrations of arsenate and arsenite. The results in Fig. 5A show that the absolute values of the negative peaks at 208 and 220 nm, that are related to the α-helical content in the protein structure, were progressively increasing at higher arsenate and arsenite concentrations revealing that *in vitro* As determines changes in *Tt*SmtB secondary structure. In more detail, the values of the molar ellipticity per residue at 208 nm, plotted against arsenate and arsenite concentrations, gave binding curves with an apparent equilibrium dissociation constant K_d of approximately 0.06 and 0.25 mM respectively (not shown). These values so diverse suggest that arsenate and arsenite cause structural rearrangements with different effect on *Tt*SmtB activity.

Finally, to analyse whether arsenic influences *Tt*SmtB DNA-binding ability, we performed EMSAs in which DNA-*Tt*SmtB complex formation was tested after pre-incubation of the protein with arsenate and arsenite each at 50-fold and 100-fold molar excess. The results reported in Fig. 5B and C demonstrate that the interaction of arsenate and arsenite with *Tt*SmtB hampers binding to *0354*p although at different concentrations; in fact, in the presence of arsenate, complex 2 is disrupted at lower concentration than when arsenite is used (Fig. 5B and C lanes 3–4).

TtSmtB is a repressor in vivo

To confirm the *in vivo* role of *Tt*SmtB as a repressor, we used qRT-PCR to evaluate in the *ΔsmtB* mutant the expression of the arsenic resistance genes *TTC0354* and *TtarsC* (Del Giudice *et al.*, 2013). As shown in Fig. 6, 3-fold and 2-fold increases in *TTC0354* and *TtarsC* expressions, respectively, were detected in *ΔsmtB* compared to the wild-type strain. These results further suggest that the genes investigated are functional targets of *Tt*SmtB that exerts a negative regulation of their transcription *in vivo*.

Discussion

Thermophilic microorganisms play important roles in arsenic bioavailability in thermal environments and represent good models either to investigate the molecular mechanisms of response to metal stress

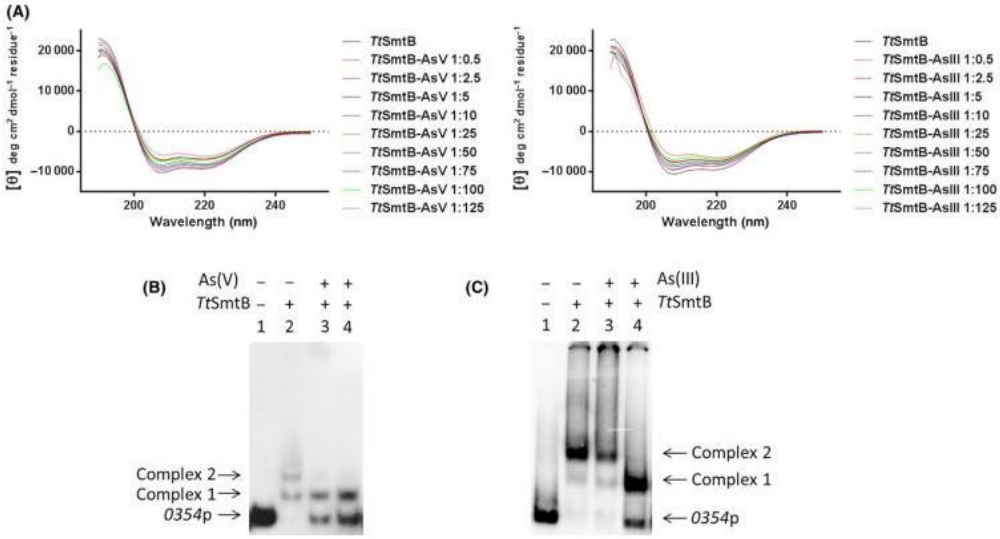


Fig. 5. *Tt*SmtB interaction with arsenic. (A). Far-UV CD spectrum of *Tt*SmtB with increasing amounts of As(V) and As(III). (B). Binding of *Tt*SmtB to *0354p* without (lane 2) and with arsenate at molar ratio of 1:50 and 1:100 (lanes 3–4). (C). Binding of *Tt*SmtB to *0354p* without (lane 2) and with arsenite at molar ratio of 1:50 and 1:100 (lane 3–4).

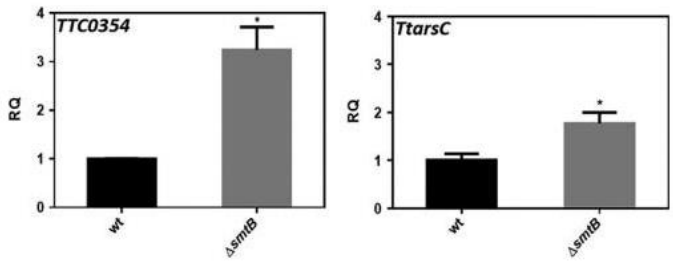


Fig. 6. qRT-PCR analysis of *TTC0354* and *TtarsC* in *T. thermophilus* HB27 (wt) and Δ*smtB* grown without As. The error bars indicate the standard deviation of the average values of two independent experiments in triplicate samples (**P* < 0.05).

(Bartolucci *et al.*, 2013) or for the development of robust biosensors for arsenic monitoring (Politi *et al.*, 2015; Fernandez *et al.*, 2016). In a recent study, we demonstrated that the Gram-negative hyperthermophilic bacterium *T. thermophilus* HB27 can tolerate high concentrations of arsenic and that the genes putatively involved in arsenic resistance are scattered in the chromosome; in particular, a thermostable arsenate reductase (*Tt*ArsC), an enzyme capable of reducing pentavalent arsenate to trivalent arsenite, was characterized (Del Giudice *et al.*, 2013; Politi *et al.*, 2016).

In the present study, we demonstrate that *Tt*SmtB, encoded by *TTC0353*, is an ArsR/SmtB transcriptional factor exerting negative regulation on two genes, *TtarsC* and *TTC0354*, involved in arsenic resistance. In the genome annotation, *TTC0353* is adjacent to *TTC0354* encoding an ATPase stimulated by cations, particularly Cu^+, with a heavy-metal-associated domain (Schurig-Briccio and Gennis, 2012). The genomic localization of *TTC0354* and structure predictions of the encoded protein (Fig. S1C) led us to hypothesize that TTC0354 could function *in vivo* as the arsenite efflux transporter, so we further investigated this point. Inactivation of

TTC0354 was decisive to point out its role in arsenic resistance; in fact, *T. thermophilus TTC0354::pK18* is significantly more sensitive to arsenic treatment.

The *ΔsmtB* mutant strain was also obtained; in this strain, the expression levels of *TTC0354* are higher than in the wild type providing experimental evidence that *Tt*SmtB represses gene transcription *in vivo.* Although the derepression was not particularly strong, the data were statistically significant. Moreover, qRT-PCR experiments performed on the wild type showed that *TtsmtB* transcription is increased upon arsenate/arsenite treatment, confirming its involvement in the arsenic response. Nevertheless, as judged by primer extension experiments, *TtsmtB* expression under basal conditions turned to be low, congruent with its role as a transcriptional regulator. Therefore, the steep induction of *TtsmtB* transcription upon exposure to arsenate/ite is functional to initially cope with the increase in arsenic concentration and afterwards to restore the negative transcriptional control over the key genes once the arsenic has been cleared out.

To elucidate the *in vitro* role of *Tt*SmtB in arsenic sensing and gene regulation, we produced a recombinant form of the protein and demonstrated that, as other members of its family, it is a dimer in solution and has three cysteine residues in a reduced form available to coordinate arsenic binding. EMSAs showed that *Tt*SmtB can form multiple complexes and binds to the different promoters that we tested with high cooperativity; the different shifted bands might represent different forms of the DNA–protein complex as already proposed for the SmtB protein of *Synechococcus* PCC7942 (Kar *et al.*, 2001). According to *in vivo* expression data, cooperativity could guarantee that even small variations in intracellular protein concentration can regulate tunable occupancy of DNA target sequences. *Tt*SmtB more strongly and specifically interacts with *TTC0354* promoter through binding to a sequence, identified through DNase I footprinting, corresponding to a consensus palindromic ArsR/SmtB-binding site, overlapping putative basal transcription elements; this mode of DNA recognition is another common feature of the ArsR/SmtB protein family (Osman and Cavet, 2010). Sequence alignments showed that *Tt*SmtB has one of the three cysteine residues required to form the As(III)-binding site at the N-terminus end outside the metal-binding box (Fig. S1A), so it is conceivable that the geometry of the metal-binding site could favour arsenate binding (Shi *et al.*, 1994). Circular dichroism spectroscopy experiments indicate that both the arsenic ions determine changes in *Tt*SmtB secondary structure and suggest that the protein binds arsenate more efficiently than arsenite. Furthermore, both arsenate and arsenite are able to trigger *Tt*SmtB release from DNA, suggesting that *in vivo* the transcription of target genes could be repressed under basal conditions and allosterically activated upon increase in intracellular arsenate/ite concentration. Interestingly, arsenate binding by *Tt*SmtB hampers DNA interaction better than arsenite, further indicating that in this system, arsenate could be the actual effector of *Tt*SmtB. To our knowledge, this is the first demonstration of an ArsR protein preferentially interacting with arsenate.

Based on the results presented here, a model for the role of *Tt*SmtB in As resistance in *T. thermophilus* HB27 is proposed in Fig. 7. Under basal conditions, *Tt*SmtB exerts autoregulation expressing itself at low levels and represses *TTC0354* and *TtarsC* (Fig. 7A). The entry of arsenate into the cells determines structural changes in *Tt*SmtB with the effect of derepressing at first the low-affinity promoters, with consequent increase in free *Tt*SmtB concentration and in the arsenate reductase *Tt*ArsC (Fig. 7B). Further accumulation of arsenate and arsenite, the latter as product of the *Tt*ArsC-catalyzed reaction, leads to the full derepression of the TTC0354 efflux pump allowing the efficient transport of arsenite out of the cells (Fig. 7C).

Our findings do not exclude that *TtsmtB/TTC0354/TtArsC* system is the only responsible to cope with arsenic toxicity and suggest that alternative regulatory circuits or protein partners could participate, as already seen in other microorganisms (Wang *et al.*, 2004). Moreover, both *Tt*SmtB and TTC0354 might play a broader role in response to metal stress.

Altogether these results give insights into the mechanisms of metal-regulated gene expression in *T. thermophilus* pointing out to substantial differences with better characterized *ars* systems (Fernandez *et al.*, 2016); the role of *Tt*SmtB and TTC0354 and their regulatory sequences in arsenic sensing add a new piece in the puzzle of the molecular machinery of *T. thermophilus* arsenic resistance and represent an important progress both for the development of effective, safe and stable whole-cell arsenic biosensors and/or for the exploitation of novel bioremediation processes.

Experimental procedures

Bacterial strains and culture conditions

Escherichia coli strains were grown in Luria Bertani (Miller, 1972) medium at 37 °C with 50 μg ml^{-1} kanamycin and/or 33 μg ml^{-1} chloramphenicol and/or hygromycin B as required. *T. thermophilus* HB27 wild-type strain (purchased from DSMZ) was grown aerobically at 70 °C in TM medium without or with 8 mM $NaAsO_2$ (Sigma, referred to throughout this article as arsenite) or 12 mM KH_2AsO_4 (Sigma, referred to throughout this article as arsenate) as described (Del Giudice *et al.*, 2013); 8 mM arsenite and 12 mM arsenate were chosen because

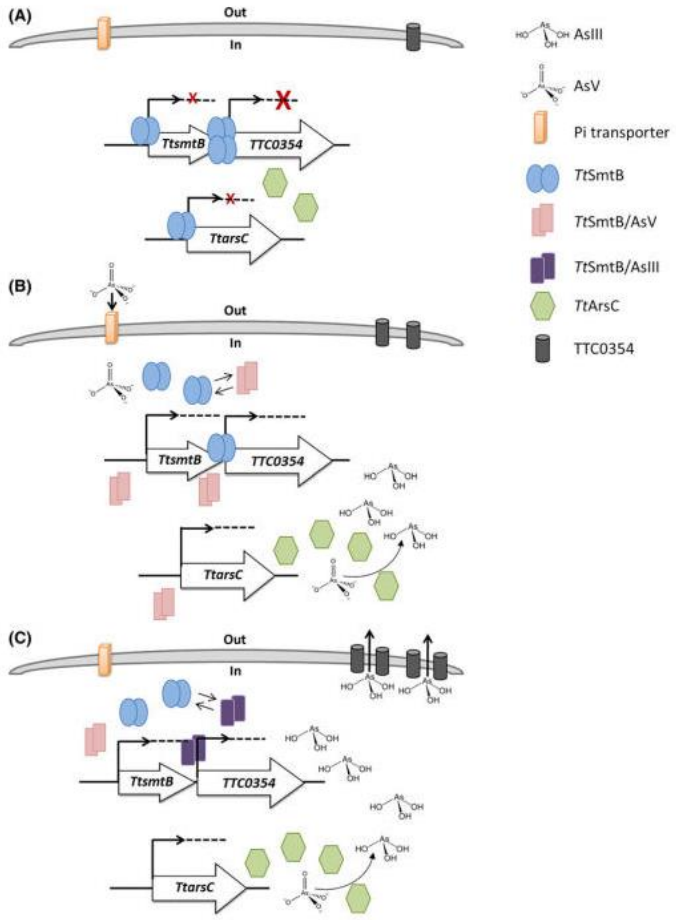


Fig. 7. Cartoon of the role of *Tt*SmtB in arsenic resistance. (A). In the absence of arsenic, the dimeric *Tt*SmtB is expressed at low levels and binds target promoters of the indicated genes eliciting transcriptional repression. As demonstrated by *in vitro* DNA-binding experiments, *0354p* is more tightly repressed. (B). When arsenate enters the cells, it binds to *Tt*SmtB lowering its affinity towards DNA and inducing derepression of *TtsmtB* and *TtarsC*; derepression of *TtarsC* allows a catalytic conversion of arsenate to arsenite. (C). Further intracellular accumulation of arsenite determines complete release of *0354p*, transcriptional activation of the pump and extrusion of arsenite outside the cells. The reduction in intracellular arsenate levels coupled to the efficient export of arsenite restores transcriptional repression.

they correspond to subinhibitory concentration values determined in a previous work (Del Giudice *et al.*, 2013). *T. thermophilus ΔsmtB::kat* and *TTC0354::pK18* mutants were grown aerobically at 70 °C in TM medium containing kanamycin (30 μg ml^{-1}).

For qRT-PCR experiments, 50 ml cultures of *T. thermophilus* HB27 and *ΔsmtB::kat* were grown up to 0.5 $OD_{600\ nm}$ and harvested at 0 and 45 min after the addition of 8 mM $NaAsO_2$ or 12 mM KH_2AsO_4 and immediately spun down, and pellets were kept at −80 °C.

For the determination of minimum inhibitory concentration (MIC), exponentially growing cultures were diluted to OD_{600} = 0.08 in 24-well plates (Corning, New York, USA) in TM medium with increasing concentrations of arsenic (0–50 mM arsenate or 0–45 mM arsenite) as described in the Manual of Antimicrobial Susceptibility Testing (Coyle, 2005) and grown at 70 °C for 18 h; for each determination, two independent experiments with triplicate samples were carried out. Minimum inhibitory concentration (MIC) was determined as the lowest concentration of arsenic

that completely inhibited the growth of the strain as evaluated by $OD_{600\ nm}$ after incubation for 18 h under optimal conditions. MIC definition is different from that previously described (Del Giudice *et al.*, 2013).

Strain genotypes and sources are summarized in Table S1.

DNA and RNA extraction

Thermus thermophilus HB27 genomic DNA was prepared following reported procedures (Pedone *et al.*, 2014). Total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany). The extracted RNA samples (20 μg) were then diluted to 0.2 mg ml^{-1} for DNase I treatment with the Ambion® TURBO™ DNase according to the manufacturer's instructions.

qRT-PCR

To determine whether the expression *of TtsmtB* and *TTC0354* genes was induced by arsenic, and to verify the expression of *TTC0354* and *TTC1502* (*TtArsC*) in the *ΔsmtB* strain, qRT-PCRs were performed using the StepONEPlus Real-Time PCR system (Applied Biosystems, Foster City, California, USA) and the SYBR Select Master Mix kit (Applied Biosystems). Total RNA extracted from *T. thermophilus* HB27 and *ΔsmtB* cells was digested with TURBO DNase, RNase-free (Contursi *et al.*, 2010). The cDNAs were synthetized using a mixture of the corresponding reverse primer (*0353rv*, *0354realrv* or *arsCrealrv*) and the 16S reverse primer (*16Sthrv*), used as internal control for normalization. The specific cDNAs synthesized were amplified using the following primers: *smtBrealfw* and *0353rv*; *0354realfw* and *0354realrv*; *arsCrealfw* and *arsCrealrv;* or *16Sthfw* and *16Sthrv* (Table S2) designed using Primer Express 2.0 software (ABI Biosystems), and amplified 107 bp, 89 bp and 100 bp specific products of *TtsmtB*, *TTC0354* and *TtarsC*, respectively. Reaction optimization was assessed for each template to generate standard curves and calculate the amplification efficiency.

For the amplification of the specific cDNAs, 25 ng from the RT-reaction mixture was used, whereas 0.002 ng was used to amplify the 16S fragment. DNA contamination was tested by the inclusion of a control without reverse transcriptase for each RNA sample. Two independent experiments were performed, and each sample was always tested in triplicate. The amplification data were analysed using the StepONE software (Applied Biosystems), and induction folds were calculated by the comparative Ct method. The relative expression ratio of the target gene, *TtsmtB* or *TTC0354*, versus that of the *16S* rRNA gene was calculated as described (Pfaffl, 2001).

Primer extension analysis of transcription start site

To determine the first transcribed nucleotides of *TtsmtB* and *TTC0354*, total RNA extracted from *T. thermophilus* HB27 cells was subjected to primer extension analysis as described (Fiorentino *et al.*, 2011), using the primers *0353pr(ext) rv 2 and 0354rv* (Table S2). The same primers were used to produce a sequence ladder using the Thermo Sequenase Cycle Sequencing Kit (Affymetrix, Santa Clara, California, USA) according to the manufacturer's instructions to locate the products on 6% urea polyacrylamide gels.

Cloning, expression and purification of TtSmtB

The gene encoding *Tt*SmtB was amplified by PCR from *T. thermophilus* HB27 genomic DNA, using Taq DNA polymerase (ThermoFisher Scientific, Waltham, Massachusetts, USA) and the primers containing the NdeI (*smtBfw*) and HindIII (*smtBrv*) sites at the 5′ and 3′ ends respectively. Amplified fragments were purified, digested and cloned into NdeI/HindIII-digested pET28b(+) vector (Novagen). For protein expression, *E. coli* BL21-CodonPlus (DE3)-RIL cells transformed with pET28/*TtsmtB* were grown in LB medium containing kanamycin (50 μg ml^{-1}), chloramphenicol (33 μg ml^{-1}) and 0.25 mM $ZnSO_4$. When the culture reached 0.7 $OD_{600\ nm}$, protein expression was induced by the addition of 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) and the bacterial culture was grown for 16 h at 22 °C. Cells were harvested and lysed by sonication in 50 mM Tris–HCl, pH 7, as described before (Fiorentino *et al.*, 2014). The recombinant protein was purified to homogeneity through a heat treatment of the cell extract (65 °C for 10 min) followed by HiTrap Heparin chromatography (5 ml; GE Healthcare) connected to an AKTA Explorer system (GE Healthcare). The fractions containing His-*Tt*SmtB were pooled, concentrated by ultrafiltration and dialysed for 16 h at 4 °C against 50 mM Tris–HCl, pH 7.0, and 0.15 M NaCl. To prevent protease activity, an inhibitor cocktail (Roche) was added at each step. The histidine tag was removed using 10 U of thrombin (Sigma, St. Louis, Missouri, USA) for 1 mg of His-*Tt*SmtB yielding *Tt*SmtB, a protein of 13.5 kDa. The purified proteins were stored in aliquots at −20 °C.

TtSmtB structural characterization

To determine the quaternary structure of *Tt*SmtB, the native molecular mass was determined by loading the purified protein to an analytical Superdex PC75 column (0.3 × 3.2 cm) in 50 mM Tris–HCl, pH 7, and 0.2 M NaCl. The column was calibrated using a set of gel filtration markers (low range; GE Healthcare),

including bovine serum albumin (67.0 kDa), ovalbumin (43.0 kDa), chymotrypsinogen A (25.0 kDa) and RNase A (13.7 kDa).

Determination of putative disulfide bonds was obtained by mass spectrometry through MALDI mapping after *Tt*SmtB carbamidomethylation with iodoacetamide, and trypsin and chymotrypsin digestion to detect carbamidomethyl-cysteine-containing peptides as described (Moinier *et al.*, 2014).

Analysis of secondary structure was performed by registering far-UV circular dichroism spectra in a Jasco J-815 CD spectrometer, equipped with a Peltier-type temperature control system (PTC-423S/15 model) using protein concentration of about 3 μM in a 25 mM Na-phosphate, pH 7.0 buffer. CD spectra were recorded as described (Prato *et al.*, 2008). Spectra were analysed for secondary structure amount according to the Selcon method using Dichroweb (Whitmore and Wallace, 2008). CD spectra were also registered to titrate *Tt*SmtB with increasing amounts of arsenate and arsenite: 6, 30, 60, 118, 230, 590, 890, 1200, 1500 μM. CD titration curves (obtained in triplicate) were fitted in Prism 6.0 using the equation for one binding site to determine the dissociation constants.

Electrophoretic mobility shift assay

To determine the binding of *Tt*SmtB to the putative promoter regions of *TtsmtB*, *TTC0354* and *TtarsC*, electrophoretic mobility shift assays (EMSAs) were performed. The promoter regions were amplified by PCR using specific primer pairs (one radiolabelled with γ^{32}P dATP and polynucleotide kinase): *0353pr(ext)fw* and *0353pr(ext)rv2*, *0354footprint fw* and *0354footprint rv*, *ArsCprfw* and *ArsCprrv* (Table S2) giving 149, 143 and 78 bp fragments respectively. EMSA reactions were set up as described (Fiorentino *et al.*, 2007), using 2.5 or 7.5 μM of proteins. Sequence-specific binding to *0354*p was evaluated in a competition assay using competitors at molar ratio of 1:200 and 1:400. As aspecific DNA, a 150 bp coding region from *Sulfolobus solfataricus* was amplified with *VP2 fw* and *VP2 rv* primers (Fusco *et al.*, 2013). To quantify the interaction between *Tt*SmtB/*TtsmtB*p and *Tt*SmtB/*0354*p, the DNA was incubated with increasing amounts of the protein (0–15 μM); the complexes were separated, and the gels were processed and analysed by phosphor imaging using Quantity One software (Bio-Rad) as already described (Fiorentino *et al.*, 2011).

To determine whether arsenite and arsenate were ligands for *Tt*SmtB and evaluate their possible effect on binding to the target *0354*p, 2.5 μM of protein was pre-incubated with $NaAsO_2$ or KH_2AsO_4 at molar ratio of 1:50 and 1:100 (considering *Tt*SmtB as a dimer).

DNase I footprinting

A probe containing the promoter region of *TTC0354* was produced by PCR using a combination of *0354footprint fw* and *0354footprint rv*; the latter was 5′ end labelled with T4 polynucleotide kinase and [γ-^{32}P] ATP. The labelled PCR product of 143 bp (about 40 nM) was incubated with 4 μg of pure *Tt*SmtB in binding buffer and digested with three units of DNase I (Roche) for 1 min at 37 °C. Subsequent steps were performed as described (Fiorentino *et al.*, 2011). Labelled primer was as also used to generate a dideoxynucleotide sequence ladder with Thermo Sequenase Cycle Sequencing Kit (Affymetrix) using 0.1 pmol of the same PCR fragment as the template and following the manufacturer's instructions.

Construction of T. thermophilus *mutants*

To obtain a *ΔsmtB::kat* deletion mutant of *T. thermophilus* HB27, the chromosomal *TtsmtB* gene was replaced with the kanamycin nucleotidyltrasferase gene (*kat*) cassette by double homologous recombination. Two regions upstream and downstream of *TtsmtB* (arm UP and arm DW) were amplified by PCR using *T. thermophilus* HB27 genomic DNA as template. For arm UP, the forward primer (UP fw SmtB *Eco*RI) and the reverse primer (New UP rv SmtB *Xba*I) contained EcoRI and XbaI sites respectively (Table S2). For arm DW, the forward primer (New DW fw SmtB *Xba*I) and the reverse primer (DW rv SmtB *Hind*III) contained XbaI and HindIII sites. The resulting products (909 bp arm UP and 1014 bp arm DW) were digested, purified, ligated (in 1:1 molar ratio) and cloned into pUC19, giving the pUC19*ΔsmtB* vector. The *kat* cassette, extracted from pUC19-*kat* after XbaI digestion, was inserted at the XbaI site of pUC19*ΔsmtB*. The resulting vector was named pUC19*ΔsmtB::kat*. The orientation of *kat* cassette was confirmed by restriction analysis. The pUC19Δ*smtB*::*kat* plasmid was HindIII-digested and used in linear form to transform *T. thermophilus* HB27 as described above.

The replacement of the *TtsmtB* gene was verified by PCR on the genomic DNA of the transformants; three primer sets (Table S2) were used: one pair (*0351promfw*/*0351promrv*) amplified a region in both deleted and wild-type strains; another one (*0352fw*/*0353rv*) amplified a region only in the wild-type strain; the last pair (*smtBfw*/*smtBrv*) amplified a fragment of 1125 bp corresponding to the *kat* gene in the mutant strain compared to the *TtsmtB* gene of 372 bp in the wild type. The *kat* insertion was further confirmed by DNA sequencing.

The *TTC0354* mutant was obtained by following a single-recombination strategy. For this, a 1407 bp internal fragment of the *TTC0354* gene was amplified with

primers *0354Eco* and *0354Hind* and further digested with EcoRI and HindIII restriction enzymes, which targets were included in the primer's sequence (Table S2). The fragment was subsequently cloned into the same sites of suicide vector pK18, conferring thermostable resistance to kanamycin, and the resulting plasmid pK18-*Δ0354* was used to transform *T. thermophilus* as described (Blesa *et al.*, 2017). Selection on kanamycin TM plates allowed the isolation of *TTC0354::pK18* mutants, in which only a C-terminal deletion form of the protein lacking 121 amino acid could be produced.

Acknowledgements

This study was supported by grants from the Regione Campania, legge 5 (Italy, CUP number E69D15002 10002) to SB, and from the project BIO2016-77031-R of the Spanish Ministry of Economy and Competitiveness to JB.

Conflict of interest

None declared.

References

Bartolucci, S., Contursi, P., Fiorentino, G., Limauro, D., and Pedone, E. (2013) Responding to toxic compounds: a genomic and functional overview of Archaea. *Front Biosci* **18:** 165–189.

Blesa, A., Baquedano, I., Quintáns, N.G., Mata, C.P., Castón, J.R., and Berenguer, J. (2017) The transjugation machinery of *Thermus thermophilus*: identification of TdtA, an ATPase involved in DNA donation. *PLoS Genet* **13:** e1006669.

Chauhan, S., Kumar, A., Singhal, A., Tyagi, J.S., and Prasad, H.K. (2009) CmtR, a cadmium-sensing ArsR-SmtB repressor, cooperatively interacts with multiple operator sites to autorepress its transcription in *Mycobacterium tuberculosis*. *FEBS J* **276:** 3428–3439.

Contursi, P., Cannio, R., and She, Q. (2010) Transcription termination in the plasmid/virus hybrid pSSVx from Sulfolobus islandicus. *Extremophiles* **14:** 453–463.

Contursi, P., Fusco, S., Limauro, D., and Fiorentino, G. (2013) Host and viral transcriptional regulators in *Sulfolobus*: an overview. *Extremophiles* **17:** 881–895.

Cook, W.J., Kar, S.R., Taylor, K.B., and Hall, L.M. (1998) Crystal structure of the cyanobacterial metallothionein repressor SmtB: a model for metalloregulatory proteins. *J Mol Biol* **275:** 337–346.

Coyle, M.B. (2005) Manual of Antimicrobial susceptibility testing. Washington, DC: American Society for Microbiology, pp. 53–62.

Del Giudice, I., Limauro, D., Pedone, E., Bartolucci, S., and Fiorentino, G. (2013) A novel arsenate reductase from the bacterium *Thermus thermophilus* HB27: its role in arsenic detoxification. *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1834:** 2071–2079.

Donahoe-Christiansen, J., D'Imperio, S., Jackson, C.R., Inskeep, W.P., and McDermott, T.R. (2004) Arsenite-oxidizing Hydrogenobaculum strain isolated from an acid-sulfate-chloride geothermal spring in Yellowstone National Park. *Appl Environ Microbiol* **70:** 1865–1868.

Fernandez, M., Udaondo, Z., Niqui, J.L., Duque, E., and Ramos, J.L. (2014) Synergic role of the two ars operons in arsenic tolerance in *Pseudomonas putida* KT2440. *Environ Microbiol Rep* **6:** 483–489.

Fernandez, M., Morel, B., Ramos, J.L., and Krell, T. (2016) Paralogous regulators ArsR1 and ArsR2 of *Pseudomonas putida* KT2440 as a basis for arsenic biosensor development. *Appl Environ Microbiol* **82:** 4133–4144.

Fiorentino, G., Ronca, R., Cannio, R., Rossi, M., and Bartolucci, S. (2007) MarR-like transcriptional regulator involved in detoxification of aromatic compounds in *Sulfolobus solfataricus*. *J Bacterial* **189:** 7351–7360.

Fiorentino, G., Del Giudice, I., Bartolucci, S., Durante, L., Martino, L., and 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.

Fiorentino, G., Del Giudice, I., Petraccone, L., Bartolucci, S., and Del Vecchio, P. (2014) Conformational stability and ligand binding properties of BldR, a member of the MarR family, from *Sulfolobus solfataricus*. *BBA-Proteins Proteom* **1844:** 1167–1172.

Fusco, S., She, Q., Bartolucci, S., and Contursi, P. (2013) T (lys), a newly identified Sulfolobus spindle-shaped virus 1 transcript expressed in the lysogenic state, encodes a DNA-binding protein interacting at the promoters of the early genes. *J Virol* **87:** 5926–5936.

Guerra, A.J., and Giedroc, D.P. (2012) Metal site occupancy and allosteric switching in bacterial metal sensor proteins. *Arch Biochem Biophtys* **519:** 210–222.

Indriolo, E., Na, G., Ellis, D., Salt, D.E., and Banks, J.A. (2010) A vacuolar arsenite transporter necessary for arsenic tolerance in the arsenic hyperaccumulating fern *Pteris vittata* is missing in flowering plants. *Plant Cell* **22:** 2045–2057.

Jacobson, T., Navarrete, C., Sharma, S.K., Sideri, T.C., Ibstedt, S., Priya, S., *et al.* (2012) Arsenite interferes with protein folding and triggers formation of protein aggregates in yeast. *J Cell Sci* **125:** 5073–5083.

Kamerlin, S.C.L., Sharma, P.K., Prasad, R.B., and Warshel, A. (2013) Why nature really chose phosphate. *Q Rev Biophs* **46:** 1–132.

Kar, S.R., Lebowitz, J., Blume, S., Taylor, K.B., and Hall, L.M. (2001) SmtB-DNA and protein-protein interactions in the formation of the cyanobacterial metallothionein repression complex: Zn^{2+} does not dissociate the protein-DNA complex in vitro. *Biochemistry* **40:** 15869.

Kulp, T.R., Hoeft, S.E., Asao, M., Madigan, M.T., Hollibaugh, J.T., Fisher, J.C., *et al.* (2008) Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science* **321:** 967–970.

Lin, Y.-F., Walmsley, A.R., and Rosen, B.P. (2006) An arsenic metallochaperone for an arsenic detoxification pump. *Proc Natl Acad Sci USA* **103:** 15617–15622.

van Lis, R., Nitschke, W., Duval, S., and Schoepp-Cothenet, B. (2013) Arsenics as bioenergetic substrates. *BBA-Bioenergetics* **1827:** 176–188.

Liu, Z.J., Boles, E., and Rosen, B.P. (2004) Arsenic trioxide uptake by hexose permeases in *Saccharomyces cerevisiae*. *J Biol Chem* **279:** 17312–17318.
Mandal, S., Chatterjee, S., Dam, B., Roy, P., and Das Gupta, S.K. (2007) The dimeric repressor SoxR binds cooperatively to the promoter(s) regulating expression of the sulfur oxidation (sox) operon of *Pseudaminobacter salicylatoxidans* KCT001. *Microbiol-Sgm* **153:** 80–91.
Meng, Y.L., Liu, Z.J., and Rosen, B.P. (2004) As(III) and Sb (III) uptake by G1pF and efflux by ArsB in *Escherichia coli*. *J Biol Chem* **279:** 18334–18341.
Miller, J.H. (1972) Experiments in molecular genetics. Cold Spring Harbor: Cold Spring Harbor Laboratory.
Moinier, D., Slyemi, D., Byrne, D., Lignon, S., Lebrun, R., Talla, E., and Bonnefoy, V. (2014) An ArsR/SmtB family member is involved in the regulation by arsenic of the arsenite oxidase operon in *Thiomonas arsenitoxydans*. *Appl Environ Microbiol* **80:** 6413–6426.
Morby, A.P., Turner, J.S., Huckle, J.W., and Robinson, N.J. (1993) SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene *smtA* – identification of a Zn inhibited DNA-protein complex. *Nucleic Acids Res* **21:** 921–925.
Ordonez, E., Thiyagarajan, S., Cook, J.D., Stemmler, T.L., Gil, J.A., Mateos, L.M., and Rosen, B.P. (2008) Evolution of metal(loid) binding sites in transcriptional regulators. *J Biol Chem* **283:** 25706–25714.
Osman, D., and Cavet, J.S. (2010) Bacterial metal-sensing proteins exemplified by ArsR-SmtB family repressors. *Nat Prod Rep* **27:** 668–680.
Pedone, E., Fiorentino, G., Pirone, L., Contursi, P., Bartolucci, S., and Limauro, D. (2014) Functional and structural characterization of protein disulfide oxidoreductase from *Thermus thermophilus* HB27. *Extremophiles* **18:** 723–731.
Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29:** e45.
Politi, J., Spadavecchia, J., Fiorentino, G., Antonucci, I., Casale, S. and 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.
Politi, J., Spadavecchia, J., Fiorentino, G., Antonucci, I. and 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:** pii: 20160629.
Prato, S., Vitale, R.M., Contursi, P., Lipps, G., Saviano, M., Rossi, M., and Bartolucci, S. (2008) Molecular modeling and functional characterization of the monomeric primase-polymerase domain from the *Sulfolobus solfataricus* plasmid pIT3. *FEBS J* **275:** 4389–4402.
Qin, J., Rosen, B.P., Zhang, Y., Wang, G.J., Franke, S., and Rensing, C. (2006) Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proc Natl Acad Sci USA* **103:** 2075–2080.
Rosen, B.P. (1999) Families of arsenic transporters. *Trends Microbiol* **7:** 207–212.
Rosen, B.P. (2002) Biochemistry of arsenic detoxification. *FEBS Lett* **529:** 86–92.
Schurig-Briccio, L.A., and Gennis, R.B. (2012) Characterization of the P-IB-Type ATPases present in *Thermus thermophilus*. *J Bacterial* **194:** 4107–4113.
Shi, W.P., Wu, J.H., and Rosen, B.P. (1994) Identification of a putative metal-binding site in a new family of metalloregulatory proteins. *J Biol Chem* **269:** 19826–19829.
Tawfik, D.S., and Viola, R.E. (2011) Arsenate replacing phosphate: alternative life chemistries and ion promiscuity. *Biochemistry* **50:** 1128–1134.
VanZile, M.L., Cosper, N.J., Scott, R.A., and Giedroc, D.P. (2000) The zinc metalloregulatory protein *Synechococcus* PCC7942 SmtB binds a single zinc ion per monomer with high affinity in a tetrahedral coordination geometry. *Biochemistry* **39:** 11818–11829.
Wang, G., Kennedy, S.P., Fasiludeen, S., Rensing, C., and DasSarma, S. (2004) Arsenic resistance in *Halobacterium* sp. strain NRC-1 examined by using an improved gene knockout system. *J Bacteriol* **186:** 3187–3194.
Whitmore, L., and Wallace, B.A. (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89:** 392–400.
Wu, J., and Rosen, B.P. (1991) The ArsR protein is a trans-acting regulatory protein. *Mol Microbiol* **5:** 1331–1336.
Wysocki, R., Clemens, S., Augustyniak, D., Golik, P., Maciaszczyk, E., Tamas, M.J., and Dziadkowiec, D. (2003) Metalloid tolerance based on phytochelatins is not functionally equivalent to the arsenite transporter Acr3p. *Biochem Bioph Res Com* **304:** 293–300.
Yang, H.C., and Rosen, B.P. (2016) New mechanisms of bacterial arsenic resistance. *Biomed J* **39:** 5–13.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Table S1. Strains used in this work classified according to their genotype.

Table S2. Oligonucleotides used in this work classified according to their purpose.

Fig. S1. (A) Multiple sequence alignment by Clustal W of TtSmtB with SmtB/ArsR members. Sequences of the protein used and percentages of identity are: ArsR of T. thermophilus SG0.5JP17-16, 87%; ArsR of *T. oshimai* JL-2, 81%; SmtB of *T. scotoductus* SA-01, 46%; CadC of Clostridium perfrigens, 34%; ZiaR of Synechocystis PCC6803/KAZUSA, 46%; SmtB of Synechococcus PCC7942, 50% and ArsR of *E. coli*, 41%. The ELCVCD motif is highlighted by a red box. HTH domain is indicated. Cys 10 is indicated by a green arrow. Red arrows indicate conserved Cys 62 and Cys 64. Blue arrows indicate residues putatively involved in ligand binding. The secondary structure elements of TtSmtB are depicted above the sequences. (B) Structural model predicted for TtSmtB dimer. (C) Structural model predicted for TTC0354. Modeling of the structures were made on the basis of amino acid sequence by the software I-TASSER.

Fig. S2. Generation of smtB::kat mutant.

Fig. S3. Generation of TTC0354 single recombination mutants.

Fig. S4. TtSmtB interaction with target promoter.

Table S1. Strains used in this work classified according to their genotype.

Strain	Genotype	Source
T. thermophilus HB27	Wild type	DSMZ
T. thermophilus ΔsmtB::*kat*	*T. thermophilus* HB27 deletion mutant of the *TtsmtB* gene, Kan^r	This study
T. thermophilus TTC0354::pK18	*T. thermophilus* HB27 insertion mutant defective in *TTC0354,* Kan^r	This study
E. coli TOP10F'	F'{ $lacI^qTn10$ (Tet^R)} mcrAΔ(mrr-hsdRMSmcrBC) ϕ80ΔlacZΔM15 Δ lacX74 recA1 deoR araD139 Δ(ara-leu) 7697galU galKrpsL(Str^R) endA1 nupG	Invitrogen
E. coli BL21-Codon Plus (DE3) RIL	F– ompThsdS(r_B– m_B–) dcm+ Tet^r gal λ (DE3) endAHte [argUileYleuWCam^r]	Stratagene

Table S2. Oligonucleotides used in this work classified according to their purpose.

Primer names	**Target region**	**Sequences** (the restriction sites are underlined) 5'-3'
qRT-PCR		
0354realfw	*TTC0354*	GACCTCTGCTTTGGCTTTG
0354realrv	*TTC0354*	CTACCTGCCAACTCCTCCA
16Sthfw	*Tt16S*	TAGTCCACGCCCTAAACGAT
16Sthrv	*Tt16S*	CCTTTGAGTTTCAGCCTTGC
smtBrealfw	*TtsmtB*	GAGAGGGTGGTCAAGGAG
0353rv	*TtsmtB*	TCGCAGACGCAAAGCTCC
ArsCrealfw	*TtarsC*	GGAAACCCCTGGAGGAGTG
ArsCrealrv	*TtarsC*	TCGTCGCTGGGAAGCCTTC
Construction of *T.thermophilus ΔsmtB::kat and T.thermophilus ΔTTC0354*		
UP fw SmtB EcoRI	arm UP	AAGAATTCTGGAGGATGGAGGGC
NewUPrvSmtBXbaI	arm UP	AAATCTAGAGCAGACCGCCCTAC
NewDWfwSmtBXbaI	arm DW	AAATCTAGATACTACCGCCTCGCC
DW rv SmtB HindIII	arm DW	AAAAAGCTTAAGGCATCCACCACC
0351promfw	0351prom	GAAGCTTGAAGGGGGCTC
0351promrv	0351prom	GCGAGCCACAATACC
0354Eco	0354 internal fw	AAAGAATTCCTCGCCTGGGTTTGGGA
0354Hind	0354 internal rv	AAAAAGCTTGTCCACGGGAGAAAGGCCA

Cloning of the *TtsmtB* gene in pET plasmid		
smtBfw	*TtsmtB*	GTCCAAGGAGGAGGAAACATATGCCAAGCGGGG
smtBrv	*TtsmtB*	GCATCATTTGAGCAAGCTTTCAAGTGTTTTCTTCCGC
EMSA		
0353pr(ext) fw	0353prom	CATGGTGGGCATCAACCT
0353pr(ext) rv2	0353prom	CTCCTTGACCACCCTCTCCT
0354footprint fw	0354prom	CGCCTCGCCGACCGGCA
0354footprint rv	0354prom	CCATGCCCTCTACCCGGAAG
ArsCpr fw	ArsCprom	GGTGGCCCAGCTTGAGCG
ArsCpr rv	ArsCprom	AGAGGACGAGGACCCGCAT
Primer extension		
0353pr(ext) rv 2	*TtsmtB*	CTCCTTGACCACCCTCTCCT
0354 rv	*TTC0354*	CTCGCGTAGTTCACCTGGAC
Footprinting		
0354footprint fw	0354 prom	CGCCTCGCCGACCGGCA
0354footprint rv	0354 prom	CCATGCCCTCTACCCGGAAG

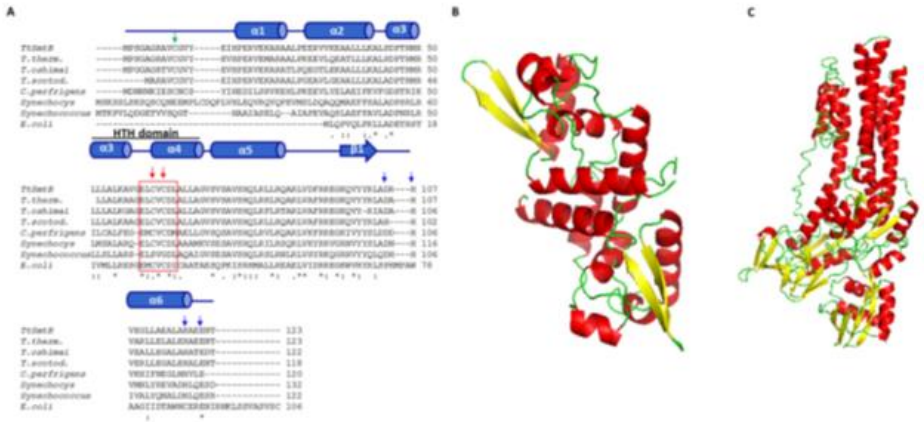


Fig. S1.

A. Multiple sequence alignment by Clustal W of *Tt*SmtB with SmtB/ArsR members. Sequences of the protein used and percentages of identity are: ArsR of *T. thermophilus* SG0.5JP17-16, 87%; ArsR of *T. oshimai* JL-2, 81%; SmtB of *T. scotoductus* SA-01, 46%; CadC of *Clostridium perfrigens*, 34%; ZiaR of *Synechocystis* PCC6803/KAZUSA, 46%; SmtB of *Synechococcus* PCC7942, 50% and ArsR of *E. coli,* 41%. The ELCVCD motif is highlighted by a red box. HTH domain is indicated. Cys 10 is indicated by a green arrow. Red arrows indicate conserved Cys 62 and Cys 64. Blue arrows indicate residues putatively involved in ligand binding. The secondary structure elements of *Tt*SmtB are depicted above the sequences.

B. Structural model predicted for *Tt*SmtB dimer.

C. Structural model predicted for TTC0354.

Modeling of the structures were made on the basis of amino acid sequence by the software I-TASSER.

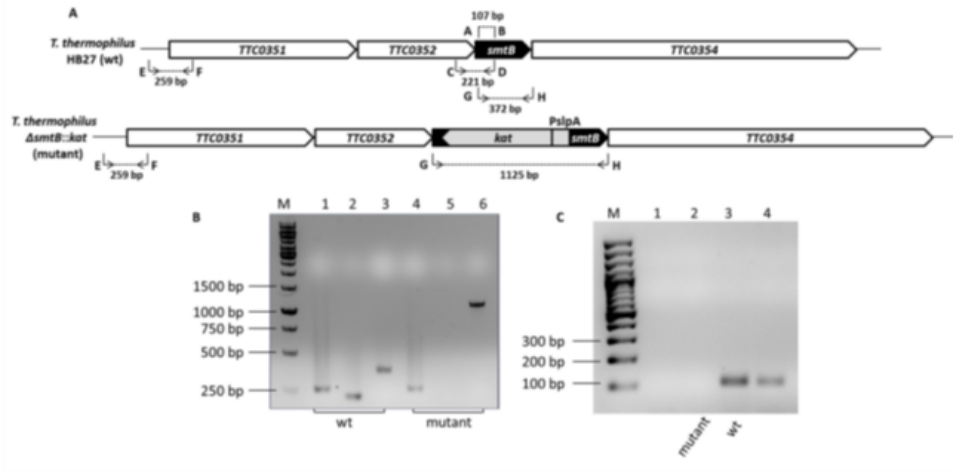


Fig. S2. Generation of *ΔsmtB::kat* mutant.

A. Scheme showing *TtsmtB* and its replacement with the *kat* cassette. Arrows indicate annealing positions and orientation of the primers used to confirm the insertion (Table S2).

B. PCR analysis. Lanes 1 to 3 correspond to DNA of the wild type strain (wt) and lanes 4 to 6 to the mutant. Amplifications were carried out to generate EF (lanes 1,4), CD (lanes 2, 5), and GH fragments (lanes 3,6). Lane M: 1Kb DNA ladder.

C. RT-PCR analysis of the *TtsmtB* gene; lane 1 negative control; lane 2: RT-PCR on *TtsmtB* cDNA from *T. thermophilus ΔsmtB*::*kat*; lane 3: RT-PCR on *TtsmtB* cDNA from the wild type parental strain; lane 4) PCR on genomic DNA of the wild type parental strain. Lane M: 100bp DNA ladder.

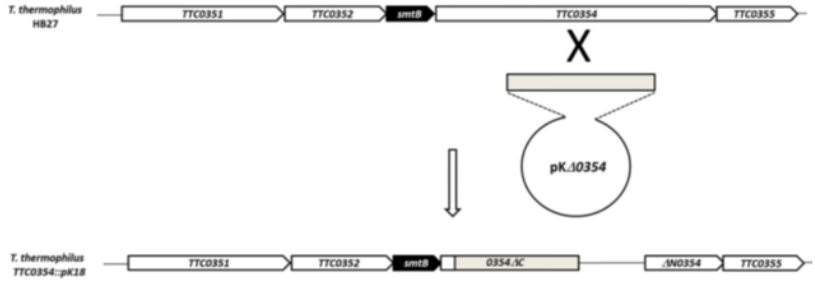


Fig. S3. Generation of *TTC0354* single recombination mutants.

An internal fragment to the targeted gene (*TTC0354*) is cloned into the suicide vector pK18 (containing the kanamycin resistance cassette) to generate pKΔ*0354*. The plasmid is inserted in *T. thermophilus* by single recombination, generating a non functional gene deleted at its C-terminus.

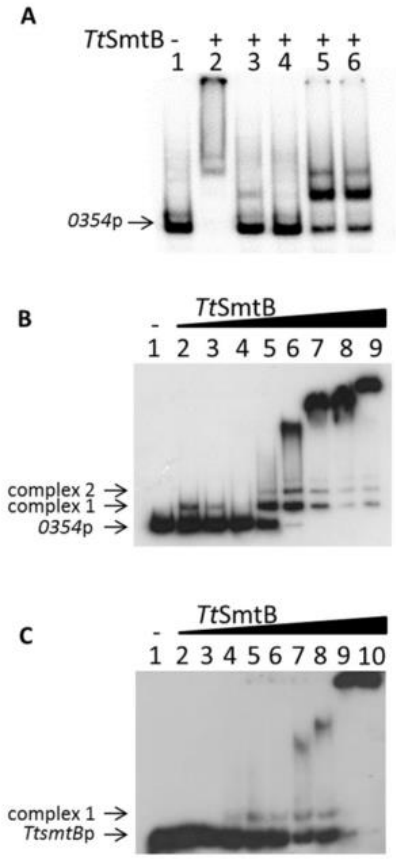


Fig. S4. *Tt*SmtB interaction with target promoter.
A. EMSA competition assays: EMSA with 2.5 µM *Tt*SmtB on *0354*p promoter (1 nM) in the absence (lane 2) or in the presence of unlabeled DNA competitor, either specific (same as the probe, lanes 3,4) or unspecific (lanes 5,6) in a molar ratio of 1:200 and 1:400.B. Binding of *Tt*SmtB to *0354*p: protein concentrations (µM) are: 0 (lane 1), 0.1 (lane 2), 0.2 (lane 3), 0.4 (lane 4), 1.2 (lane 5), 2.5 (lane 6), 5.0 (lane 7), 10.0 (lane 8), 15.0 (lane 9).
C. Binding of *Tt*SmtB to *TtsmtB*p: protein concentrations (µM) are: 0 (lane 1), 0.5 (lane 2), 0.1 (lane 3), 1.5 (lane 4), 2.0 (lane 5), 2.5 (lane 6), 3.7 (lane 7), 5.0 (lane 8), 10.0 (lane 9), 15.0 (lane 10).

International Journal of Biological Macromolecules 138 (2019) 1056–1063

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: http://www.elsevier.com/locate/ijbiomac

A physicochemical investigation on the metal binding properties of *Tt*SmtB, a thermophilic member of the ArsR/SmtB transcription factor family

Giovanni Gallo [a], Immacolata Antonucci [a], Luciano Pirone [b], Angela Amoresano [c], Patrizia Contursi [a], Danila Limauro [a], Emilia Pedone [b], Simonetta Bartolucci [a], Gabriella Fiorentino [a,*]

[a] Department of Biology, University of Naples Federico II, Napoli, Italy
[b] Institute of Biostructure and Bioimaging, CNR, Napoli, Italy
[c] Department of Chemical Sciences, University of Naples Federico II, Napoli, Italy

ARTICLE INFO

Article history:
Received 6 June 2019
Received in revised form 12 July 2019
Accepted 25 July 2019
Available online 26 July 2019

Keywords:
Thermostability
Transcriptional repressor
Metal-binding site
Thermus thermophilus
ArsR SmtB family

ABSTRACT

The transcription factors of the ArsR/SmtB family are widespread within the bacterial and archaeal kingdoms. They are transcriptional repressors able to sense a variety of metals and undergo allosteric conformational changes upon metal binding, resulting in derepression of genes involved in detoxification. So far, the molecular determinants of specificity, selectivity, and metal binding mechanism have been scarcely investigated in thermophilic microorganisms. *Tt*SmtB, the only ArsR/SmtB member present in the genome of *Thermus thermophilus* HB27, was chosen as a model to shed light into such molecular mechanisms at high temperature. In the present study, using a multidisciplinary approach, a structural and functional characterization of the protein was performed focusing on its metal interaction and chemical-physical stability. Our data demonstrate that *Tt*SmtB has two distinct metal binding sites per monomer and interacts with di-tri-penta-valent ions with different affinity. Detailed knowledge at molecular level of protein-metal interaction is remarkable to design metal binding domains as scaffolds in metal-based therapies as well as in metal biorecovery or biosensing in the environment.

1. Introduction

The physiological concentration of metal ions in a cell is maintained by a coordinated action of intracellular metal-sensing and membrane transporter proteins that play crucial roles in discriminating metal ions, controlling the cellular levels of essential metals, and exporting/detoxifying metal(loid)s [1]. Notably, metal sensors are not totally selective for their cognate metal [2].

Among families of metal-sensor proteins, those belonging to the ArsR/SmtB comprehend a group of prokaryotic transcription factors able to modulate the efflux of a wide variety of metal ions [3]. They are widespread homodimeric proteins with a typical structural folding (α1-α2-α3-α4-β1-β2-α5) containing a helix-turn-helix DNA binding domain (α3-turn-α4) interacting with the operator/promoter of target genes [4]. Upon sensing metal ions such as zinc, cadmium, arsenic, antimony *etc.*, these transcriptional repressors dissociate from DNA promoting the expression of genes encoding proteins that expel, chelate, or change the oxidation state of the metal ions, thus reducing cellular toxicity. A large number of metal-responsive transcriptional regulatory proteins has been described, like SmtB, in *Synechococcus* species [5] and ArsR in *Escherichia coli* [6]. Besides sharing many common features, ArsR/SmtB members display great diversity in metal-sensing motifs as well as in the number of metals binding sites, determining a big variation in metal selectivity and metal coordination geometry. This is consistent with the hypothesis that metal binding sites in DNA binding proteins have evolved convergently in response to environmental pressure [7–9].

In a recent work, the abundance of ArsR/SmtB proteins in the genomes of many archaea and bacteria has been correlated to their pivotal role in fighting metal stress. Moreover, based on the locations of sensing amino acids within known secondary structure elements, a multiplicity of metal sensory motifs has been identified, further expanding knowledge of metal binding sites, mechanisms and evolutionary relationships [10].

Among environmental microorganisms, those with an optimum growth temperature higher than 60 °C are designated as thermophiles and their natural niches are mostly terrestrial volcanic sites, hot springs, solfataric fields or submarine hydrothermal places like submarine volcanoes, fumaroles and vents [11,12]. Thermophilic microorganisms have been of interest to many protein biochemists and molecular biologists over the years, both to understand the molecular basis of

* Corresponding author.
E-mail address: fiogabri@unina.it (G. Fiorentino).

https://doi.org/10.1016/j.ijbiomac.2019.07.174
0141-8130/

adaptation to extreme conditions [13–17] and as a reservoir of biocatalysts and bioactive molecules [18,19]; in fact, thermophilic proteins not only are more stable to high temperature than their mesophilic counterparts, but also possess further advantages such as higher solvent tolerance and stability to denaturing agents [14,20–23].

The majority of thermophilic habitats also exhibits great metal and metalloid ion concentrations [24,25]. As part of adaptation to such environments, thermophilic microbes have evolved peculiar metabolisms and molecular mechanisms to utilize and control the intracellular concentration of highly soluble toxic metal ions [26–28]. In particular, it has been proposed that microorganisms thriving in harsh conditions include more genes encoding ArsR/SmtB proteins or ArsR/SmtB members with multiple metal binding motifs in comparison to some pathogenic microorganisms that are typically endowed with many representatives in their genomes [29].

The first thermophilic member of the ArsR/SmtB family to be characterized is *Tt*SmtB, discovered in the bacterium *Thermus thermophilus* HB27 sharing 50% of identity with SmtB of *Synechococcus elongatus* PCC7942 [30]. It is a dimeric protein binding to a consensus regulatory sequence located upstream of the gene encoding the membrane ATPase *Tt*ArsX, involved in metal efflux outside the cell [27]. *Tt*SmtB also regulates the expression of *TtarsC*, a gene encoding an arsenate reductase involved in the process of metal ion detoxification [31,32]. Likewise its mesophilic counterparts, *Tt*SmtB is a repressor *in vivo* and its binding to target regulatory regions is hindered *in vitro* upon interaction with As(V), As(III) and Cd(II) [27,33].

In this work, to give insight into the molecular mechanisms of metal binding by ArsR/SmtB members at high temperature, we performed a structural and functional characterization of *Tt*SmtB by multidisciplinary approaches, focusing on its metal interaction and chemical-physical stability.

2. Methodology

2.1. Purification of TtSmtB

Recombinant *Tt*SmtB was purified to homogeneity from BL21-Codon Plus(DE3)-RIL cells transformed with pET28b(+)*Tt*SmtB grown with supplemented ($ZnSO_4$) using the procedure already described, consisting of a thermo-precipitation of the *E. coli* cell extract followed by an HiTrap Heparin chromatography [27]. To prevent protease activity and the oxidation of *Tt*SmtB an inhibitor cocktail (Roche) and 1 mM DTT were added respectively in all the buffers at each step and the solutions were nitrogen flushed for a few minutes prior to use. The histidine tag was removed from purified *Tt*SmtB by thrombin digestion (Sigma-Aldrich) and the complete removal was assessed by Western blotting using anti-poly-Histidine-Peroxidase antibodies (Sigma-Aldrich).

The concentration of the purified protein was determined by UV measurements using a theoretical, sequence-based extinction coefficient of 4595 M^{-1} cm^{-1} at 280 nm calculated on the dimeric protein [34]. *Tt*smtB was stored in aliquots at −20 °C.

2.2. In silico studies

BlastP analyses were performed using the Blast program (https://blast.ncbi.nlm.nih.gov/Blast.cgi) [35]. Multiple sequence alignments were performed with CLC Sequence viewer 8 (https://www.qiagenbioinformatics.com/products/clc-sequence-viewer/). 3D model predictions of *Tt*SmtB were generated through I-TASSER web server (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) [36] using as input the complete primary structure of *Tt*SmtB. The model was selected according to its similarity to the available crystal structure of SmtB from *Synechococcus elongatus* (PDB 1R22) [30].

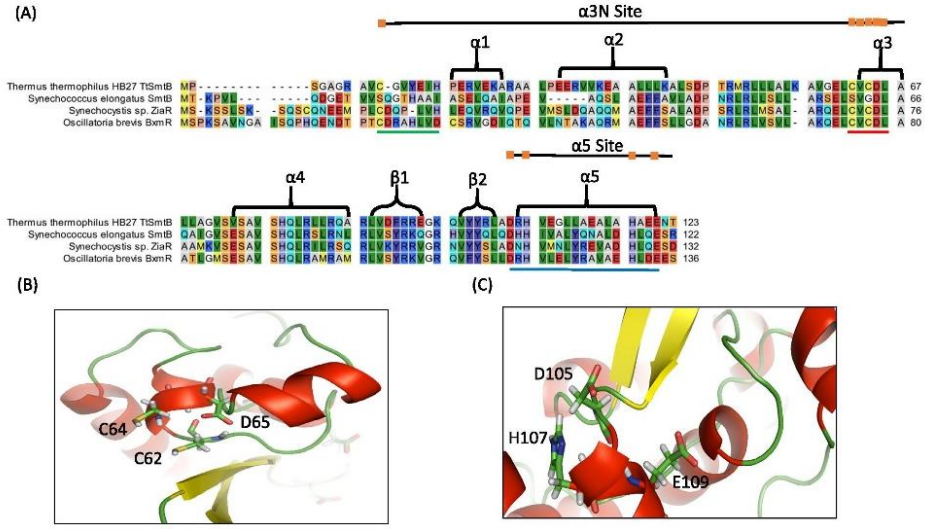


Fig. 1. A) Multiple sequence alignment of functionally characterized SmtB/ArsR proteins: ZiaR and BxmR belonging to the MX subgroup, and SmtB from *Synechococcus elongatus* belonging to the α5 subgroup (as defined in Roy et al. [10]). *Tt*SmtB α3N and α5 metal binding sites are shown together with secondary sequence designation above the alignment; the invariant amino acids within each motif are indicated by orange squares. Sequences Cx_5H (at N-terminal), CxC (α3 helix) and $DxHx_{10}Hx_2H$ (α5 helix) are underlined. The DNA recognition helix containing conserved VSHQLR sequence is α4. B) Molecular docking of putative *Tt*SmtB Cd(II) binding site showing coordination by Cys 62, Cys 64 and Asp65 of the CVCDL motif. C) Molecular docking of putative *Tt*SmtB Zn(II) binding site within the α5 helix showing metal coordination by Asp 105, His 107 and Glu 109.

Molecular docking calculations were carried out with Hex 8.0 CUDA available at http://hex.loria.fr/ [37] using a dimeric model of *Tt*SmtB, Zn(II) and Cd(II); 100 rigid body docking solutions were generated and the 10 best solutions were refined by energy minimization. The proposed model corresponds to the structure with the smallest distance between the metal and amino acids, after energy minimization step.

2.3. Intrinsic fluorescence spectroscopy

Measurements were made on a Cary Eclipse Fluorescence spectrophotometer at 25 °C, using an excitation wavelength of 225 nm. Tyrosine emission spectra were recorded from 250 to 400 nm. Spectra were taken in 1-nm increments with 1 s of integration per increment and slit widths of 5 nm. Protein samples of 0.62 μM were prepared in

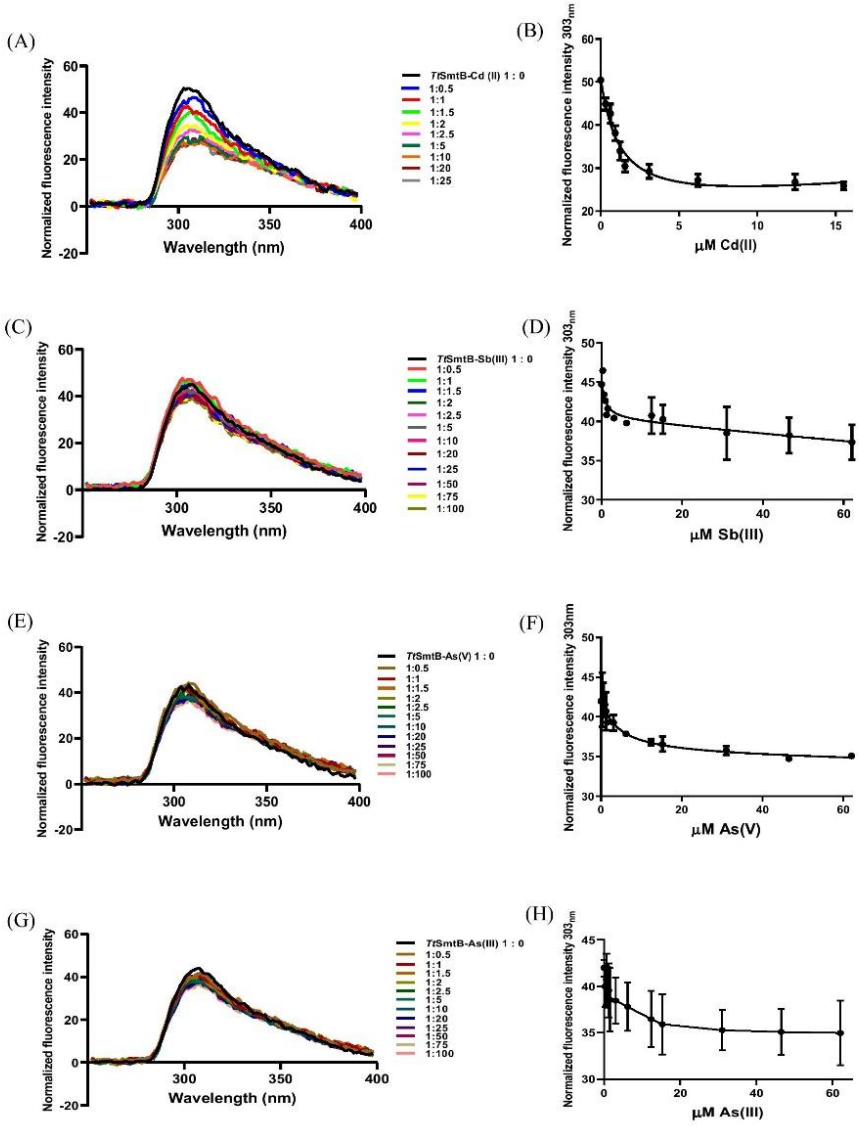


Fig. 2. Intrinsic fluorescence analysis of *Tt*SmtB binding to Cd(II), Sb(III), As(V) and As(III). A), C), E), G). Spectra measured at increasing concentrations of $CdCl_2$, $SbCl_3$, Na_2HAsO_4, $NaAsO_2$, respectively. B), D), F), H). Decrease in maximum fluorescence intensity as a function of $CdCl_2$, $SbCl_3$, Na_2HAsO_4, and $NaAsO_2$ concentration.

50 mM Tris-HCl pH 7.0, 1 mM DTT and titrations were conducted with freshly prepared solutions of $CdCl_2$, $SbCl_3$, $NaAsO_2$, and Na_2HAsO_4 added in increasing concentrations from 0.31 μM up to 15.5 μM (for $CdCl_2$) or up to 62 μM (for $SbCl_3$, $NaAsO_2$, Na_2HAsO_4) corresponding to 1:0.5 up to 1:100 protein-metal molar ratio. As a control bovine serum albumin (BSA, Sigma-Aldrich) at 5 μM was incubated with different concentrations of $CdCl_2$ using the same protein-metal molar ratio. Quenching of the tyrosine emission was performed by successive additions of the ligands to the cuvette, followed by gentle sample mixing and spectrum recording. Before the analysis of the quenching data, the fluorescence intensities were corrected for absorption of the exciting light and reabsorption of the emitted light to decrease the inner filter effect by using the equation $Fcorr = Fobs \times 10^{(Aexc+Aem)/2}$, where F_{corr} is the corrected fluorescence value, F_{obs} is the measured fluorescence value, A_{exc} is the absorption value at the excitation wavelength, and A_{em} is the absorption value at the emission wavelength [38]. Fluorescence titration curves were fitted in GraphPad Prism 7.0 by using the equation for one binding site. All the experiments were performed in triplicate.

2.4. Circular dichroism

CD spectra were recorded in a Jasco J-815 CD spectrometer, at 25 °C using protein samples of 3 μM in 25 mM Na-Phosphate, pH 7.0 1 mM DTT and cells of 0.1 path length [39]. CD values are reported as molar ellipticity per mean residue, $[\theta]$ in deg cm^2 $dmol^{-1}$, calculated from the equation: $[\theta] = [\theta]_{obs}\, mrw / 10 \times l \times C$; where $[\theta]_{obs}$ is the ellipticity measured in degrees, *mrw* is the mean residue molecular weight of the protein (27017.58 Da) for *Tt*SmtB, *C* is the protein concentration in g/mL and *l* is the optical path length of the cell in cm. To analyze whether interaction between *Tt*SmtB and Cd(II) or Sb(III) ions had an effect on *Tt*SmtB secondary structure, titrations were conducted in duplicate with freshly prepared solutions of $CdCl_2$ or $SbCl_3$ and added in increasing concentrations from 1.5 μM up to 150 μM corresponding to 1:0.5 up to 1:50 protein-metal molar ratio for $CdCl_2$ and $SbCl_3$. All the data were corrected for the baseline contribution of the buffer and the metal under all conditions. Spectra were analyzed for secondary structure amount according to the Selcon 3 method [40,41] using DichroWeb analysis [42,43].

To evaluate *Tt*SmtB stability to chemical denaturation, protein samples were incubated at 4 °C for 16 h with increasing concentrations of guanidinium chloride (GuHCl) from 1 M to 7 M and then analyzed recording the spectra. For each concentration, two independent experiments were performed. To examine whether the interaction between *Tt*SmtB and Cd(II), Sb(III), As(V) and As(III) had an effect on protein stability the following experiment was set-up: 3 μM *Tt*SmtB was incubated at 4 °C for 16 h with 3 M GuHCl and 150 μM $CdCl_2$, $SbCl_3$, Na_2HAsO_4 or $NaAsO_2$. Thermal curves were recorded in the temperature mode, by following the change in the CD signal at 222 nm with a scan rate of 1.0 °C/min (range of temperature from 30 °C to 105 °C). Unfolding reversibility was tested by lowering the temperature after denaturation and registering CD spectra. The thermal denaturation curves were obtained by plotting the molar ellipticity per mean residue at 222 nm in function of temperature and fitted to an equation for a two-state transition to identify the T_m (midpoint temperature of the unfolding transition) as described before [39].

2.5. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) studies were performed at 20 °C with an ITC200 calorimeter (MicroCal/GE Healthcare). $CdCl_2$ (150 μM) was titrated into a solution of *Tt*SmtB (10 μM) in 50 mM Tris pH 7.0, 150 mM NaCl. Data were analyzed and fitted using the Microcal Origin version 7.0 software package. Dissociation constants and stoichiometry were determined by fitting the data using a one-set-of-site binding model. ITC runs were repeated twice to evaluate the reproducibility of the results.

2.6. Electrophoretic mobility shift assays (EMSA)

To evaluate the effect of Sb(III) or Zn(II) on *Tt*SmtB/DNA complex formation EMSA assays were performed following procedures described elsewhere [44]. Specifically, the protein (5 μM) was tested for its binding to the regulatory region upstream of *TtarsX* (*TtarsXp*) that is the gene encoding the arsenic and cadmium membrane ATPase exporter, in the presence or not of Sb(III) or Zn(II) at a molar ratio of 1:20 or 1:50 (considering *Tt*SmtB as a dimer) and then incubated with 50 ng of *TtarsXp*, the 143 bp promoter region amplified by PCR using specific primers [33].

2.7. Inductively coupled plasma mass spectrometry (ICP-MS)

*Tt*SmtB (3,7 μM) was dissolved in 2% analytical grade HNO_3 acid for analysis. ICP-MS analysis was performed with an Agilent 7700 ICP-MS instrument (Agilent Technologies) equipped with a frequency-matching radio frequency (RF) generator and 3rd generation Octopole Reaction System (ORS3), operating with helium gas in ORF. ICP-MS data were collected to evaluate the amount of Zn(II) encapsulated within the cage. The following parameters were used: RF power: 1550 W, plasma gas flow: 14 L min^{-1}; carrier gas flow: 0.99 L min^{-1}; He gas flow: 4.3 mL min^{-1}. 103 Rh was used as an internal standard (final concentration: 50 μg L^{-1}).

3. Results and discussion

3.1. In silico analysis

Detailed analysis of its primary structure clearly showed that *Tt*SmtB belongs to the "mixed" (MX) group according to a recent categorization of ArsR/SmtB metalloproteins [10]. Indeed, it possesses conserved α3N-α5 sequence signatures encompassing two pairs of metal binding sites per dimer. Previously characterized proteins of this group are the cyanobacterial proteins BxmR from *Oscillatoria brevis* [45] and ZiaR from *Synechocystis* sp. [46]. These mesophilic proteins bind monovalent

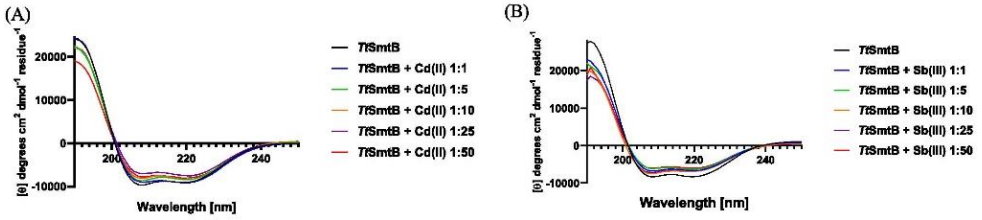


Fig. 3. A. Far-*UV CD* analysis of *Tt*SmtB binding to Cd(II) and Sb(III). A) Spectra measured at increasing concentrations of $CdCl_2$. B) Spectra measured at increasing concentrations of $SbCl_3$.

or divalent ions in trigonal and tetrahedral symmetry such as Ag (I), Cd(II) at the α3N site, and smaller metal ions, primarily Zn (II), at the second C-terminal α5 site; playing a structural or a regulatory role [47]. *Tt*SmtB first metal binding site, Cx_5H in amino-terminal and CxC in α3 helix, is composed of a cysteine residue, Cys10, at the N-terminus and a Cys62/Cys64 pair in the α3 helix, whereas the second metal site ($DxHx_{10}Hx_2H$) is located within the α5 helix and is devoid of thiolate ligands (Fig. 1A). Docking simulations highlighted that Cd(II) can be coordinated by thiol groups of Cys62/Cys64 and Asp 65 in the α3 helix (Fig. 1B), whereas Zn(II) preferentially interacts with histidine and carboxylate ligands (Asp 105, His 107 and Glu 109) derived from the C-terminal α5 helices on opposite subunits at the dimer interface (Fig. 1C). *Tt*SmtB DNA binding domain includes the α3 and α4 helices as part of the projected helix-turn-helix (α3-turn-α4) DNA binding motif with α4 being the recognition helix (Fig. 1A); the sequence VSHQLR in the DNA recognition helix is also highly conserved throughout the SmtB/ArsR family and is, in fact, another distinguishing feature that defines membership [3]. The *in silico* analysis indicates the formation of two distinct metal coordination complexes, one at the α3N site, which could be responsible for allosteric metalloregulation, and the other at α5 site, which could preferentially bind zinc ions [48].

Some experimental observations support our hypothesis that Zn(II) could play a stabilizing rather than a regulatory role: a) we failed to obtain a soluble apo form of *Tt*SmtB after attempting dialysis in the presence of different concentrations of EDTA (from 0.5 mM to 5 mM); even the small amount of protein which did not precipitate contained zinc, as assessed through ICP-MS. b) The yield of soluble recombinant *Tt*SmtB greatly increased when *E. coli* cells were grown in the presence of $ZnSO_4$. c) ICP-MS analysis performed on the purified recombinant protein revealed the presence of Zn(II). d) Mobility shift assay indicated that *Tt*SmtB does not dissociate from the cognate DNA in the presence of increasing concentration of Zn(II) (see Fig. 6B).

Therefore, all the subsequent characterizations described were performed with a Zn-bound protein.

Table 1
Deconvolution of CD spectra for analysis of *Tt*SmtB secondary structure in the presence of increasing concentration of: A) Cd(II); B) Sb(III); C) As(V) and D) As(III) using Selcon3 algorithm [40,41] and DichroWeb [42,43].

	α helix (%)	β sheet (%)	Turn (%)	Random (%)
A)				
*Tt*SmtB	32,8	17,2	20,2	29,8
*Tt*SmtB-Cd(II) 1:1	32,5	17,7	19,9	29,9
*Tt*SmtB-Cd(II) 1:5	30,9	18,6	20,4	30,1
*Tt*SmtB-Cd(II) 1:10	27,6	19,8	22,1	30,5
*Tt*SmtB-Cd(II) 1:25	22,7	21,6	23,9	31,8
B)				
*Tt*SmtB	32,8	17,2	20,2	29,8
*Tt*SmtB-Sb(III) 1:1	30,2	18,2	19,9	30,5
*Tt*SmtB-Sb(III) 1:5	33,4	18,0	18,5	30,5
*Tt*SmtB-Sb(III) 1:10	27,7	17,6	18,1	35,9
*Tt*SmtB-Sb(III) 1:25	30,0	18,3	21,2	31,1
C)				
*Tt*SmtB	32,8	17,2	20,0	29,8
*Tt*SmtB-As(V) 1:1	31,0	17,4	20,6	29,8
*Tt*SmtB-As(V) 1:5	28,8	18,0	20,4	30,0
*Tt*SmtB-As(V) 1:10	23,9	19,6	22,2	31,7
*Tt*SmtB-As(V) 1:20	23,0	21,5	22,5	31,2
D)				
*Tt*SmtB	32,8	17,2	20,0	29,8
*Tt*SmtB-As(III) 1:1	28,0	18,6	21,7	30,2
*Tt*SmtB-As(III) 1:5	27,6	18,4	21,6	30,7
*Tt*SmtB-As(III) 1:10	24,2	20,1	21,7	31,6
*Tt*SmtB-As(III) 1:20	22,9	20,7	21,8	32,4

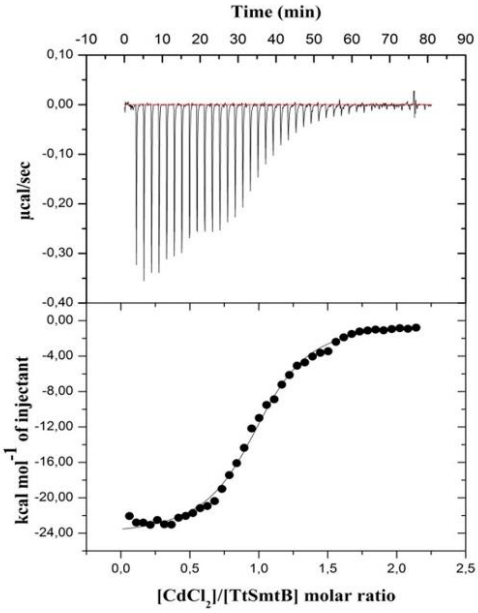


Fig. 4. ITC analysis of *Tt*SmtB binding to Cd(II). Top and bottom panel report raw and integrated data, respectively.

3.2. Fluorescence analysis of *TtSmtB* binding to metal ions

In previous papers we demonstrated the involvement of the transcription factor *Tt*SmtB in mediating the response to As(III), As(V) and Cd(II) *in vivo* [27,33]. In this work, to get insight into the structural determinants of metal specificity, we analyzed conformational changes of *Tt*SmtB upon Cd(II), As(V), As(III) and Sb(III) binding through intrinsic fluorescence spectroscopy, taking advantage of three tyrosine residues located in proximity of the metal binding motifs (Tyr 13 is within the α3N metal binding site whereas Tyr 100 and Tyr 101 are close to α5 metal binding site, see Fig. 1) and considering that *Tt*SmtB does not contain any Trp residue. Sb(III) was included in this analysis because some ArsR/SmtB family members are known to regulate

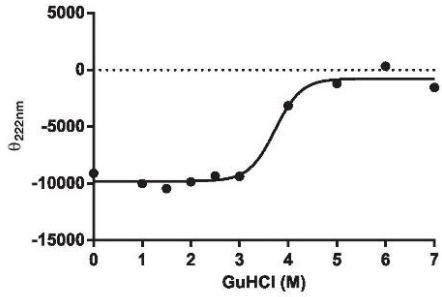


Fig. 5. GuHCl-induced unfolding curve of *Tt*SmtB.

Table 2
Denaturation temperatures of *Tt*SmtB bound to metal ions determined by circular dichroism.

*Tt*SmtB-	*Tt*SmtB Cd(II)	*Tt*SmtB-As(V)	*Tt*SmtB-As(III)	*Tt*SmtB-Sb(III)
93 ± 2.8 °C	98 ± 2.8 °C	92.5 ± 2.1 °C	92.5 ± 2.1 °C	89 ± 1.4 °C

antimony response [49]. The fluorescence spectra are shown in Fig. 2. Titration experiments showed that *Tt*SmtB undergoes a progressive decrease in fluorescence emission as a function of metal concentration; since the fluorescence quenching could not be observed when titrating a control protein (Bovine Serum Albumin with a ratio Tyr/Trp of 20/2 [50] with the investigated metals, we concluded that the interaction with all the metal ions truly takes place and that the protein undergo a metal induced allosteric switch towards a more compact homodimer. Interestingly, a higher effect of Cd(II) on the tertiary structure is observed at lower molar ratio than the other ions suggesting a higher affinity towards the former. Moreover, since no λ shift could be observed in the fluorescence spectra upon metal binding it was possible to argue that conformational rearrangements due to the exposition of hydrophobic amino acids are not required for transducing the allosteric response [51].

3.3. Circular dichroism and ITC analyses of TtSmtB binding to metal ions

In this study, the effect of Cd(II) and Sb(III) binding on *Tt*SmtB secondary structure was also analyzed by circular dichroism spectroscopy and compared to that previously reported for As(V) and As(III) [27]. Spectra showed a gradual decrease of the molar ellipticity at 208 nm at increasing Cd(II) concentration (Fig. 3A) whereas less pronounced spectral differences were observed in the presence of Sb(III) (Fig. 3B). These results suggest the occurrence of changes in *Tt*SmtB α-helix content mainly upon interaction with Cd(II).

DichroWeb analysis obtained by deconvolution of spectra measured at all the tested metal concentrations using the Selcon3 software (Table 1) confirmed a gradual rearrangement in the secondary structural elements at increasing metals concentration; in particular, at 1:25 protein-metal ratio the percentage of α-helix decreased about 10% in the presence of Cd(II), As(V) and As(III) while that of β-sheets increased about 5%, but non-significant variations were observed in the presence of Sb(III). The lower content of α helices could be traced back to the reduction of the surface of DNA recognition helix upon metal ion binding and therefore to a possible role of the interacting metals as effectors for *Tt*SmtB [27].

Overall, the results obtained by both fluorescence and CD spectroscopic analyses pointed up to cadmium as the metal ion determining major changes in the spectroscopic signals, suggesting that it could be the best effector. For this reason *Tt*SmtB interaction with Cd(II) was also analyzed by ITC experiments: the K_d value of 0.5 ± 0.04 μM calculated confirmed the high affinity (Fig. 4). The best fit binding isotherm was found to have a stoichiometry of 1:1, consistent with the binding of 2 cadmium ions per dimer and in agreement with our hypothesis that only one of the two metal binding sites is employed for Cd(II) interaction.

In order to verify a correlation between conformational changes and regulatory metal ion binding we ascertained whether Cd(II), As(V), As(III) and Sb(III) affected *Tt*SmtB thermal stability; to this purpose, we compared thermal denaturation profiles of the metal-bound and not-bound proteins by means of CD spectroscopy [52]. Due to the very high thermal protein stability, it was not possible to acquire a complete thermal unfolding curve, as the transition was not achieved at 105 °C (temperature limit for the instrument). For this reason, we followed the thermal denaturation in the presence of GuHCl as a chemical denaturing agent. To estimate the optimal concentration to employ, a GuHCl-induced denaturation curve was acquired by measuring changes in the molar ellipticity at 222 nm in the presence of increasing concentrations of GuHCl. Fig. 5 shows a typical sigmoid shaped curve, indicative of a cooperative two-state transition between folded dimers and unfolded monomers. The midpoint concentration (C_m) of GuHCl of ~3.7 M, highlights a high *Tt*SmtB resistance to the denaturant and indicates that at 3.0 M of denaturant, the protein retains its secondary structure. Afterwards, we collected thermal denaturation curves incubating the protein with 3.0 M GuHCl in the presence or absence of saturating Cd(II), Sb(III), As(V) and As(III) concentrations (in a molar ratio of 1:50). The Tm values were extrapolated from the sigmoidal curves, and the denaturation temperatures obtained without or with metal ions are shown in Table 2. Intriguingly, *Tt*SmtB thermal stability remained unchanged upon addition of As(V) or As(III), slightly reduced in the presence of Sb(III) and sensibly enhanced (5 °C) upon incubation with Cd(II), proving that the latter has a more stabilizing effect on the protein structure.

This result, in agreement with fluorescence and CD analyses suggested that the interaction with Cd(II) induces structural rearrangements towards a more compact homodimer, whereas interaction with Sb(III) induces structural rearrangements towards a less stable conformation, and hence pointed to Cd(II) as the major allosteric effector. These results also highlight that a combination of methods investigating changes in spectroscopic properties can provide a valuable tool for evaluating protein metal interaction and determining affinities.

3.4. Binding of TtSmtB to DNA in the presence of metal ions

Previous EMSA results indicated that binding of Cd(II) As(V) and As(III) to *Tt*SmtB reduced the ability to recognize specific DNA sequences [33]. To better investigate on the functional role of *Tt*SmtB in the presence of the other metal ions, the effect of Zn(II) and Sb(III) on *in vitro* binding of the protein to its cognate DNA was determined in gel shift

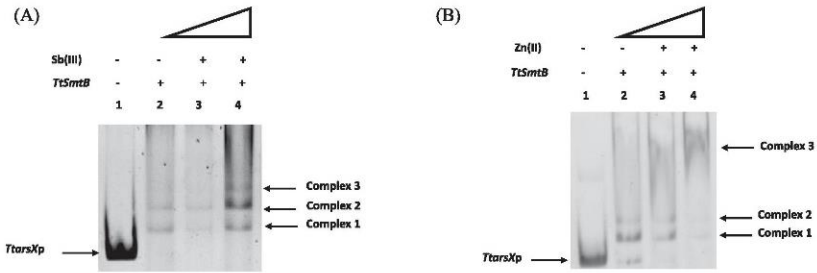


Fig. 6. EMSA analysis of *Tt*SmtB binding to cognate DNA in the presence of Sb(III) and Zn(II). A) Binding of *Tt*SmtB to *TtarsXp* without (lane 2) and with antimony at a molar ratio of 1:20 and 1:50 (lanes 3–4). B) Binding of *Tt*SmtB to *TtarsXp* without (lane 2) and with zinc at a molar ratio of 1:20 and 1:50 (lane 3–4).

assays. Fig. 6A shows that Sb(III) has only a negligible effect on DNA binding, proving that *Tt*SmtB is not able to respond efficiently to this metal. Interestingly, Zn(II) not only does not dissociate *Tt*SmtB/DNA complexes, but even increases the ability of *Tt*SmtB to bind target regulatory sequences as indicated by the formation of the slow-mobility DNA-*Tt*SmtB/complex (Fig. 6B). Altogether these data confirm previous indications that Sb(III) is a weak ligand and a poor effector and Zn(II) plays mainly a stabilizing structural effect on the conformation of DNA-bound-*Tt*SmtB.

4. Conclusion

To the best of our knowledge, *Tt*SmtB is the first thermophilic member deeply characterized within the ArsR/SmtB family and this work provides new insights into the molecular mechanisms of protein-metal interaction at high temperature. For mesophilic members, it has been demonstrated that the interaction with the regulatory metal drives allosteric conformational changes in the dimer that lowers the DNA binding affinity. Our data indicate conservation in the amino acids of the metal binding pockets and general chemical properties shared with mesophilic counterparts of the same subgroup [10] Notably, *Tt*SmtB which belongs to the α3N-α5 class, is the only ArsR/SmtB member present in the genome of *T. thermophilus* HB27. Therefore, it is conceivable that this protein has evolved as a sensor of an expanded array of heavy metal ions found in hot environments that are naturally rich in highly soluble metals [27,33]. It is known that the distinct metal binding properties displayed by α3N-α5 class members are mirrored by differences in the α3N site (amino acid composition and sequence length) making it structurally flexible and functionally versatile in adopting a wide range of coordination chemistries for detecting and discerning among metals [47].

Our data demonstrate that *Tt*SmtB functionally interacts with di-tri-penta-valent ions with different affinity and that the conformational properties of *Tt*SmtB are mainly affected upon binding with Cd(II), highlighting the Cd(II) as the best effector.

On the other hand, *Tt*SmtB interaction with Zn(II) stabilizes protein/DNA complexes suggesting a structural rather than a regulatory role for this ion. It might be speculated that *Tt*SmtB has evolved in such a way to exploit the α3N site as the sole determinant for functional metal selectivity, whereas the α5 could have lost its primary regulatory function and acquired a stabilizing role of the protein and/or of the protein-DNA complexes under the evolutionary pressure of high-temperature conditions. Further investigation on the role of the residues constituting and surrounding the metal binding site, is required to verify this hypothesis.

The detailed characterization of *Tt*SmtB paves the way to investigate at the molecular level the mechanisms responsible for the tuning of metal specificity in extreme environments. Such knowledge is fundamental not only to gain more insights into its role in biological systems but also to develop innovative strategies for biosensing, and/or bioremediation of toxic heavy metals or even for biomining those of economic value. In this context, metalloproteins working at high temperatures have the additional advantage of stability and resistance to harsh operational conditions.

Declaration of Competing Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by Università degli Studi di Napoli Federico II that financed the original research project: "Immobilization of ENzymes on hydrophobin-functionalized NAnomaterials" and by ERA-NET Cofund MarTERA who funded the project "FLAshMoB: Functional Amyloid Chimera for Marine Biosensing".

References

[1] D.P. Giedroc, A.I. Arunkumar, Metal sensor proteins: nature's metalloregulated allosteric switches, Dalton Trans. (29) (2007) 3107–3120.
[2] D. Osman, A.W. Foster, J. Chen, K. Svedaite, J.W. Steed, E. Lurie-Luke, T.G. Huggins, N.J. Robinson, Fine control of metal concentrations is necessary for cells to discern zinc from cobalt, Nat. Commun. 8 (1) (2017) 1884.
[3] L.S. Busenlehner, M.A. Pennella, D.P. Giedroc, The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance, FEMS Microbiol. Rev. 27 (2–3) (2003) 131–143.
[4] D. Osman, J.S. Cavet, Bacterial metal-sensing proteins exemplified by ArsR–SmtB family repressors, Nat. Prod. Rep. 27 (5) (2010) 668–680.
[5] G. Ybarra, R. Webb, Effects of divalent metal cations and resistance mechanisms of the cyanobacterium Synechococcus sp. strain PCC 7942, J. Hazard. Subst. Res. 2 (1) (1999) 1.
[6] J. Wu, B.P. Rosen, The arsD gene encodes a second trans-acting regulatory protein of the plasmid-encoded arsenical resistance operon, Mol. Microbiol. 8 (3) (1993) 615–623.
[7] A.J. Guerra, D.P. Giedroc, Metal site occupancy and allosteric switching in bacterial metal sensor proteins, Arch. Biochem. Biophys. 519 (2) (2012) 210–222.
[8] W.C. Warren, L.W. Hillier, J.A.M. Graves, E. Birney, C.P. Ponting, F. Grützner, K. Belov, W. Miller, L. Clarke, A.T. Chinwalla, Genome analysis of the platypus reveals unique signatures of evolution, Nature 453 (7192) (2008) 175.
[9] E. Ordóñez, S. Thiyagarajan, J.D. Cook, T.L. Stemmler, J.A. Gil, L.M. Mateos, B.P. Rosen, Evolution of metal (loid) binding sites in transcriptional regulators, J. Biol. Chem. 283 (37) (2008) 25706–25714.
[10] R. Roy, S. Samanta, S. Patra, N.K. Mahato, R.P. Saha, In silico identification and characterization of sensory motifs in the transcriptional regulators of the ArsR-SmtB family, Metallomics 10 (2018) 1476–1500.
[11] L.J. Rothschild, R.L. Mancinelli, Life in extreme environments, Nature 409 (6823) (2001) 1092.
[12] P. Contursi, S. Fusco, R. Cannio, Q. She, Molecular biology of fuselloviruses and their satellites, Extremophiles 18 (3) (2014) 473–489.
[13] R. Ladenstein, G. Antranikian, Proteins from hyperthermophiles: stability and enzymatic catalysis close to the boiling point of water, Biotechnology of Extremophiles, Springer 1998, pp. 37–85.
[14] G. Fiorentino, R. Cannio, M. Rossi, S. Bartolucci, Decreasing the stability and changing the substrate specificity of the Bacillus stearothermophilus alcohol dehydrogenase by single amino acid replacements, Protein Eng. 11 (10) (1998) 925–930.
[15] P. Contursi, S. Fusco, D. Limauro, G. Fiorentino, Host and viral transcriptional regulators in Sulfolobus: an overview, Extremophiles 17 (6) (2013) 881–895.
[16] P. Contursi, R. Cannio, S. Prato, Q. She, M. Rossi, S. Bartolucci, Transcriptional analysis of the genetic element pSSVx: differential and temporal regulation of gene expression reveals correlation between transcription and replication, J. Bacteriol. 189 (17) (2007) 6339–6350.
[17] S. Prato, R.M. Vitale, P. Contursi, G. Lipps, M. Saviano, M. Rossi, S. Bartolucci, Molecular modeling and functional characterization of the monomeric primase–polymerase domain from the Sulfolobus solfataricus plasmid pIT3, FEBS J. 275 (17) (2008) 4389–4402.
[18] R. Gaglione, L. Pirone, B. Farina, S. Fusco, G. Smaldone, M. Aulitto, E. Dell'Olmo, E. Roscetto, A. Del Gatto, R. Fattorusso, Insights into the anticancer properties of the first antimicrobial peptide from Archaea, Biochim. Biophys. Acta Gen. Subj. 1861 (9) (2017) 2155–2164.
[19] J. Politi, J. Spadavecchia, G. Fiorentino, I. Antonucci, L. De Stefano, 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 (123) (2016), 20160629.
[20] M. Aulitto, S. Fusco, G. Fiorentino, D. Limauro, E. Pedone, S. Bartolucci, P. Contursi, Thermus thermophilus as source of thermozymes for biotechnological applications: homologous expression and biochemical characterization of an α-galactosidase, Microb. Cell Factories 16 (1) (2017) 28.
[21] F.A. Fusco, G. Fiorentino, E. Pedone, P. Contursi, S. Bartolucci, D. Limauro, Biochemical characterization of a novel thermostable β-glucosidase from Dictyoglomus turgidum, Int. J. Biol. Macromol. 113 (2018) 783–791.
[22] S. Elleuche, C. Schäfers, S. Blank, C. Schröder, G. Antranikian, Exploration of extremophiles for high temperature biotechnological processes, Curr. Opin. Microbiol. 25 (2015) 113–119.
[23] M. Lanzilli, G. Donadio, F.A. Fusco, C. Sarcinelli, D. Limauro, E. Ricca, R. Isticato, Display of the peroxiredoxin Bcp1 of Sulfolobus solfataricus on probiotic spores of Bacillus megaterium, New Biotechnol. 46 (2018) 38–44.
[24] P. Ranawat, S. Rawat, Metal-tolerant thermophiles: metals as electron donors and acceptors, toxicity, tolerance and industrial applications, Environmental Science and Pollution Research 2017, pp. 1–29.
[25] C. Castro, M. Urbieta, J.P. Cazón, E. Donati, Metal biorecovery and bioremediation: whether or not thermophilic are better than mesophilic microorganisms, Bioresour. Technol. 279 (2019) 317–326.
[26] G.H. Wheaton, A. Mukherjee, R.M. Kelly, Transcriptomes of the extremely thermoacidophilic archaeon Metallosphaera sedula exposed to metal "shock" reveal generic and specific metal responses, Appl. Environ. Microbiol. 82 (15) (2016) 4613–4627.
[27] I. Antonucci, G. Gallo, D. Limauro, P. Contursi, A.L. Ribeiro, A. Blesa, J. Berenguer, S. Bartolucci, G. Fiorentino, An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in Thermus thermophilus HB27, Microb. Biotechnol. 10 (6) (2017) 1690–1701.

[28] G. Gallo, R. Puopolo, D. Limauro, S. Bartolucci, G. Fiorentino, Metal-tolerant thermophiles: from the analysis of resistance mechanisms to their biotechnological exploitation, Open Biochem. J. 12 (1) (2018).
[29] R.P. Saha, S. Samanta, S. Patra, D. Sarkar, A. Saha, M.K. Singh, Metal homeostasis in bacteria: the role of ArsR-SmtB family of transcriptional repressors in combating varying metal concentrations in the environment, Biometals 30 (4) (2017) 459–503.
[30] C. Eicken, M.A. Pennella, X. Chen, K.M. Koshlap, M.L. VanZile, J.C. Sacchettini, D.P. Giedroc, A metal–ligand-mediated intersubunit allosteric switch in related SmtB/ArsR zinc sensor proteins, J. Mol. Biol. 333 (4) (2003) 683–695.
[31] I. Del Giudice, D. Limauro, E. Pedone, S. Bartolucci, G. Fiorentino, A novel arsenate reductase from the bacterium Thermus thermophilus HB27: its role in arsenic detoxification, Biochim. Biophys. Acta Protein Proteomics 1834 (10) (2013) 2071–2079.
[32] E. Pedone, D. Limauro, R. D'alterio, M. Rossi, S. Bartolucci, Characterization of a multifunctional protein disulfide oxidoreductase from Sulfolobus solfataricus, FEBS J. 273 (23) (2006) 5407–5420.
[33] I. Antonucci, G. Gallo, D. Limauro, P. Contursi, A.L. Ribeiro, A. Blesa, J. Berenguer, S. Bartolucci, G. Fiorentino, Characterization of a promiscuous cadmium and arsenic resistance mechanism in Thermus thermophilus HB27 and potential application of a novel bioreporter system, Microb. Cell Factories 17 (1) (2018) 78.
[34] G. Fiorentino, I. Del Giudice, L. Petraccone, S. Bartolucci, P. Del Vecchio, Conformational stability and ligand binding properties of BldR, a member of the MarR family, from Sulfolobus solfataricus, Biochim. Biophys. Acta Protein Proteomics 1844 (6) (2014) 1167–1172.
[35] M. Johnson, I. Zaretskaya, Y. Raytselis, Y. Merezhuk, S. McGinnis, T.L. Madden, NCBI BLAST: a better web interface, Nucleic Acids Res. 36 (suppl_2) (2008) W5–W9.
[36] Y. Zhang, I-TASSER server for protein 3D structure prediction, BMC Bioinf. 9 (1) (2008) 40.
[37] G. Macindoe, L. Mavridis, V. Venkatraman, M.-D. Devignes, D.W. Ritchie, HexServer: an FFT-based protein docking server powered by graphics processors, Nucleic Acids Res. 38 (suppl_2) (2010) W445–W449.
[38] M. Van De Weert, Fluorescence quenching to study protein-ligand binding: common errors, J. Fluoresc. 20 (2) (2010) 625–629.
[39] G. Fiorentino, I. Del Giudice, S. Bartolucci, L. Durante, L. Martino, P. Del Vecchio, Identification and physicochemical characterization of BldR2 from Sulfolobus solfataricus, a novel archaeal member of the MarR transcription factor family, Biochemistry 50 (31) (2011) 6607–6621.
[40] N. Sreerama, R.W. Woody, A self-consistent method for the analysis of protein secondary structure from circular dichroism, Anal. Biochem. 209 (1) (1993) 32–44.
[41] N. Sreerama, S.Y. Venyaminov, R.W. Woody, Estimation of the number of α-helical and β-strand segments in proteins using circular dichroism spectroscopy, Protein Sci. 8 (2) (1999) 370–380.
[42] L. Whitmore, B.A. Wallace, Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases, Biopolymers 89 (5) (2008) 392–400.
[43] L. Whitmore, B. Wallace, DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data, Nucleic Acids Res. 32 (suppl_2) (2004) W668–W673.
[44] G. Fiorentino, R. Ronca, R. Cannio, M. Rossi, S. Bartolucci, MarR-like transcriptional regulator involved in detoxification of aromatic compounds in Sulfolobus solfataricus, J. Bacteriol. 189 (20) (2007) 7351–7360.
[45] T. Liu, S. Nakashima, K. Hirose, M. Shibasaka, M. Katsuhara, B. Ezaki, D.P. Giedroc, K. Kasamo, A novel cyanobacterial SmtB/ArsR family repressor regulates the expression of a CPx-ATPase and a metallothionein in response to both Cu (I)/Ag (I) and Zn (II)/Cd (II), J. Biol. Chem. 279 (17) (2004) 17810–17818.
[46] C. Thelwell, N.J. Robinson, J.S. Turner-Cavet, An SmtB-like repressor from Synechocystis PCC 6803 regulates a zinc exporter, Proc. Natl. Acad. Sci. 95 (18) (1998) 10728–10733.
[47] T. Liu, X. Chen, Z. Ma, J. Shokes, L. Hemmingsen, R.A. Scott, D.P. Giedroc, A Cu(I)-sensing ArsR family metal sensor protein with a relaxed metal selectivity profile, Biochemistry 47 (40) (2008) 10564–10575.
[48] L.S. Busenlehner, N.J. Cosper, R.A. Scott, B.P. Rosen, M.D. Wong, D.P. Giedroc, Spectroscopic properties of the metalloregulatory Cd (II) and Pb (II) sites of S. aureus pI258 CadC, Biochemistry 40 (14) (2001) 4426–4436.
[49] L. López-Maury, F.J. Florencio, J.C. Reyes, Arsenic sensing and resistance system in the cyanobacterium Synechocystis sp. strain PCC 6803, J. Bacteriol. 185 (18) (2003) 5363–5371.
[50] N. Zhdanova, E. Maksimov, A. Arutyunyan, V. Fadeev, E. Shirshin, Tyrosine fluorescence probing of conformational changes in tryptophan-lacking domain of albumins, Spectrochim. Acta A Mol. Biomol. Spectrosc. 174 (2017) 223–229.
[51] E. Pedone, M. Saviano, M. Rossi, S. Bartolucci, A single point mutation (Glu85Arg) increases the stability of the thioredoxin from Escherichia coli, Protein Eng. 14 (4) (2001) 255–260.
[52] T. Mondol, J. Åden, P. Wittung-Stafshede, Copper binding triggers compaction in N-terminal tail of human copper pump ATP7B, Biochem. Biophys. Res. Commun. 470 (3) (2016) 663–669.

Chapter IV

Characterization of a promiscuous cadmium and arsenic resistance mechanism in *Thermus thermophilus* HB27 and potential application of a novel bioreporter system

In this chapter we analyze the biological role of another component of the detoxification system which's responsible for the export of arsenic. This protein named *Tt*ArsX is a P_{1B}-type membrane ATPase responsible for the efflux of heavy metals.
Bioinformatic and biochemical data prove that *Tt*ArsX is capable of also exporting cadmium supporting the idea that the arsenic resistance system in *T. thermophilus* HB27 is promiscuous. In this chapter, it's also demonstrated that *Tt*SmtB is capable of interacting with Cd(II).
Based on the knowledge obtained in the previous and in the present chapter, we have developed a biosensing system: in which in the plasmid pMHp*nqo*, has been cloned a cassette containing the promoter recognized by *Tt*SmtB, downstream of the reporter gene coding for a thermostable version of β-galactosidase, and an expression cassette for the factor *Tt*SmtB under the control of the constitutive promoter p*nqo*. Using this reporter system, we have obtained a first example of *whole-cell* biosensor for heavy metals, obtained transforming thermophilic cells of *T. thermophilus* HB27 with this construct.
This paper outlines the importance of a detailed characterization of the molecular components (intrinsic promoter activity, repressor/promoter and repressor/metal(s) binding affinities) and points to *T. thermophilus* as suitable chassis cell for design and development of robust metal biosensors. In this context, advantageous modifications can be programmed to increase biosensor sensitivity, selectivity and/or ability to detect metal mixtures.

Antonucci et al. Microb Cell Fact (2018) 17:78
https://doi.org/10.1186/s12934-018-0918-7

Microbial Cell Factories

RESEARCH

Open Access

Characterization of a promiscuous cadmium and arsenic resistance mechanism in *Thermus thermophilus* HB27 and potential application of a novel bioreporter system

Immacolata Antonucci[1], Giovanni Gallo[1], Danila Limauro[1], Patrizia Contursi[1], Ana Luisa Ribeiro[2], Alba Blesa[2], José Berenguer[2], Simonetta Bartolucci[1] and Gabriella Fiorentino[1]*

Abstract

Background: The characterization of the molecular determinants of metal resistance has potential biotechnological application in biosensing and bioremediation. In this context, the bacterium *Thermus thermophilus* HB27 is a metal tolerant thermophile containing a set of genes involved in arsenic resistance which, differently from other microbes, are not organized into a single operon. They encode the proteins: arsenate reductase, *Tt*ArsC, arsenic efflux membrane transporter, *Tt*ArsX, and transcriptional repressor, *Tt*SmtB.

Results: In this work we show that the arsenic efflux protein *Tt*ArsX and the arsenic responsive transcriptional repressor *Tt*SmtB are required to provide resistance to cadmium. We analyzed the sensitivity to Cd(II) of mutants lacking *Tt*ArsX, finding that they are more sensitive to this metal than the wild type strain. In addition, using promoter probe reporter plasmids, we show that the transcription of *TtarsX* is also stimulated by the presence of Cd(II) in a *Tt*SmtB-dependent way. Actually, a regulatory circuit composed of *Tt*SmtB and a reporter gene expressed from the *TtarsX* promoter responds to variation in Cd(II), As(III) and As(V) concentrations.

Conclusions: Our results demonstrate that the system composed by *Tt*SmtB and *Tt*ArsX is responsible for both the arsenic and cadmium resistance in *T. thermophilus*. The data also support the use of *T. thermophilus* as a suitable chassis for the design and development of As-Cd biosensors.

Keywords: Thermophiles, Cadmium and arsenic resistance, Thermophilic reporter systems

Background

Toxic metals and metalloids such as cadmium (Cd) and arsenic (As) are widespread environmental contaminants that pose risks to human health [1]. Microorganisms are endowed with multiple molecular mechanisms to handle exposure to these toxic compounds. In general, microbial resistance is achieved through three main mechanisms: transformation of the metals through reduction to a different oxidation state, efflux outside the cell by transporters, and/or sequestration/biosorption [2]. Common reduction mechanisms include for example the conversion of Hg^{2+} to Hg^0, AsO_4^{3-} to AsO_3^{3-}. Facilitated efflux transporters fall into two wide, functionally and evolutionary distinct membrane protein families, the P-type ATPases and the Major Facilitator Superfamily (MFS) of transporters [3–5]. P-type ATPases use ATP hydrolysis to transport ions across cellular membranes and are composed of three conserved domains: (1) the transmembrane (TM) helix bundle, allowing substrate translocation; (2) the soluble ATP binding domain (ATPBD) that contains the transiently phosphorylated Asp residue; (3) the soluble actuator domain (AD) [6]. Over the years, on the basis

*Correspondence: fiogabri@unina.it
[1] Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Complesso Universitario Monte S. Angelo, Naples, Italy
Full list of author information is available at the end of the article

of sequence similarity and overall architecture, they have been divided into different classes: those belonging to P_{1B}-type are capable to drive the efflux out of cells of both essential transition metal ions (e.g., Zn^{2+}, Cu^{+}, and Co^{2+}) and toxic metal ions (e.g., Ag^{+}, Cd^{2+}, Pb^{2+}) contributing to their homeostasis maintenance. A recent study on a huge number of P_{1B}-type ATPase sequences combined with available biochemical data classifies them into seven distinct subfamilies ($1_{B\text{-}1}$ $1_{B\text{-}7}$) on the basis of conserved motifs in TM4, TM5 and TM6, but the molecular basis of metal ion specificity remains unclear [7]. All metal efflux transporters characterized so far are tightly controlled at transcriptional level by metalloregulatory proteins which bind DNA sequences and dissociate following metal binding, thus ensuring derepression of genes encoding efflux proteins [8]. Several regulatory systems dedicated to metal/metalloid sensing have also been characterized; for instance, transcription factors of the ArsR/SmtB family are small dimeric proteins with a winged helix-turn-helix DNA binding domain controlling gene expression in response to divalent metals (e.g. zinc, cadmium) as well as metalloids (e.g. arsenic and antimonite) through an allosteric switching mechanism [9]. The exploration of life in extreme environments has led to the isolation of many thermophilic microorganisms occupying diverse extreme habitats like hot hydrothermal fluids containing high concentrations of toxic metals. For this reason they are able to cope with toxic metals, which are more soluble at high temperatures, or even to use them for their metabolism [10, 11] and are currently exploited in some bioprocesses such as biomining and bioremediation [12]. A detailed understanding of the molecular mechanisms responsible for resistance to toxic metals is also crucial for engineering organisms to develop sensitive biosensors for the detection of chemicals in the environment and to enhance bioremediation strategies [13, 14].

A significant number of whole-cell arsenic and/or cadmium biosensors has been already described in literature and is based on the realization of reporter systems containing regulatory cis-acting sequences interacting with a transcriptional repressor belonging to the ArsR/SmtB family [15, 16]. Biosensors are not intended to fully replace chemical methods but have the advantage of lower fabrication cost and higher stability and can offer on-site monitoring of even trace levels of targeted compounds in comparison with non-portable analytical methodologies [17]. To date, the major challenges in biosensor development regard the screening or modification of efficient regulator protein/promoter pair for increased sensitivity and specificity [18, 19] as well as biosensor stability over time and simultaneous monitoring of multiple environmental parameters [20].

Thermus thermophilus HB27 is a thermophilic aerobic Gram-negative bacterium capable of growing in the presence of arsenic concentrations that are lethal for other microorganisms [21]. In recent studies we demonstrated that the arsenic resistance system is not clustered in a single *ars* operon as in other organisms, but the genes are spread in the genome: *TTC1502* encoding a cytoplasmic arsenate reductase (*Tt*ArsC) able to reduce arsenate to arsenite, *TTC0354*, encoding a P_{1B}-type membrane ATPase responsible for the efflux (herein named *Tt*ArsX) and *TTC0353* encoding a repressor (*Tt*SmtB) sensitive to both As(V) and As(III) [22–24]. *Tt*SmtB is a member of the ArsR/SmtB family, sharing 50% identity with the well characterized SmtB of *Synechococcus* PCC7942 [25]. It is a dimeric protein containing three Cys residues in a reduced state and a conserved metal binding box presumably involved in As(V) and As(III) binding. The protein can bind to a consensus regulatory sequence located upstream of *TtarsX* preferentially in an un-metallated state and in vivo *Tt*SmtB regulates *TtarsX* transcription upon arsenic interaction through a derepression mechanism [24].

In the present study, we evaluated the contribution of *Tt*SmtB and *Tt*ArsX in cadmium detoxification using a combination of genetic and physiological approaches. The results obtained support that in *T. thermophilus* the mechanism employed for survival to cadmium and arsenic exposure is promiscuous, suggesting that the evolution of shared/common defense mechanisms represents an adaptation strategy to cope with toxic metals at high temperatures whereas keeping a reduced genome. In addition, to analyze the cadmium/arsenic response of *TtarsX* promoter in dependency on varying *Tt*SmtB concentration, we settled up different β-gal reporter systems with the final goal of evaluating the utilization of thermophilic molecular components and thermostable chassis cells in biosensing applications.

Methods

Culture conditions and determination of minimum inhibitory concentration

Strains, genotypes and sources are summarized in Additional file 1: Table S1.

T. thermophilus HB27 wild type strain, *T. thermophilus ΔsmtB::kat* and *TTC0354::pK18* mutants were grown aerobically at 70 °C in TM medium as described [24].

Minimum inhibitory concentration (MIC) was determined as the lowest concentration of cadmium that completely inhibited the visible growth of the strains after overnight incubation as indicated by the lack of

turbidity. Basically, exponentially growing cultures of *T. thermophilus* HB27, *T. thermophilus ΔsmtB::kat* and *TTC0354::pK18* were diluted to 0.08 $OD_{600\ nm}$ in 24 well plates containing increasing concentrations of $CdCl_2$ ranging from 0 mM to 5 mM and incubated at 70 °C or 60 °C for 18 h; depending on the strain tolerance, the concentration interval was narrowed in consecutive experiments; the MIC endpoint was considered as the lowest Cd(II) concentration at which there was a difference between grown and start culture lower than $0.01OD_{600\ nm}$. The values reported are the average of two independent determinations.

In silico analysis

BlastP analyses were performed using the Blastall (v.2.2.25) program. The predicted presence and number of TMs, in the full length *Tt*ArsX, was determined using the TMHMM 2.0 online (http://www.cbs.dtu.dk/services/TMHMM) and the TM-pred online servers (http://www.ch.embnet.org/software/TMPRED_form.html) [26, 27]. Conserved motifs in TM helices of *Tt*ArsX were determined by manual inspection referring to those identified by Smith et al. [7]. The presence of Metal Binding Domain(s) (MBD) was determined by manually looking at a CXXC motif in the protein sequence and fold prediction performed at (http://pfam.sanger.ac.uk).

Models of the MBD were generated using I-TASSER web server [28] (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) using as input the first 91 amino acids of *Tt*ArsX. The model was selected according to its similarity to available crystal structures of MBD of P_{1B}-type ATPases from other species [6].

The docking calculations were obtained using Hex Protein Docking server [29] with *Tt*ArsX MBD and As(III) or Cd(II); 100 rigid body docking solutions were generated and the 10 best refined by energy minimization. The proposed model for the metal docked into the MBD is the structure with the smallest distance between the metal and cysteine [2.61 Å for As(III) or 2.35 Å for Cd(II)], after an energy minimization step.

Bioreporter constructions

The regulatory region upstream of *TtarsX* (*TtarsXp*) previously named *TTC0354p*, was amplified by PCR using the primer pairs *new 0354pr fwEcoRI* and *R0354NdeI*, respectively (Additional file 1: Table S2); the region extends from −73 to +1 from the transcription start site [24]. The primers introduced *Eco*RI and *Nde*I restriction sites, so that the amplified fragment could be cloned in the adapted pMHbgaA plasmid [30]. The new vector was named pMH*TtarsXp*bgaA.

To obtain the plasmid pMH*TtarsXp*bgaA-nqo*Tt*smtB, the *pnqo-TtsmtB* gene cassette, where *TtsmtB* is under the control of the *nqo* promoter, was cloned into pMH*TtarsXp*bgaA vector; in *T. thermophilus*, the *nqo* promoter drives the expression of the operon encoding the major respiratory complex I during aerobic growth [31]. In particular, pET28*/TtsmtB* was digested with *Nde*I *Hind*III [24] and cloned into the corresponding sites of pMKpnqo-bgaA [31] giving pMKpnqo-*TtsmtB*; afterwards, the plasmid was digested with *Xba*I *Hind*III and the gel purified *pnqo-TtsmtB* cassette subjected to a fill-in reaction and cloned into the filled-in *Hind*III site of pMH*TtarsXp*bgaA.

pMH*TtarsXp*bgaA-nqo*Tt*SmtB was used to transform *T. thermophilus* HB27 and *TTC0354::pK18* mutants in the conditions described [24]. The pMHPnorbgaA vector was also used to transform the same strains and used as negative control [32]. Cells were then incubated for 24–48 h at 60 °C on TM plates containing hygromycin (100 μg/mL). All the plasmids used in this study are described in Additional file 1: Table S3.

β-galactosidase assays

For measuring reporter β-galactosidase activity, the growing transformants were diluted to 0.1 $OD_{600\ nm}$ in TM medium in the presence of hygromycin (100 μg/ml), treated with different concentrations of $NaAsO_2$, KH_2AsO_4 and $CdCl_2$ (Sigma) as source of As(III), As(V) and Cd(II) respectively, and grown at 60 °C for 16 h. β-galactosidase activity assays were carried out on permeabilized cells in 96-well microplates at 70 °C with a Synergy H4 microplate reader (BioTeK) as described by Miller [33].

Miller units (U) were calculated by the equation:

$$U = \frac{OD_{420} - (1.75 \times OD_{550})}{t(\text{min}) \times \text{V(mL)} \times OD_{600}}$$

where: OD_{420} = OD of the chromogenic product, OD_{550} = OD of the cellular debris, t = time of reaction, V = volume of used cells and OD_{600} = OD of the cell culture. β-galactosidase activity of *T. thermophilus* negative control (transformed with pMHPnorbgaA vector) was subtracted from that of the samples. The activity reported is the average of two or three independent experiments each made in triplicate. The error bars indicate the standard deviation of the average values. Miller Units expressed as a percentage, were calculated assuming that the Miller Units value of not treated cells (control) was 100%. Statistical analysis was performed using a Student's t test; significant differences are indicated as: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

Purification of *Tt*SmtB

Recombinant *Tt*SmtB was purified to homogeneity using the procedure already described, consisting of

thermo-precipitation of the *E. coli* BL21-Codon Plus (DE3) RIL/*Tt*SmtB cell extract followed by HiTrap Heparin chromatography. The histidine tag was removed from purified *Tt*SmtB by thrombin digestion (Sigma). The purified protein was stored in aliquots at −20 °C [24].

Electrophoretic mobility shift assay (EMSA)

To determine if cadmium was ligand of *Tt*SmtB, electrophoretic mobility shift assays (EMSA) were performed. The *TtarsX* promoter region was amplified by PCR using specific primer pair: *0354footprint fw* and *0354footprint rv*, (Additional file 1: Table S2). EMSA reactions were set up as described [34], using 5 μM of proteins pre-incubated or not with Cd(II) at molar ratio of 1:20 and 1:50 (considering *Tt*SmtB as a dimer).

Results and discussion

Domain organization and subfamily classification of *Tt*ArsX

Our recent work showed that TTC0354, herein named *Tt*ArsX, is a membrane metal-transporter involved in arsenic detoxification; in fact, a mutant strain in which this gene had been knocked out, was about 15-fold more sensitive to both As(III) and As(V) treatment than the wild type [24]. In order to analyse the putative role of this protein in the detoxification of metals different from As(III)/As(V), in particular Cd(II), at first some in silico studies were performed aimed at identifying divergence or not in conserved motifs. TM helix prediction tools suggested that *Tt*ArsX is composed of six TM helices containing all the conserved signatures of the characterized P_{1B-2}-subclass members which generally display a dual role in Zn^{2+} transport and toxic metal ion detoxification, especially Cd(II) [7, 35]. They are: SXP and CPC motifs in TM4, a $T(X)_5QN(X)_7K$ motif in TM5 and a $DXG(X)_7N$ in TM6 (Fig. 1a in bold). A previous work had already shown that *Tt*ArsX is a P_{1B}-type ATPase stimulated in vitro by Zn^{2+}/Cd^{2+} cations but with an unclear role in their tolerance in vivo [23]. The overall topology of *Tt*ArsX (Fig. 1b) obtained integrating results from a 3D model of the protein [24] also helped to identify a soluble MBD containing the CXXC motif that, as indicated by docking analyses, could be responsible for both Cd(II) and As(III) recognition inside the cell (Fig. 1c).

Taken together, the results of in silico analysis assign *Tt*ArsX to the well characterized P_{1B-2}-subfamily, show the presence of a soluble heavy metal associated domain (Pfam: PF 00403) and suggest a wider metal ion specificity than that previously known.

Contribution of *Tt*ArsX and *Tt*SmtB to the Cd(II) tolerance mechanism

To analyse the role of *Tt*ArsX and the transcriptional repressor *Tt*SmtB [24] in resistance to cadmium, we used different physiological and genetic approaches.

First, we compared the growth of *T. thermophilus* HB27 and two mutants defective in *Tt*ArsX or *Tt*SmtB (*T. thermophilus TTC0354::pK18* and *T. thermophilus ΔsmtB::kat*, respectively) in the presence of different Cd(II) concentrations, and determined the MIC values which are reported in Table 1. *T. thermophilus TTC0354::pK18* revealed a 15-fold increase in cadmium sensitivity in comparison to the wild type, thus showing a role of *Tt*ArsX in Cd(II) resistance; on the other hand, *T. thermophilus ΔsmtB::kat*, showed a 3-fold increase in cadmium sensitivity as compared to wild type HB27. One possible explanation for this unexpected result is a polar effect of the *TtsmtB::kat* mutation: in the mutant strain the kanamycin resistance gene is in counter sense; therefore, the basal levels of *Tt*ArsX could be lower than in the wild type, making the cells more sensitive.

In addition, the comparison of MIC values with those previously reported for As(V) and As(III) showed that *T. thermophilus* HB27 is almost 14-fold more sensitive to Cd(II) than to arsenic (Table 1).

To analyse in vivo whether *TtarsX* promoter had Cd(II) responsive activity, the regulatory region was cloned in a promoter probe vector upstream of the *bgaA* gene, encoding a thermostable β-galactosidase [30]. The plasmid pMH*TtarsXp*bgaA (Fig. 2a) was transformed into *T. thermophilus* HB27 and the β-gal activity was measured also upon Cd(II) treatment at 10, 20 and 100 μM (see "Methods"). The results in Fig. 2b show that transcription from *TtarsX* promoter in the reporter system is activated by Cd(II) by twofold at 20 μM (633 ± 84 Miller Units, MU) in comparison to values in untreated cells (393 ± 40 MU). At 100 μM Cd(II) transcription is also increased in comparison to the control, but at lower levels; the reduced activity could be the consequence of partial toxicity, since reduced growth rates were observed under these conditions (Additional file 1: Figure S1).

To check whether *Tt*SmtB was the transcriptional regulator of *Tt*ArsX in the presence of Cd(II), we performed in vivo and in vitro experiments. The comparison of β-gal activities in *T. thermophilus* HB27-pMH*TtarsXp*bgaA and *T. thermophilus ΔsmtB::kat*-pMH*TtarsXp*bgaA grown in the absence of metals (393 ± 40 and 887 ± 138 MU, respectively), is consistent with the in vivo role of *Tt*SmtB as a repressor. Moreover, EMSA assays were carried out with recombinant *Tt*SmtB on the *TtarsX* promoter after preincubation with or without Cd(II). As shown in Fig. 3, our results confirm that the protein interacts with *TtarsXp* and indicate that, upon binding with Cd(II), binding of *Tt*SmtB to the promoter is hampered. *Tt*SmtB DNA-binding behavior upon interaction with Cd(II) is similar to that already observed with As(III) and As(V) [24], suggesting that metal

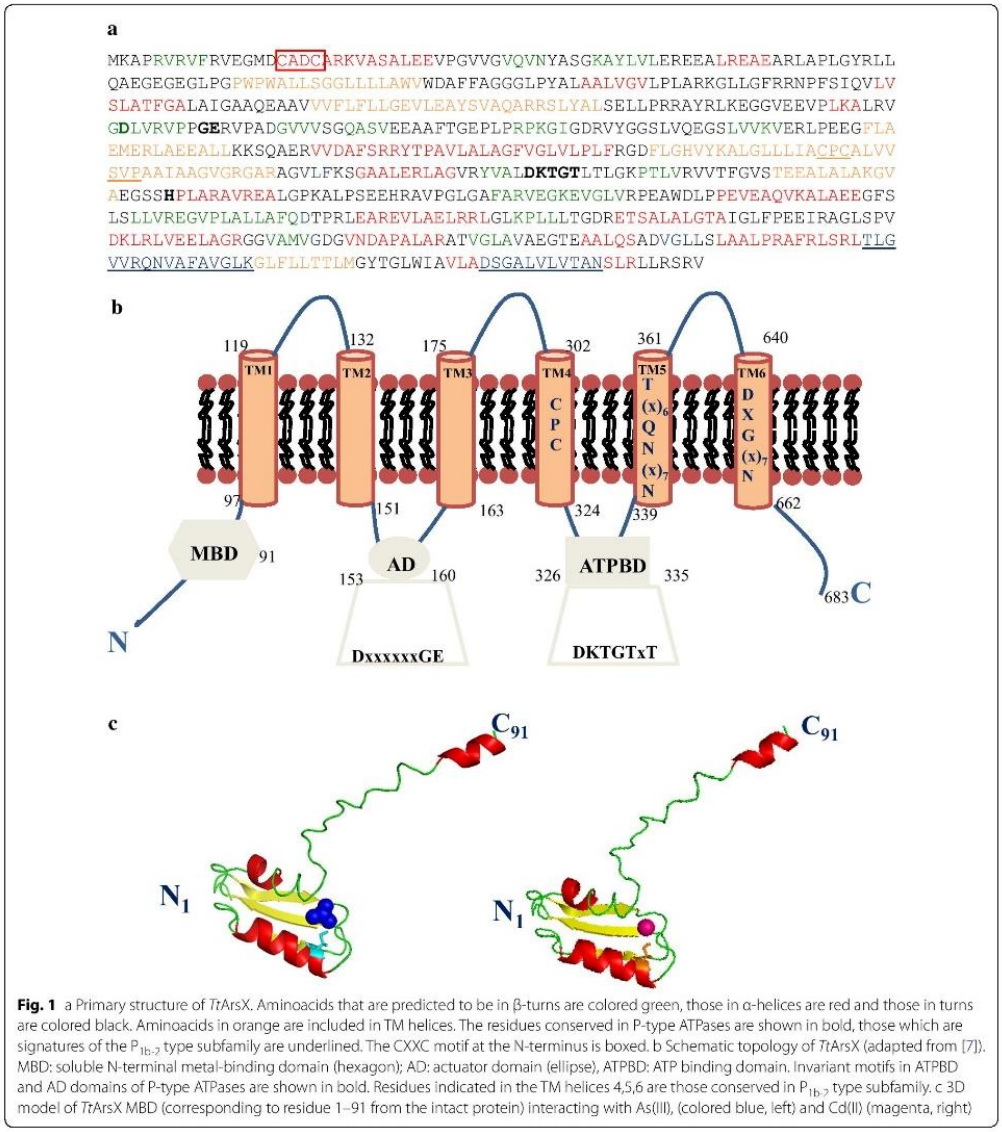


Fig. 1 a Primary structure of *Tt*ArsX. Aminoacids that are predicted to be in β-turns are colored green, those in α-helices are red and those in turns are colored black. Aminoacids in orange are included in TM helices. The residues conserved in P-type ATPases are shown in bold, those which are signatures of the P_{1b-2} type subfamily are underlined. The CXXC motif at the N-terminus is boxed. b Schematic topology of *Tt*ArsX (adapted from [7]). MBD: soluble N-terminal metal-binding domain (hexagon); AD: actuator domain (ellipse), ATPBD: ATP binding domain. Invariant motifs in ATPBD and AD domains of P-type ATPases are shown in bold. Residues indicated in the TM helices 4,5,6 are those conserved in P_{1b-2} type subfamily. c 3D model of *Tt*ArsX MBD (corresponding to residue 1–91 from the intact protein) interacting with As(III), (colored blue, left) and Cd(II) (magenta, right)

ions could take contacts with the same protein motif. Moreover, these results strongly support the role of *Tt*SmtB in Cd(II) mediated *TtarsX* transcription where it works as repressor.

Altogether, these results indicate that *Tt*SmtB regulates cadmium tolerance by controlling at transcriptional level the metal efflux gene, adopting a derepression mechanism similar to that employed for arsenic detoxification [24].

Table 1 Bacterial resistance to cadmium and arsenic

MIC			
Strain	Cd(II) (mM)	As(III)[a] (mM)	As(V)[a] (mM)
T. thermophilus HB27	3	40	44
T. thermophilus TTC0354::pK18	0.2	3	3
T. thermophilus ΔsmtB::kat	1	32	18

MICs were determined in TM medium with increasing concentrations of cadmium (0–5 mM) as already measured for As(V) and As(III) [23[a]]

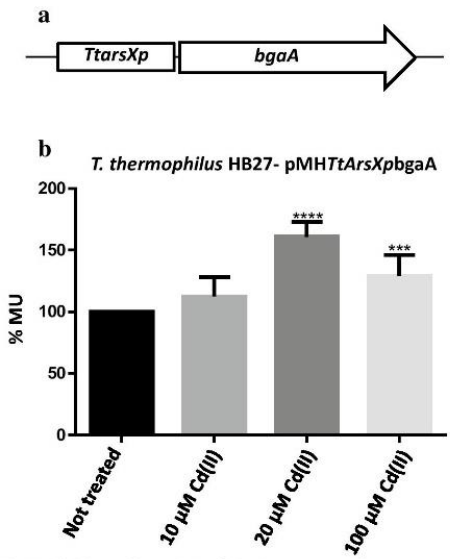


Fig. 2 Cd(II) dependent activity of *TtarsX* promoter with transcriptional *bgaA* fusions. a Schematic representation of the reporter system used to transform *T. thermophilus* HB27 and *TTC0354::pK18*. b β-gal activity expressed in % MU of *T. thermophilus* HB27-pMH*TtarsXp*bgaA treated or not with 10, 20 and 100 µM of $CdCl_2$. Statistical analysis was performed using the Student's t test; significant differences are indicated as: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

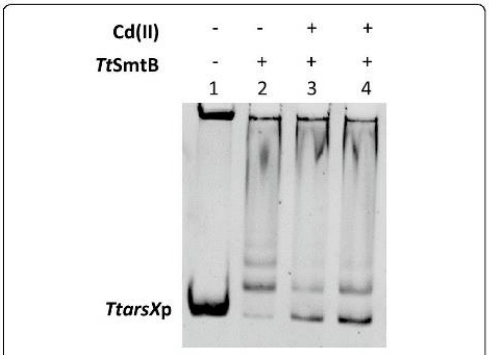


Fig. 3 EMSA analysis of *Tt*SmtB in the presence of Cd(II). Binding of *Tt*SmtB to *TtarsXp* without (lane 2) and with $CdCl_2$ at molar ratio of 1:20 and 1:50 (lanes 3–4)

Bioreporter construction and characterization

To evaluate the potential of the pair composed by *Tt*SmtB and *TtarsX* responsive promoter as components of a bioreporter system for toxic metal detection, we aimed at characterizing their cadmium and arsenic sensing potential in engineered *T. thermophilus* cells. Hence, we generated a new reporter plasmid, pMH*TtarsXp*bgaA-nqo*Tt*SmtB, to measure the transcription of the reporter gene from the *TtarsX* promoter in a context in which the *Tt*smtB gene was expressed constitutively from the nqo promoter (Fig. 4a). We expected that an increase in the intracellular concentration of the transcription factor allowed a more efficient repression of the system and, as a consequence, an activity of the reporter gene mainly depending on metal concentration. As shown in Fig. 4b, the system responds to increasing concentrations of Cd(II) almost in a linear form in a wide range, but with a slope too flat to measure any concentration accurately. As we suspected that this was due to the presence of *Tt*ArsX actively extruding Cd(II), we assayed the bioreporter in *T. thermophilus TTC0354::pK18* defective mutant. As expected (Fig. 4c), the increase in bioreporter activity was higher in the mutant strain than in the parental one; in fact at 10 µM Cd(II) a twofold induction was measured in the first, whereas no significant induction occurs in the wild type; furthermore, the mutant strain was able to detect Cd(II) at 5 µM, a concentration value that is fourfold lower than that required to get a similar signal in the parental strain. However, the tradeoff of this system is that in the mutant strain the response decreases above 20 µM Cd(II). Therefore, it could be envisioned to use both host systems to cope with a wider range of Cd(II) concentrations. The increase in the sensitivity of a biosensor using strains with mutations in the arsenic efflux pump in place of the wild type has been also reported for *P. putida* [36].

As previous functional studies showed that *Tt*SmtB and *Tt*ArsX are part of the arsenic resistance system, we evaluated the bioreporter system for its capability to detect arsenate and arsenite in *T. thermophilus TTC0354::pK18*. Figure 5 shows that the first significant reporter signal was obtained at 20 µM As(V) and 10 µM As(III), demonstrating that inactivation of Cd/As efflux pump translates in higher response sensitivities to arsenic

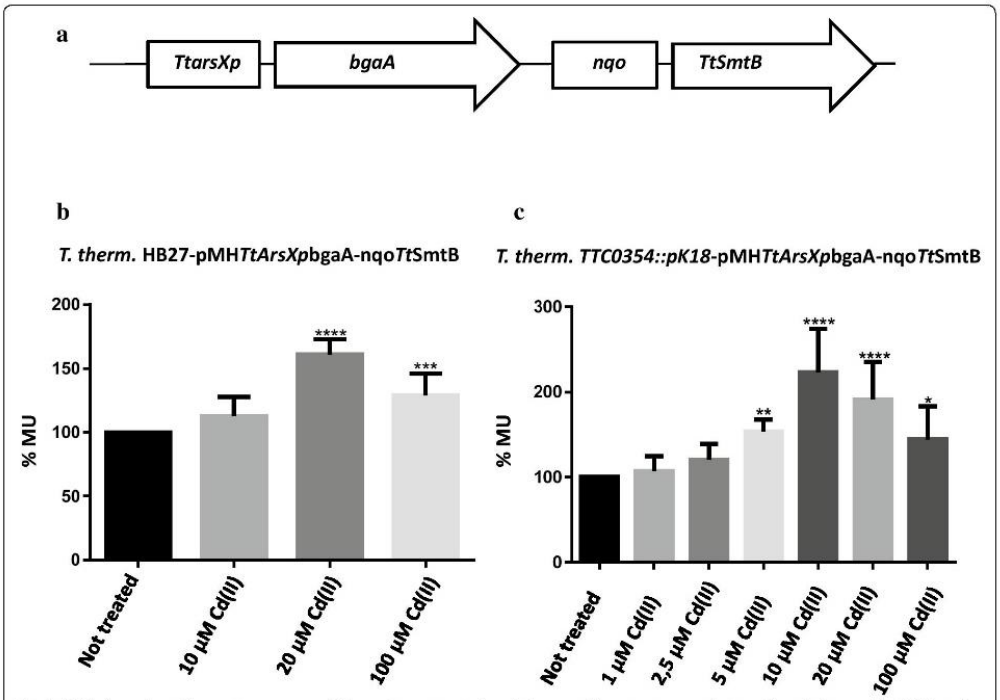


Fig. 4 Cd(II) dependent bioreporter response. a Schematic representation of the recombinant vector used to transform *T. thermophilus* HB27 and *TTC0354::pK18*. b β-gal activity expressed in % MU of *T. thermophilus* HB27-pMH*TtarsXp*bgaA-nqo*Tt*SmtB treated or not with 10, 20 and 100 µM of $CdCl_2$. c β-gal activity expressed in % MU of *T. thermophilus TTC0354::pK18*-pMH*TtarsXp*bgaA-nqo*Tt*SmtB treated or not with 10, 20 and 100 µM of cadmium. Statistical analysis was performed using a Student's t test; significant differences are indicated as: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

too. As also observed for Cd(II), the reporter activity increases to a maximum level [385 ± 22 MU for As(V) and 563 ± 65 MU for As(III)] after which it decreases likely because of the metal toxicity. Overall, this study underlines the advantage of using: (i) an extremophilic microorganism more resistant to stress conditions for the development of biosensor for in situ monitoring of contaminated sites [37]; (ii) a bioreporter system where the interaction between the regulator and its target promoter has been analyzed and defined [38]; (iii) a strain genetically modified more sensitive to detect metals [36]. In addition, these results encourage to improve the newly developed cell-based system for the realization of a robust biosensor for multiple metal detection [39]. In this context, a system to be used as a screening tool to detect metals in environmental samples could significantly help to find metal polluted areas, or to monitor in situ cleanup during bioremediation.

Conclusions

This work reports for the first time the identification of a molecular mechanism responsible for cadmium tolerance of the thermophilic bacterium *T. thermophilus*, taking advantage of the utilization of suitable genetic tools. Interestingly, the system includes part of the machinery (*Tt*ArsX and *Tt*SmtB, membrane efflux ATPase and ArsR/SmtB transcriptional repressor, respectively) used to cope with arsenic; to the best of our knowledge this is the first functional characterization in a bacterium of a common/promiscuous mechanism to defend from both arsenic and cadmium, giving new venues for the understanding of the metal response evolution and the adaptation of environmental thermophilic microorganisms to deal with high concentrations of metals under enhanced solubilization conditions [12]. Notably, from an evolutionary point of view it can be speculated that in the absence of an *ars* operon, cells may have evolved arsenic

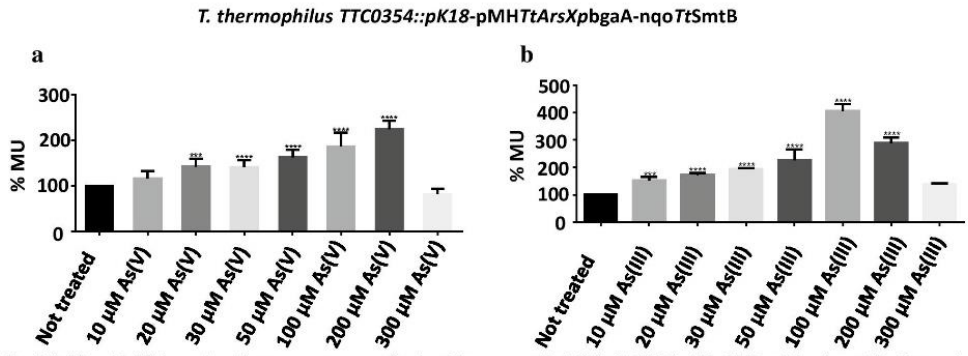


Fig. 5 As(V) and As(III) dependent bioreporter response. a β-gal activity expressed in % MU of *TTC0354::pK18*-pMH*TtarsXp*bgaA-nqo*Tt*SmtB treated or not with 10, 20, 30, 50, 100, 200 and 300 µM of As(V). b β-gal activity expressed in % MU. of *T. thermophilus TTC0354::pK18*-pMH*TtarsXp*bgaA -nqo*Tt*SmtB treated or not with 10, 20, 30, 50, 100, 200 and 300 µM of As(III). Statistical analysis was performed using a Student's t test; significant differences are indicated as: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$

resistance from preexisting metal detoxification systems. It is also plausible that promiscuous detoxification systems have developed according to the hypothesis that in the genomes of thermophilic microorganisms the genetic information is condensed. Moving to biotechnologically relevant applications in the improvement of biosensor field, this study outlines the importance of a detailed characterization of the molecular components (intrinsic promoter activity, repressor/promoter and repressor/metal(s) binding affinities) and points to *T. thermophilus* as suitable chassis cell for design and development of robust metal biosensors. In this context, advantageous modifications can be programmed to increase biosensor sensitivity, selectivity and/or ability to detect metal mixtures.

Additional file

Additional file 1: Table S1. Strains used in this work classified according to their genotype. Table S2. Oligonucleotides used in this work. Table S3. Plasmids used in this work classified according to their features. Figure S1. Growth curves of *T. thermophilus* HB27 transformed with the vector pMH*TtarsXp*bgaA in the absence (circle) and presence of 100 µM Cd(II) (triangle).

Authors' contributions
IA, GG and ALR performed experiments. DL, PC, AB, JB, SB and GF supervised the project. IA and GF drafted the manuscript. All authors read and approved the final manuscript.

Author details
[1] Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Complesso Universitario Monte S. Angelo, Naples, Italy. [2] Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Campus Universidad Autónoma de Madrid, 28049 Madrid, Spain.

Acknowledgements
The authors thank Dr. Emilia Pedone of the CNR Institute of Biostructures and Bioimages in Napoli for helpful discussions during the preparation of the manuscript.

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Consent for publication
Not applicable.

Ethics approval and consent to participate
Not applicable.

Funding
This research was carried out in the frame of the Project "Immobilization of ENzymes on hydrophobin-functionalized NAnomaterials" funded by the University of Napoli Federico II.

Publisher's Note

Received: 12 February 2018 Accepted: 3 May 2018
Published online: 18 May 2018

References
1. Jarup L, Akesson A. Current status of cadmium as an environmental health problem. Toxicol Appl Pharmacol. 2009;238:201–8.
2. Das S, Dash HR, Chakraborty J. Genetic basis and importance of metal resistant genes in bacteria for bioremediation of contaminated environments with toxic metal pollutants. Appl Microbiol Biotechnol. 2016;100:2967–84.
3. Wong K, Ma J, Rothnie A, Biggin PC, Kerr ID. Towards understanding promiscuity in multidrug efflux pumps. Trends Biochem Sci. 2014;39:8–16.

4. Prasad R, Rawal MK. Efflux pump proteins in antifungal resistance. Front Pharmacol. 2014;5:202.
5. Paul S, Moye-Rowley WS. Multidrug resistance in fungi: regulation of transporter-encoding gene expression. Front Physiol. 2014;5:143.
6. Smith AT, Barupala D, Stemmler TL, Rosenzweig AC. A new metal binding domain involved in cadmium, cobalt and zinc transport. Nat Chem Biol. 2015;11:678.
7. Smith AT, Smith KP, Rosenzweig AC. Diversity of the metal-transporting P 1B-type ATPases. J Biol Inorg Chem. 2014;19:947–60.
8. Chandrangsu P, Rensing C, Helmann JD. Metal homeostasis and resistance in bacteria. Nat Rev Microbiol. 2017;15:338.
9. Ma Z, Jacobsen FE, Giedroc DP. Coordination chemistry of bacterial metal transport and sensing. Chem Rev. 2009;109:4644–81.
10. Pedone E, Bartolucci S, Fiorentino G. Sensing and adapting to environmental stress: the archaeal tactic. Front Biosci. 2004;9:2909–26.
11. Bartolucci S, Contursi P, Fiorentino G, Limauro D, Pedone E. Responding to toxic compounds: a genomic and functional overview of Archaea. Front Biosci. 2013;18:165–89.
12. Ranawat P, Rawat S. Metal-tolerant thermophiles: metals as electron donors and acceptors, toxicity, tolerance and industrial applications. Environ Sci Pollut Res 2018;25:4105–33.
13. Fiorentino G, Ronca R, Bartolucci S. A novel E. coli biosensor for detecting aromatic aldehydes based on a responsive inducible archaeal promoter fused to the green fluorescent protein. Appl Microbiol Biotechnol. 2009;82:67–77.
14. 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.
15. Choe S-I, Gravelat FN, Al Abdallah Q, Lee MJ, Gibbs BF, Sheppard DC. Role of Aspergillus niger acrA in arsenic resistance and its use as the basis for an arsenic biosensor. Appl Environ Microbiol. 2012;78:3855–63.
16. Ravikumar Y, prabhu Nadarajan S, Lee CS, Yun H. Engineering an FMN-based iLOV protein for the detection of arsenic ions. Anal Biochem. 2017;525:38–43.
17. PrÛvÛral S, Brutesco C, Descamps EC, Escoffier C, Pignol D, Ginet N, Garcia D. A bioluminescent arsenite biosensor designed for inline water analyzer. Environ Sci Pollut Res. 2017;24:25–32.
18. Peng Z, Yan Y, Xu Y, Takeo M, Yu H, Zhao Z, Zhan Y, Zhang W, Lin M, Chen M. Improvement of an E. coli bioreporter for monitoring trace amounts of phenol by deletion of the inducible σ54-dependent promoter. Biotech Lett. 2010;32:1265–70.
19. Behzadian F, Barjeste H, Hosseinkhani S, Zarei A. Construction and characterization of Escherichia coli whole-cell biosensors for toluene and related compounds. Curr Microbiol. 2011;62:690–6.
20. Chen J, Rosen BP. Biosensors for inorganic and organic arsenicals. Biosensors. 2014;4:494–512.
21. 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. 2013;1834:2071–9.
22. 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.
23. Schurig-Briccio LA, Gennis RB. Characterization of the P1B-Type ATPases present in Thermus thermophilus. J Bacteriol. 2012;194:4107–13.
24. Antonucci I, Gallo G, Limauro D, Contursi P, Ribeiro AL, 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;10:1690–701.
25. VanZile ML, Cosper NJ, Scott RA, Giedroc DP. The zinc metalloregulatory protein Synechococcus PCC7942 SmtB binds a single zinc ion per monomer with high affinity in a tetrahedral coordination geometry. Biochemistry. 2000;39:11818–29.
26. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. J Mol Biol. 2001;305:567–80.
27. Hofmann K. TMbase-A database of membrane spanning proteins segments. Biol Chem Hoppe-Seyler. 1993;374:166.
28. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinform. 2008;9:40.
29. Macindoe G, Mavridis L, Venkatraman V, Devignes M-D, Ritchie DW. HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res. 2010;38:W445–9.
30. Cava F, Zafra O, Da Costa MS, Berenguer J. The role of the nitrate respiration element of Thermus thermophilus in the control and activity of the denitrification apparatus. Environ Microbiol. 2008;10:522–33.
31. 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;64:630–46.
32. Alvarez L, Bricio C, Gómez MJ, Berenguer J. Lateral transfer of the denitrification pathway genes among thermus thermophilus strains. Appl Environ Microbiol. 2011;77:1352–8.
33. Miller JH. Experiments in molecular genetics. Cold Spring Harbor Laboratory: Cold Spring Harbor; 1972.
34. 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;50:6607–21.
35. Smith AT, Ross MO, Hoffman BM, Rosenzweig AC. Metal selectivity of a Cd-, Co-, and Zn-transporting P1B-type ATPase. Biochemistry. 2016;56:85–95.
36. Fernandez M, Morel B, Ramos JL, Krell T. Paralogous regulators ArsR1 and ArsR2 of Pseudomonas putida KT2440 as a basis for arsenic biosensor development. Appl Environ Microbiol. 2016;82:4133–44.
37. Kim HJ, Jeong H, Lee SJ. Synthetic biology for microbial heavy metal biosensors. Anal Bioanal Chem. 2018;410:1191–203.
38. Arruda LM, Monteiro LM, Silva-Rocha R. The Chromobacterium violaceum ArsR arsenite repressor exerts tighter control on its cognate promoter than the Escherichia coli system. Front Microbiol. 1851;2016:7.
39. Amaro F, Turkewitz AP, Martin-Gonzalez A, Gutierrez JC. Whole-cell biosensors for detection of heavy metal ions in environmental samples based on metallothionein promoters from Tetrahymena thermophila. Microb Biotechnol. 2011;4:513–22.

Table S1. Strains used in this work classified according to their genotype.

Strain	Genotype	Source
T. thermophilus HB27	Wild type	DSMZ
T. thermophilus ΔsmtB::kat	*T. thermophilus* HB27 deletion mutant of the *TtsmtB* gene, Kan^r	[23]
T. thermophilus TTC0354::pK18	*T. thermophilus* HB27 insertion mutant defective in *TTC0354*, Kan^r	[23]
E. coli BL21-Codon Plus (DE3) RIL - *Tt*SmtB	F^- ompT hsdS(r_B^- m_B^-) dcm+ Tet^r gal λ (DE3) endA Hte [argU ileY leuW Cam^r]	[23]

Table S2. Oligonucleotides used in this work.

Primer names	Sequences (the restriction sites are underlined) 5'-3'
0354pfwEcoRI	AAAGAATTCCTGTTGGCGGAGGCCCTG
0354prv0354NdeI	AGCCTTCATATGCCCAGGGTAGC
0354footprint fw	CGCCTCGCCGACCGGCA
0354footprint rv	CCATGCCCTCTACCCGGAAG

Table S3. Plasmids used in this work classified according to their features.

Plasmid	Key features	Source
pMHbgaA	β-galactosidase, hygromycin B resistance.	[25]
pMH*TtarsX*pbgaA	*TtarsX* promoter, β-galactosidase, hygromycin B resistance.	This work
pMKpnqo-bgaA	*nqo* promoter, β-galactosidase, kanamycin resistance.	[26]
pMKpnqo-*TtsmtB*	*nqo* promoter, *Tt*SmtB, kanamycin resistance	This work
pMH*TtarsX*pbgaA-nqo*Tt*SmtB	*TtarsX* promoter, β-galactosidase, *nqo* promoter, *Tt*SmtB, hygromycin B resistance	This work
pMHPnorbgaA	*nor* promoter, β-galactosidase, hygromycin B resistance	[27]
pET28b(+)/*Tt*SmtB	T7lac , *lacI* repressor, *Tt*SmtB , kanamycin resistance	[23]

Figure S1.

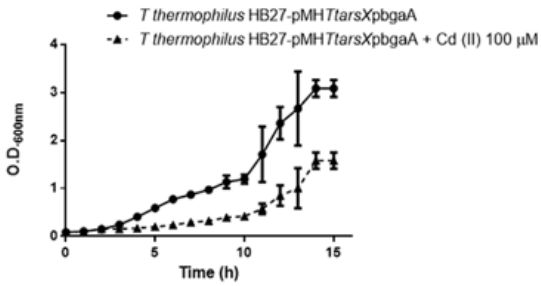


Fig. S1 Growth curves of *T. thermophilus* HB27 transformed with the vector pMH*TtarsX*pbgaA in the absence (circle) and presence of 100 µM Cd(II) (triangle).

Chapter V

Identification and characterization of a unique thermoactive arsenite methyltransferase

In this chapter it's described the approach used to identify other potential partners involved in arsenic resistance acting as *Tt*SmtB interactors. Among partners, a novel putative arsenite methyltransferase was identified using a pull-down approach combined with LC-MS/MS.

The arsenite methyltransferase encoding gene was cloned and expressed in *E. coli* and the recombinant protein, named *Tt*ArsM, was characterized. Indeed, the protein showed arsenite methyltransferase activity *in vitro* and its heterologous expression in *E. coli* increased the arsenite tolerance of the recombinant strain. In order to create strains in which the gene encoding for *Ttars*M was deleted genome editing technology based on CRISPR-Cas9 system was employed. This system was also used to create a genomic *whole-cell* biosensor, which is more stable and sensitive than that previously described.

T. thermophilus Δ*ars*M is more sensitive to arsenite than the wild type in agreement with the role of this protein in arsenite detoxification.

The discovery of *Tt*ArsM, the first thermophilic arsenite methyltransferase, adds a new piece in the puzzle of the arsenic resistance system in *T. thermophilus* HB27. Moreover, the occurrence of interaction between *Tt*SmtB and *Tt*ArsM suggests a post-transcriptional sophisticated role of the two proteins in controlling the concentration of free arsenic in the cells.

Chapter V

Identification and characterization of a unique thermoactive arsenite methyltransferase

Giovanni Gallo [1,*], Ioannis Mougiakos [2,*], Mauricio Bianco [2], Andrea Carpentieri [3], Pietro Pucci [3], Simonetta Bartolucci [1], John van der Oost [2], Gabriella Fiorentino [1]

Affiliations:
1) Department of Biology, University of Naples Federico II, Napoli, Italy
2) Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands
3) Department of Chemical Sciences, University of Naples Federico II, Napoli, Italy

Author information:
* These authors contributed equally

(Paper submitted)

Introduction

Environmental arsenic consists in many inorganic species, associated with minerals or organic substances in the soils. Inorganic trivalent As(III) is more toxic then pentavalent As(V), because the trivalent species react with thiols and protein sulfhydryl groups affecting their function[1]. As result of continual exposure to arsenic, bacteria evolved different arsenic detoxification pathways[2]; in bacteria As(V) is taken by phosphate transport system[3], where it's reduced, by an arsenate reductase[2], in As(III) that can be extruded by an efflux pump or methylated by an arsenite methyltransferase[4,5]. Methylation of arsenic by bacteria occurs in several steps which, starting from As (III) lead to mono, di and tri-methylated forms by S-adenosylmethionine methyltransferase encoded by *ars*M gene[6]. The function of the ArsM proteins was first demonstrated in an hypersensitive strain of *Escherichia coli*, where recombinant expression of an *ars*M gene conferred As resistance. Methyltransferases belong to the enzyme class of transferases (EC 2). They are widespread enzymes in nature, for example, in *Escherichia coli* approximately 2% of all proteins are methyltransferases[7].

All methyltransferases share the need for a cofactor to perform their action. S-Adenosyl-l-methionine the methyl donor used most frequently in biological systems[8]. The methylation reaction occurs through a direct SN2 mechanism with the inversion of the

configuration at the reacting carbon atom[9]. Methyltransferases can be classified either according to the substrate (DNA, RNA, proteins, lipids, carbohydrates, and small molecules) or to the type of atom (element) they methylate[10]. The methyl transfer from SAM to its substrate has a very favourable enthalpy and leads to selective methylation[10]. As a result of their versatile substrates, the roles of these enzymes are different, for example, regulation, signalling, and biosynthesis. Although the catalytic mechanism has been extensively studied and the ArsM coding genes have been found in all kingdoms, it remains unknown how these were distributed. In a lot of bacteria the arsenic resistance system is arranged in an operon[11], but in the thermophilic bacterium, *T. thermophilus* HB27 has an unusual machinery to face arsenic toxicity; in fact, the molecular components are not encoded in a single operon, but they are interspersed in the genome[12] and are: an arsenate reductase *Tt*ArsC[13], an efflux pump *Tt*ArsX [14] and the transcriptional repressor *Tt*SmtB[12,15] ; the latter regulates the expression of all these genes in dependence of arsenic intracellular concentration. In our recent past, we developed a bioreporter system based on the regulation of the expression of the β-galactosidase, as reporter gene, by *Tt*SmtB[14], this system was sensitive to the increasing concentration of As(III), As(V) and cadmium. In this study, a novel arsenic methyltransferase gene (*Ttars*M) that catalyses methylation of inorganic arsenicals was identified from *T. thermophilus* HB27. *Ttars*M was heterologous expressed in *E. coli* and purified for functional characterization of the protein.

In addition, for the present work we have developed a genome editing system based on CRISPR-Cas9 technology, in order to create strains in which the gene encoding for *Ttars*M has been deleted and also to create a genomic *whole-cell* biosensor, which should be more stable and sensitive, as upgrade of our previous work[14].

Results

Discovery of the first thermotolerant arsenite methyltransferase

Recent studies demonstrated that the key transcriptional regulator of the arsenic and cadmium resistance systems in *Thermus thermophilus* HB27 is a metalloregulatory transcriptional repressor that belongs to the ArsR/ SmtB family. *Tt*SmtB detaches from target

DNA regulatory sequences when bound to As(V), As(III) or Cd(II), thus permitting the expression of genes involved in heavy metals tolerance[12,14,15]. Since the members of the ArsR/SmtB family possess a dimerization domain that allows them to form homodimers[16]; we hypothesized that such domain could allow *Tt*SmtB to interact with other intracellular partners. A protein pull-down approach, employing purified *Tt*SmtB and *T. thermophilus* HB27 cell free extracts from cultures exposed to arsenic, followed by mass spectrometry confirmed our hypothesis that *Tt*SmtB interacts with other proteins.

Among the proteins that were identified to interact with *Tt*SmtB, TTC0109 (Accession No. AAS80457, UniProt code Q72LF0) was selected for further investigation as it is highly conserved among the members of the *Thermus* genus. Bioinformatic analysis of the *TTC0109* conserved domains lead to the identification of a thiamine-binding domain at the N-terminus of the protein and a S-adenosylmethionine-dependent methyltransferase domain at the C-terminus. This preliminary analysis indicated that TTC0109 could be a member of the methyltransferase superfamily that potentially uses arsenite as its substrate. To corroborate the hypothesis that TTC0109 is an arsenite methyltransferase, a multiple sequence alignment of characterized prokaryotic arsenite methyltransferases was performed (Fig. 1).

All the proteins included in the alignment contain a typical C-terminal Rossman fold with a characteristic GxGxG motif in a loop region, which presumably interacts with SAM, and a highly conserved acidic residue at the end of the β2 strand of the Rossman fold, which interacts with the ribose moiety of the cofactor[17,18]. In the case of TTC0109, the predicted GxGxG motif is composed of the G114, T115, G116, T117, G118 residues and the D135 aspartic acid is predicted to be the conserved acidic residue. Moreover, all the arsenite methyltransferases possess at least two catalytic cysteines, which are employed to perform arsenic methylation[19]. The alignment revealed a unique conserved cysteine in position 76 of TTC0109, suggesting the occurrence of an unknown arsenic methylation mechanism.

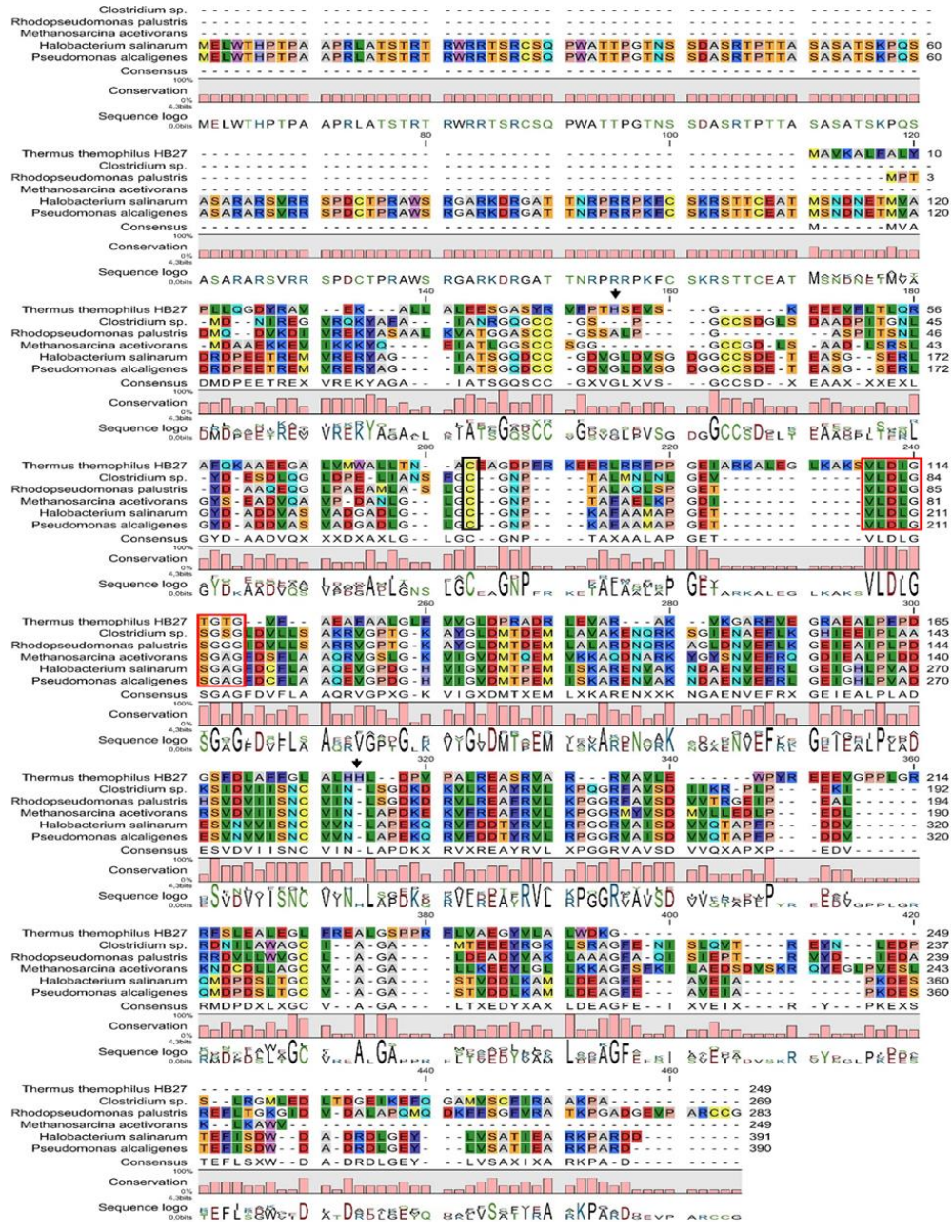


Fig. 1: Multiple sequence alignment of functionally characterized archaeal and bacterial arsenite methyltransferase (ArsM) proteins with the *Thermus thermophilus* HB27 putative arsenite methyltransferase TTC0109. The aligned sequences originate from the ArsM of *Clostridium sp.* BMX[20] (28% identity to TTC0109), *Rhodopseudomonas palustris*[21] (32% identity to TTC0109), *Methanosarcina acetivorans*[22] (29% identity to TTC0109), *Halobacterium*

salinarum[23] (25% identity to TTC0109) and *Pseudomonas alcaligenes*[24] (31% identity to TTC0109). The alignment shown here encompasses the conserved catalytic cysteine (black box); SAM-binding domain (red box) and the two-histidines interacting with As(III) (black arrows).

We further aimed to evaluate, both *in vitro* and *in vivo*, the potential arsenite methyltransferase activity of TTC0109. For this reason, a His-tagged fusion protein was generated by cloning the corresponding gene in pET30b(+) plasmid expressing it in *E. coli* BL21-CodonPlus (DE3)-RIL and purifying to homogeneity the recombinant protein (about 3 mg from 1 liter of culture) through a single purification step (Fig. 2 A).

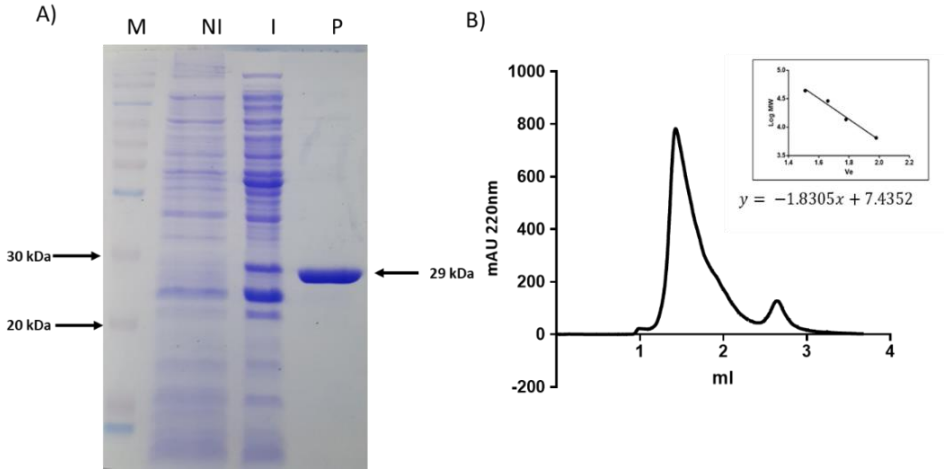


Fig. 2: A) Purification of recombinant *Tt*ArsM. Lane M, protein marker. Lane NI, protein extract from non-induced cells. Lane I, protein extract from induced cells. Lane P, pure protein after His-trap chromatography. **B)** Size exclusion chromatogram of *Tt*ArsM showing an elution volume of 1.43 ml corresponding to a dimeric form.

A colorimetric coupled enzyme assay was set up to detect the potential TTC0109 arsenite methyltransferase activity at 50°C, using saturating concentrations of SAM and As(III) and increasing concentrations of TTC0109. Indeed, the specific arsenite methyltransferase activity calculated for the TTC0109 purified enzyme, under the tested conditions, was 4 mU/mg. In parallel, the recombinant *E. coli* strain expressing TTC0109 was challenged with different concentrations of arsenite for 24 h to confirm *in vivo* the arsenite detoxification activity. Although the growth rate of the

TTC0109-expressing strain was lower than the control strain when cultured in an arsenic-free medium (Fig. 3, red and blue lines), the TTC0109-expressing strain grew better at 2.5 mM of As(III) than the control strain (Fig. 3 red and blue dashed lines). The *in vitro* and *in vivo* results demonstrate that TTC0109 is a thermotolerant arsenite methyltransferase, and from here onwards TTC0109 will be denoted as *Tt*ArsM.

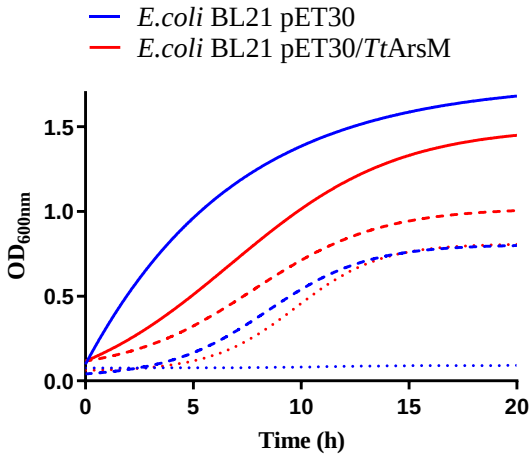


Fig 3: Growth curves of *E. coli* BL21 transformed with pET30b(+) (*E.coli* pET30) in blue and growth curves of *E.coli* BL21 transformed with pET30/*Ttars*M (pET30/*Ttars*M) in red. The cells were cultured in LB medium supplemented with 1 mM IPTG (continuous line), or LB medium supplemented with 1 mM IPTG and either 2.5 mM (dashed line) or 5 mM (dotted line) As(III). The graph reported was obtained by fitting the average OD_{600nm} value of three independent growths in a non-linear regression curve.

*Tt*SmtB; a unique transcriptional and post-translational regulator

*Tt*SmtB is a transcriptional regulator of genes involved in arsenic and cadmium resistance of *T. thermophilus* HB27[12,14,15]. Aiming to discover whether *Tt*SmtB regulates the transcription of *TtarsM* too, EMSA assays were carried out using purified *Tt*SmtB preincubated with a putative regulatory region 108 bp-long, located exactly upstream of the *TtarsM* coding region, in the presence or absence of As (III), As (V), Cd (II) and Sb(III).

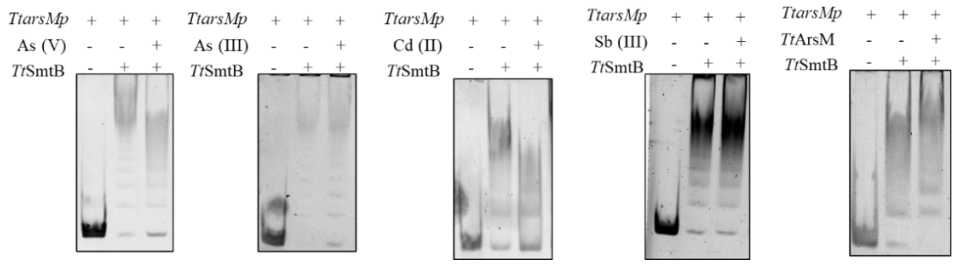


Fig. 4: Electrophoretic mobility shift assay (EMSA) on the predicted promoter of *TtarsM* using *Tt*SmtB and increasing concentrations of protein or heavy metals. Electrophoretic mobility shift assays (EMSA) of *TtarsM* promoter (*TtarsM*p) with purified *Tt*SmtB, in the presence of As(V), As(III), Cd(II), Sb(III) at molar ratio 1:25 or in the presence of *Tt*ArsM at molar ratio 1:1.

It was demonstrated that *Tt*SmtB binds to the provided *TtarsM* promoter sequence, while all the tested metal ions, with the exception of Sb(III), prevented DNA binding (Fig. 4).
As described above in this study *Tt*ArsM and *Tt*SmtB interact *in vivo*. To date, none of the characterized bacterial transcriptional regulators has post-translational interactions with the product of the gene that regulates. Hence, we reasoned to further confirm the surprising protein-protein interaction between *Tt*ArsM and its transcriptional repressor *Tt*SmtB, and to evaluate whether *Tt*SmtB effectors influence this interaction. For this purpose, a Co-IP assay was carried out upon incubation of purified *Tt*ArsM and *Tt*SmtB proteins in the presence or not of As(III), As(V), Cd(II) and Sb(III). After immunoprecipitation with anti *Tt*SmtB antibodies, the detection of his-tagged *Tt*ArsM by anti-His antibodies confirmed that the two proteins interact also *in-vitro* (Fig. 5).

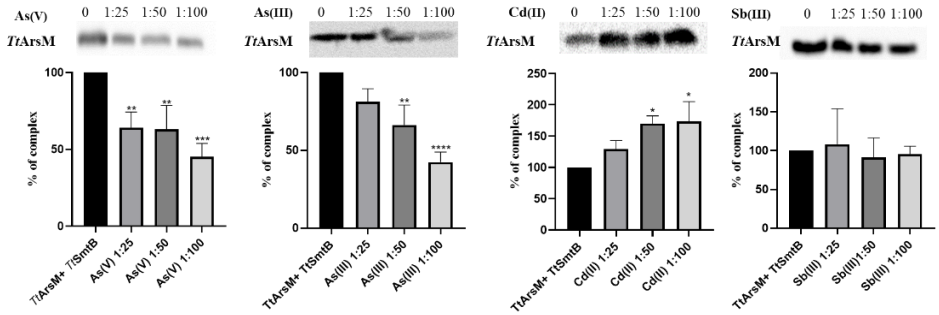


Fig. 5: Co-ImmunoPrecipitation assay using anti-*Tt*SmtB antibodies and revealing with anti-His antibodies. Upper half: Blots of the *Tt*SmtB-*Tt*ArsM co-ImmunoPrecipitation assays upon incubation with increasing molar ratio of As(V), As(III), Cd(II) and Sb(III); *Tt*SmtB and *Tt*SmtB-*Tt*ArsM complexes were immunoprecipitated with anti-*Tt*SmtB antibodies and the abundance of the complexes was revealed with anti-His antibodies. Lower half: Densitometric analysis of the corresponding blots employing the intensity of the *Tt*SmtB-*Tt*ArsM complex as reference. Statistical analysis was performed using a Student's t-Test; significant differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Notably, increasing As(III) and As(V) concentrations negatively affected the stability of the *Tt*SmtB and *Tt*ArsM interaction. Densitometric analysis revealed up to a 3-fold decrease (at 1:100 protein-arsenic ratio) of the *Tt*SmtB and *Tt*ArsM binding stability, indicating that the *Tt*SmtB/*Tt*ArsM complex formation is destabilized in the presence of both arsenite and arsenate. Interestingly, the presence of Cd(II) had the opposite effect, enhancing by up to 2-fold the *Tt*SmtB and *Tt*ArsM binding complex stability (Fig. 5). A negligible effect was observed in the presence of antimony, in agreement with previous data showing that this metal ion is not an effector for *Tt*SmtB (Fig. 5). Altogether these data confirm the formation of a *Tt*SmtB/*Tt*ArsM complex which is destabilized by arsenate and arsenite and is stabilized by cadmium. Finally, these data demonstrate that *Tt*SmtB is the first member of a novel category of bacterial regulators, that perform simultaneous transcriptional and post-translational regulation of the same gene.

Model of the *Tt*SmtB:*Tt*ArsM complex

The surprising and unique protein-protein interaction of a bacterial protein with its transcriptional repressor motivated us to further study this interaction. A *de-novo* tridimensional model of *Tt*ArsM was generated with I-TASSER, followed by molecular docking, aiming to identify the amino acid residues that are potentially involved in the binding of As(III) and SAM, as well as in the interaction with *Tt*SmtB [25].

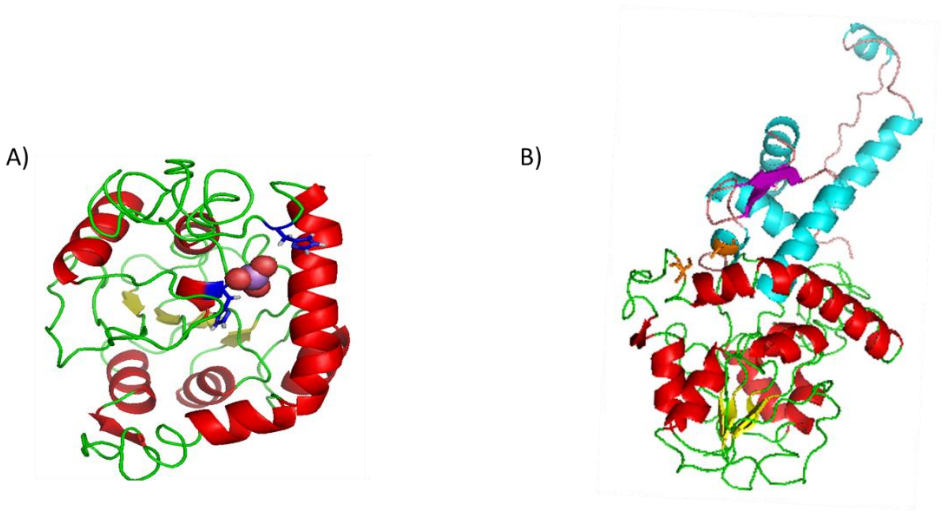
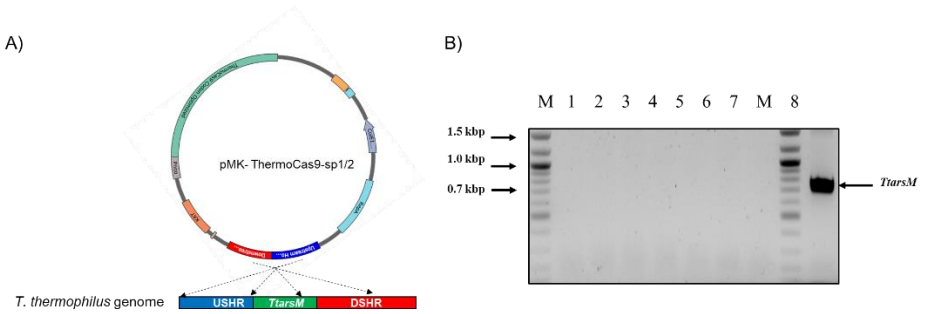


Fig 6: A) Graphical representation of the molecular docking of the TTC0109 3D model with As (III) (the arsenic atom is the purple sphere and the oxygen atoms are the red spheres). The coloured blue amino-acid residues correspond to His-40 and His-179, the responsible residues for the As (III) coordination. **B)** Graphical representation of the molecular docking of the TTC0109 3D model with the *Tt*SmtB 3D model. The docking revealed a possible interaction between the α1 helix of TTC0109 and the α3 helix of *Tt*SmtB. The proximal Cys residues from the two protein that potentially interact (77 of TTC0109 and 64 of *Tt*SmtB) are shown in orange.

Fig. 6 A shows that As(III) can be coordinated by the C77, H40 and H179 residues, strongly supporting the fact that *Tt*ArsM is an arsenite methyltransferase. Notably, H40 is encompassed in a sequence motif (34-YRVFPTHSE-42) which shares 45% identity with a sequence motif at the C-terminus of the *Tt*SmtB metal binding site[15]. Fig. 6 B shows that the interaction between the two proteins could involve the α3 helix of *Tt*SmtB, which contains the conserved N-terminal metal binding site[15], and the α1 helix of *Tt*ArsM, which contains the hypothetical catalytic cysteine. Interestingly, the model highlights that two cysteine residues, one from *Tt*SmtB and one from *Tt*ArsM, are near each other. Moreover, the bioinformatics analysis predicted that *Tt*ArsM forms homodimers and native gel filtration chromatography confirmed this prediction, showing that purified *Tt*ArsM forms a homodimer of approximately 65 kDa (Fig. 2 B).

ThermoCas9-based genome editing and evaluation of *Tt*ArsM activity in *T. thermophilus*

We aimed to further evaluate the contribution of *Tt*ArsM in the As(III) detoxification mechanism of *T. thermophilus* HB27 via the deletion of the *Ttars*M gene from the *T. thermophilus* HB27 genome. Nonetheless, the currently available genome editing tool for *T. thermophilus* is time consuming and inefficient[26]. For this purpose, we reasoned to develop a plasmid-based, homologous recombination(HR) ThermoCas9 counter-selection (CS) genome editing tool[27] for *T. thermophilus* HB27. To date, ThermoCas9 has been used for the development of bacterial editing and silencing tools at temperatures between 37°C and 55°C[27]. Moreover, it has been demonstrated that purified ThermoCas9 is active *in vitro* at temperatures at least up to 70°C, hence it has the potential to edit the *T. thermophilus* genome. We cloned the 3 components of the HR-ThermoCas9 CS system in the pMK18 vector, constructing the pMK-ThermoCas9-spNT editing vector: i) the *T. thermophilus* codon-harmonized version of the *thermocas9* gene under the control of the constitutive *nqo* promoter[28], ii) the sgRNA expressing module -with a non-targeting spacer- under the transcriptional control of the *T. thermophilus* 16S-rRNA promoter, and iii) the HR-template for the deletion of the *TtarsM* gene, composed of the fused 1kb upstream and 1kb downstream genomic regions of the *Ttars*M gene. We then constructed 2 additional editing vectors, denoted as pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2, substituting the non-targeting spacer of pMK-ThermoCas9-spNT with 2 *Ttars*M targeting spacers (Fig. 7 A). The 2 editing vectors were transformed into *T. thermophilus* cells grown at 70°C, while the recovery and plating phases were performed at 60°C and 65°C respectively, due to the sensitivity of pMK18 at temperatures above 65°C. All the colonies transformed with both targeting vectors were Δ*Ttars*M, upon colony PCR screening with genome specific primers and subsequent sequencing verification (Fig. 7 B-C). On the contrary, all the colonies transformed with the pMK-ThermoCas9-spNT non-targeting vector were wild type. In sum, a, highly efficient 65°C, ThermoCas9-based genome editing tool for *T. thermophilus* was developed and a *T. thermophilus* Δ*Ttars*M strain was constructed in less than a week-time.



C)

Plasmid	Recovery	Plating	Tested	Δ*ars*M	Efficiency
pMK-ThermoCas9-sp1	70°C	60°C	17	17	100%
pMK-ThermoCas9-sp2	70°C	60°C	9	9	100%
pMK-ThermoCas9-sp1	70°C	65°C	4	4	100%
pMK-ThermoCas9-sp2	70°C	65°C	9	9	100%

Fig. 7: A) Schematic overview of the basic pMK-ThermoCas9-sp construct. The thermocas9 gene was introduced either to the pMK vector. Homologous recombination flanks were introduced 1 kb upstream (blue) and downstream (red) region of the gene of interest (*Ttars*M) in the targeted genome. A sgRNA-expressing module was introduced downstream the *thermocas9* gene. **B)** Agarose gel electrophoresis showing the resulting products from genome-specific PCR on seven colonies from the ThermoCas9-based *Ttars*M deletion process from the genome *T. thermophilus* HB27. All seven colonies contained the Δ*arsM* genotype. **C)** Mutagenesis efficiency of ThermoCas9 system in *T. thermophilus* HB27.

The so-far described results support the hypothesis that *Tt*ArsM plays an important role in the arsenic detoxification of *T. thermophilus*. To confirm this hypothesis, we determined the minimum inhibitory concentration (MIC) values of As(III) and As(V) for the *T. thermophilus ΔTtarsM* and wild-type strains, as previously described[49]. As shown in Table 1, the resistance of the *T. thermophilus ΔarsM* strain to As(III) was significantly lower compared to the wild-type strain, while the resistance of both strains to As(V) was similar. In sum, we confirmed that *Tt*ArsM is an arsenite methyltransferase involved in the detoxification of *T. thermophilus* HB27.

	MIC		
Strain	**As(III)**	**As(V)**	**Cd(II)**
T. thermophilus HB27*	40 mM	44 mM	3 mM
T. thermophilus ΔarsM	18 mM	42 mM	3 mM
T. thermophilus ΔarsX(syfp)	3 mM	3 mM	0.2 mM
T. thermophilus ΔarsM-ΔarsX(syfp)	0.5 mM	2 mM	0.2 mM

Table 1: MIC of *T. thermophilus* wild type and mutant strain to arsenic. *(values already reported in Antonucci et al.[12-14])

Development of a fluorescent bioreporter for arsenic and cadmium detection

In a previous study we developed a *T. thermophilus* HB27- based fluorescence bioreporter system for toxic metal ions. Nonetheless, the plasmid-based nature of this bioreporter negatively affected the stability and sensitivity of the system, and required antibiotic selection[14]. Hence, we reasoned to create a stable fluorescence reporter system in order to overcome these problems. For this purpose, we employed the ThermoCas9-based editing tool to exchange the *TtarsX* gene in the genomes of the *T. thermophilus* HB27 wild-type and Δ*arsM* strains with a sYFP coding sequence[29], obtaining the *T. thermophilus* HB27 Δ*arsX* (*syfp*) and *T. thermophilus* HB27 Δ*arsM*-Δ*arsX* (s*yfp*) strains.

We determined the MIC values of the Δ*arsX* (*syfp*) and Δ*arsM*-Δ*arsX* (s*yfp*) strains on As(V), As(III) and Cd(II) and compared them to other *T. thermophilus* HB27 strains. The MIC values of the Δ*arsX*(*syfp*) strain were similar to those of the previously constructed *T. thermophilus* Δ*arsX* (also denoted as *TTC0354::pK18*) strain[12,15], demonstrating that the *arsX* to s*yfp* substitution and the sYFP expression does not impose a burden on cells' viability. On the contrary, the MIC value of the Δ*arsM*-Δ*arsX* (*syfp*) double mutant on As(III) was further decreased when compared to the single Δ*arsX* (*syfp*) and *ΔarsM* mutant strains (Table 1), demonstrating that there is an *in vivo* synergy between *Tt*ArsM and *Tt*ArsX that enhances arsenic resistance.

We subsequently challenged the Δ*arsX* (*syfp*) and Δ*arsM*-Δ*arsX* (s*yfp*) fluorescence bioreporter strains with varying As(V), As(III) or Cd(II) concentrations at 70°C, in order to evaluate the sensitivity and

the robustness of the developed systems. The fluorescence intensity of the cultures (normalized for the corresponding OD_{600} values) was measured 1 hour after the addition of the metals. The background fluorescence for both bioreporter strains was low when cultured in the absence of metal ions, indicating low levels of "leakiness" for the system. Moreover, increasing concentrations of all the metal ions tested resulted in a dose-dependent increase (though not linear) in the *yfp* expression for both bioreporter strains. Interestingly, the emitted fluorescence from the double mutant Δ*arsM*-Δ*arsX* (s*yfp*) strain was significantly enhanced even at concentrations as low as 0.5 µM Cd(II) or 1 µM As(III)/As (V) (Fig. 8). In sum, the Δ*arsM*-Δ*arsX* (s*yfp*) bioreporter strain substantially decreased the lower metal detection limit by to 10 times for Cd(II) and 20 times for As(V) and As(III) when compared to the previously developed plasmid-based system [14].

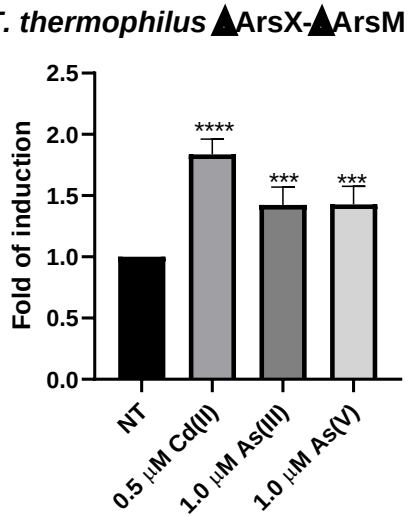


Fig. 8: Fluorescent As(III) bioreporter response of *T. thermophilus* ΔArsM-ΔArsX (*syfp*). Statistical analysis was performed using a Student's t-Test; significant differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$.

Discussion

The distribution of arsenic resistance genes reflects the ubiquitous presence of arsenic in nature, but these genes are also present in microorganisms that live in ideally arsenic-free environments.

The presence of these genes on plasmids and genomic islands underlines the involvement of horizontal gene transfer[11]. Usually in bacteria the arsenic resistance genes are arranged in ars operons, regulated by the protein ArsR, a member of the SmtB/ArsR family of metalloregulatory proteins. ArsR is a trans-acting transcriptional repressor protein that binds to the promoter region of ars operons. The interaction of ArsR with arsenite dissociates the repressor protein from the DNA thus allowing transcription of the operon. In our recent work, we have characterized the role of *Tt*SmtB in *T. thermophilus* HB27 as member of the SmtB/ArsR family[12,15]; that regulates arsenic resistance genes that are unusually arranged in the genome of *T. thermophilus* HB27.

In order to gain thorough insights into the mechanism of *Tt*SmtB regulation of arsenic tolerance we applied a proteomic approach to identify in *T. thermophilus* HB27 cells challenged with As(V) or As(III), putative *Tt*SmtB interacting proteins with a role in arsenic metabolism/detoxification. In this study, we discovered a new arsenite methyltransferase, *Tt*ArsM, as protein interacting with *Tt*SmtB.

The arsenite methyltransferases (ArsM; EC 2.1.1.137) are enzymes which catalyze the methylation of As(III) into monomethyl-arsenite MMA(III) or in DMA(III); in this work we identified and characterized the first thermophilic ArsM from bacteria. In particular, the presence of this protein in *T. thermophilus* HB27 highlights the presence of a new protein involved in the arsenic resistance system and motivates the high resistance of this bacterium to high concentrations of arsenic.

Starting from the *in vitro* characterization of this enzyme, its methyltransferase activity appeared to be evident but to obtain a *in vivo* data, we decided to generate a ΔArsM mutant strain, which is devoid of the coding gene for arsenite methyltransferase. To pursue this goal we developed a new editing system based on the CRISPR-Cas9 technology because the currently employed genome editing tools for *T. thermophilus* are well established but time consuming [12]. CRISPR-Cas based technologies have been extensively applied on many different bacteria for genome engineering purposes, substantially accelerating metabolic engineering studies. Nonetheless, till date only one thermotolerant CRISPR associated endonuclease, ThermoCas9, has been used for engineering purposes at temperatures above 37°C and up to 55°C[27,30]. Meanwhile, it was shown that ThermoCas9 can efficiently introduce

double stranded DNA breaks *in vitro* at temperatures as high as 70°C, motivating the development of a *T. thermophilus* ThermoCas9-based genome editing tool, designed to knockout *Ttars*M at 65°C. With this genetic editing system, it was possible, not only to obtain and characterize the mutant strain *T. thermophilus* ΔArsM, but also to develop a bioreporter system (plasmide free) for monitoring arsenic and cadmium [*T. thermophilus* ΔArsM-ΔArsX (*syfp*)], able to detect 0.5 µM Cd (II) and 1 µM As (III) and As (V). This result is an improvement on our previous system[14], with an increase of about 10 times in sensitivity to these two heavy metals. Furthermore, the interaction between *Tt*SmtB and *Tt*ArsM verified with Pull-Down assays was then confirmed by Co-Ip experiments, which also revealed the biological role of this interaction. Using Co-Ip experiments in the presence of different heavy metals, it is possible to hypothesize that the interaction between *Tt*SmtB and *Tt*ArsM has a regulatory function on *Tt*ArsM. In fact, in the presence of As (V) and As (III), the complex is highly destabilized, thus releasing *Tt*ArsM allowing the methylation of arsenite and detoxification. When the *Tt*SmtB-*Tt*ArsM complex is compared in the presence of Cd (II), a metal that is recognized by *Tt*SmtB[15] and detoxified by *Tt*ArsX[14], the result is an increase in the stability of the complex. This is since cadmium is not a metal recognized by *Tt*ArsM, therefore the interaction with *Tt*SmtB leads to the complete inactivation of methyltransferase. In this way any *Tt*ArsM activity is inhibited, leading to an energy saving by the cell, which in this way can divert the redox recycling path (which usually accompanies the activity of arsenite methyltransferase) due to the resistance to oxidative stress induced by the cadmium.
Another level of regulation that *Tt*SmtB operates on *Tt*ArsM, is at the transcriptional level. In fact, like all genes involved in arsenic and cadmium resistance, the *Tt*ArsM promoter is recognized by *Tt*SmtB, which in the presence of As (III), As (V) and Cd (II) detaches from the promoter thus allowing the expression of the protein. So, for the first time there is a double level of regulation operated by *Tt*SmtB, at the transcriptional and post-translational level. As has already been described for other bacterial systems, it is evident that the interaction between the transcription factor *Tt*SmtB and *Tt*ArsM, can be very important to maintain an energy balance of the cell in a condition of oxidative stress induced by the presence of heavy metals.

Materials and methods

T. thermophilus HB27 growth and cell-free extract preparation

T. thermophilus HB27 (purchased from DSMZ) was grown aerobically at 70°C in TM medium as described[13]; cells grown up to 0.5 OD_{600nm} (350 ml) were treated or not with 8 mM $NaAsO_2$ or 12 mM NaH_2AsO_4 (Sigma, referred to throughout this article as As(III) and As(V) respectively) and harvested at 0 and 60 min after their addition. Cells were lysed by sonication (10'cycle: 30''on/30''off) in 20 mM Na_3PO_4 pH 7.5 and centrifuged; a protease inhibitor cocktail (Thermo Scientific) was added to prevent protease activity in the obtained protein extracts.

Purification of recombinant *Tt*SmtB

*Tt*SmtB was purified from *E. coli* BL21-CodonPlus (DE3)-RIL cells transformed with pET28/*Ttsmt*B (E. coli BL21-*Tt*SmtB) as already described [12]. At each purification step an inhibitor protease cocktail (Roche) and 1 mM DTT were added to prevent protease activity and oxidation. *Tt*SmtB was digested with thrombin (Sigma-Aldrich) and treated with ultrafiltration in order to remove the histidine tag. The complete removal was assessed by Western blotting using anti-poly-Histidine-Peroxidase antibodies (Sigma-Aldrich)[15].

Immobilized Metal Affinity Chromatography (IMAC) and pull-down

Purified 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, 20 mM imidazole, pH 7.5 for 16 h at 4 °C and then washed three times with the same buffer to remove unbound protein. *T. thermophilus* HB27 protein extracts, prepared as described before (not treated, treated with As(V) or treated with As (III)) were incubated with the functionalized resin (Ni^{2+}-NTA/*Tt*SmtB) for 16 h at 4 °C under stirring conditions; the resin was extensively washed and the interacting proteins were eluted with the 20 mM Na_3PO_4, 0,5 M NaCl, 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 three *T. thermophilus* protein extracts.

In-situ hydrolysis and LC-MS/MS analysis

The eluted fractions were analyzed by 15 % SDS-PAGE and hydrolyzed in situ for mass spectrometry analysis. Specifically, mono-dimensional SDS-PAGE gel was colored with Coomassie

Brilliant Blue, the revealed bands were cut and de-stained with 100 µL of 0,1 M ammonium bicarbonate (AMBIC) and 130 µL of acetonitrile (ACN). Each band was subjected to in-situ hydrolysis incubating at first for 1,5 h at 4 °C and then for an additional 16 h at 37°C in 10 mM AMBIC, 0,1 µg/µl trypsin. The hydrolysis reactions were stopped by adding acetonitrile and 0,1% formic acid. The samples were then filtered and dried in a Savant vacuum centrifuge. The filtered samples were analyzed with mass spectrometry LC-MS/MS using a Q-TOF instrument. 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 nano-spray source of the mass spectrometer. The latter is represented by a hybrid Q-TOF spectrometer (MS CHIP 6520 QTOF) equipped with a chip (Agilent Technologies).

In-silico analysis

Putative *Tt*SmtB protein interactors were identified using the Mascot software (http://www.matrixscience.com/search_form_select.html). Among interactors, ORF TTC0109, (UniProt code Q72LF0) herein named *Tt*ArsM was 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).
Models of *Tt*ArsM were generated through I-TASSER[31] (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) using as input the complete sequence of *Tt*ArsM. Molecular docking between *Tt*ArsM and As(III) was obtained using Hex Protein Docking server[32]; 100 rigid body docking solutions were generated and the best 10 obtained refined by energy minimization. The proposed model for the metal ion docked into *Tt*ArsM is the structure with the smallest distance between As(III) and histidine (4.33 Å from His40 and 5.77 Å from His179 in *Tt*ArsM model). To better understand how happens the interaction between *Tt*SmtB and *Tt*ArsM, molecular docking was obtained using ClusPro server[34-36]. The proposed model for the interaction between *Tt*SmtB and *Tt*ArsM was the one with the best hydrophobic-ionic ratio.

Cloning, expression, and purification of recombinant *Tt*ArsM

The gene encoding *TtarsM* was amplified by PCR from *T. thermophilus* HB27 genomic DNA, using Taq DNA polymerase (Thermo Fisher Scientific) and the primers containing the NdeI (*ArsMfw,* Supplementary Table 1) and HindIII (*ArsMrv,* Supplementary Table 1) sites at the 5' and 3' ends, respectively. Amplified fragments were purified, digested, and cloned into NdeI/HindIII-digested pET30b(+) vector (Novagen). For protein expression, *E. coli* BL21-CodonPlus(DE3)-RIL cells transformed with pET30b(+)*/TtArsM* were grown in LB medium containing kanamycin (50 µg/ml), chloramphenicol (33 µg/ml); when the culture reached 0.7 OD_{600nm}, protein expression was induced by the addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) and the bacterial culture was grown for 16 h at 37°C. Cells were harvested and lysed by sonication in 20 mM NaP pH 7.4, 50 mM NaCl and 20 mM Imidazole. The recombinant protein was purified to homogeneity through HisTrap HP (1 ml; GE Healthcare) connected to an AKTA Explorer system (GE Healthcare). The fractions containing His-*Tt*ArsM were pooled and dialyzed for 16 h at 4°C against 20 mM NaP pH 7.4. To prevent protease activity an inhibitor cocktail (Roche) was added at each step. The identity of the purified protein was assessed by mass spectrometry and protein aliquots stored at -20°C.

*Tt*ArsM quaternary structure assessment

To determine the quaternary structure of *Tt*ArsM, the native molecular mass was determined by loading 500 µg of the purified protein to an analytical Superdex PC75 column (3.2 by 30 cm) connected to an AKTA Pure system, in 50 mM Tris-HCl, pH 7.5, 0.2 M KCl. 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 described[13].

Methyltransferase activity assay

*Tt*ArsM activity was measured using the enzyme coupled assay SAM510: SAM Methyltransferase Assay Kit (G-Biosciences) according to the manufacturer's suggestions with some modifications (temperature, SAM concentration and time). A typical reaction mixture containing 200 µM As(III), 800 µM SAM, 1x reaction buffer and 10 µg of enzyme in a final reaction volume of 115 µL was incubated for 1h at 50°C in a Synergy™ HTX Multi-Mode Microplate

Reader (BioTek) and enzyme activity was measured following the increase in absorbance at 510 nm due to the formation of the chromogen 3,5-dichloro-2-hydroxybenzene sulfonic acid (DHBS with ε_{mM}=15.0). As a negative control, the same reaction mixture was tested with 10 µg of *Tt*ArsC[13]. To determine the saturating concentrations of As(III) and SAM, they were alternatively varied from 50 µM to 300 µM and from 200 µM to 1.2 mM, respectively.

Co-Ip assay

Co-immunoprecipitation assay was performed to verify the interaction between recombinant *Tt*SmtB and *Tt*ArsM using *Tt*SmtB and His-tag antibodies (Sigma-Aldrich); 5 µg of *Tt*SmtB were incubated in the absence or presence of As(III), As(V), Cd(II) and Sb(III) in increasing molar ratio (1:25, 1:50 and 1:100) for 10 min at 60°C, then *Tt*ArsM (5 µg) was added in 1 ml of Co-Ip buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10% glycerol and 0.1% Triton X-100) and incubation was performed at 4°C for 2 h in continuous rotation. As controls, *Tt*ArsM (5 µg), *Tt*SmtB (5 µg) and *Tt*GalA[73] were also incubated alone. All the samples were immunoprecipitated with 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 continuing incubation for 16 h at 4°C. The immunocomplexes were washed with Co-Ip buffer and analysed by Western Blot on 15 % SDS-PAGE, using PVDF membranes (Millipore), anti-poly-Histidine-Peroxidase antibody (Sigma-Aldrich) diluted 1:10000 and following already described procedures[37]. Densitometric analysis was performed using the ImageJ tool, considering as 100% the band intensity corresponding to the sample containing *Tt*ArsM and *Tt*SmtB; each experiment was performed in technical and biological triplicates; statistical analysis was performed using a Student's t-Test and significative differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Electrophoretic mobility shift assay (EMSA)

To determine if *Tt*SmtB bound to the promoter region upstream *TtarsM*, electrophoretic mobility shift assays (EMSA) were performed. The *TtarsM* promoter region was amplified by PCR using specific primer pair: *ArsMp Fw* and *ArsMp Rv,* going from position -105 to position +3 from the ATG translation codon (Table S2). EMSA reactions were set up as described[38] using 3 µM of proteins pre-

incubated or not with As (III), As (V), Cd (II) and Sb(III) at molar ratios of 1:25 (considering *Tt*SmtB as a dimer). In order to determine if the complex *Tt*SmtB-*Tt*ArsM was able to bind to *TtarsM* promoter region, EMSA reactions were set up using 3 µM of each protein in 1:1 *ratio*.

Arsenic tolerance in *E. coli*

E. coli BL21-CodonPlus (DE3)-RIL, transformed with pET30/TtArsM or with empty pET30 (control), were inoculated into 10 ml of LB medium at 37°C for 16 h at 180 rpm. Cells were then diluted in 50 ml at 0.08 OD up to 0.6 OD_{600nm}, then *TtarsM* expression was induced with 1 mM IPTG for 3h. After that, the culture was diluted to 0.05 OD_{600nm} (in fresh LB medium supplied with 1 mM IPTG, kanamycin (50 µg/ml) and chloramphenicol (33 µg/ml) and 1 mL transferred in a 24-well dish containing 2.5 mM or 5 mM As(III). The growth was monitored measuring the increase in absorbance at 600 nm at 30 min intervals for 20 h at 37°C in a Synergy™ HTX Multi-Mode Microplate Reader (BioTek). Minimal inhibitory concentration (MIC) was determined as the lowest concentration of As(III) (tested in a range from 2.5 mM to 7.0 mM) that completely inhibited the growth of *E.coli* expressing or not *TtarsM* after 16 hours of incubation in LB medium at 37°C. The values reported are the average of three independent experiments.

Construction of Δ*arsM* mutant

For the realization of a Δ*arsM* mutant of *T. thermophilus* HB27, a new genome-editing strategy based on a thermophilic CRISPR-Cas9 system was used[27]. All the primers and plasmids are listed in Table S1; the cloning procedure and the control plasmids are described in Table S3. In detail, a plasmid based on the pMK-Pnqo-sYFP backbone (Table S2) was developed, containing: a codon-harmonized version of the gene encoding *thermoCas9* (synthesized by Twist Bioscience) under the control of the constitutive *nqo* promoter[28], the *sg*RNA sequence between 16S rRNA promoter of *T. thermophilus* and Rho-independent terminator of *Geobacillus thermodenitrificans* T12[27], two different 23 bp spacers targeting at different positions of the *arsM* gene and a 1.9 kbp homologous recombination fragment consisted of the fused 907 bp upstream and 1 kbp downstream genomic regions of the *TtarsM* gene; as a control, a plasmid containing non-targeting spacer was used. The fragments for assembling the plasmids were obtained via PCR amplification

with the Q5 High-Fidelity DNA polymerase (NEB) and were designed with 25 nucleotide-long complementary overhangs for NEBuilder HiFi DNA assembly (NEB). The constructed vectors were sequence verified (Macrogen Inc). *T. thermophilus* HB27 cells were transformed as previously described using 300 ng of each plasmid, including the pMK-Pnqo-sYFP vector as transformation-efficiency control[12] Since the optimal conditions for ThermoCas9 genome editing in this organism are not known, the transformed bacteria were incubated for 6 h at 70 °C and plated in TM medium with kanamycin 33 µg/ml (Sigma-Aldrich) and grown at two different temperatures (60 °C and 65°) for 48 h. The colonies were counted and the efficiency of the genome editing process in the recombinant colonies was assessed by colony PCR, using the genome specific BG16545-BG16546 primers (Table S1). The *TtarsM* deletion for the edited cells was further confirmed by PCR on purified genomic DNA (GenomeJET Genomic DNA Purification Kit, Thermo Fisher Scientific) and sequencing (Macrogen Inc).

Arsenic and cadmium resistance in *T. thermophilus*

T. thermophilus Δ*arsM* mutant strain was grown aerobically at 70°C in TM medium as described[12]. Exponentially growing cultures of *T. thermophilus* HB27 (as control) and *T. thermophilus* Δ*arsM* were diluted at 0.08 OD_{600nm} in 50 ml tubes containing increasing concentration of As(III), As(V) and Cd(II) (from 5 mM to 50 mM for arsenic ions and from 0.05 mM to 10 mM for Cd(II)) and grown at 70°C as described [12]. The MIC value was determined as the lowest concentration of As(III), As(V) and Cd(II) that completely inhibited the growth of the strain as evaluated by OD_{600nm} after incubation for 18 h at 70°C. The values reported are the average of three independent experiments.

Bioreporter design

Using the already obtained plasmids as backbone (pMK-ThermoCas9-sp1 and pMK-ThermoCas9-sp2), a new set of plasmids was constructed to obtain the insertion of s*yfp* gene into *Ttars*X. For this reason, a 2 kbp repair fragment containing 1kbp sequence upstream and 1 kbp downstream the *TtarsX* gene sequences was amplified from genomic DNA; with the encoding sequence for sYFP (721 bp) in the middle. Two different 23 bp spacers targeting at different positions of *TtarsX* were designed.

The obtained plasmids were used to transform wild-type *T. thermophilus* and Δ*arsM* as described above. The genome editing process in recombinant colonies was verified by colony PCR using BG16496-BG16497 and BG17398-BG17399 primers. The absence of *TtarsX* and the presence of *syfp*, in recombinant colonies was confirmed by PCR on purified genomic DNA (GenomeJET Genomic DNA Purification Kit, Thermo Fisher Scientific) and DNA sequencing (Macrogen Inc). The obtained bioreporter strains were named *T. thermophilus* ΔArsX (*syfp*) and *T. thermophilus* ΔArsX-ΔArsM (*syfp*), respectively.

Fluorescence measurements

To analyze reporter *sYFP* fluorescence, an overnight culture of the bioreporter strains [*T. thermophilus* ΔArsX (*syfp*) and *T. thermophilus* ΔArsX-ΔArsM (*syfp*)] were diluted to 0.08 OD_{600nm} in TM medium and grown at 70 °C until 0.5 OD_{600nm} and aliquots of 5ml were incubated in the presence of increasing concentrations of As(III), As(V) and Cd(II) (from 0.5 μM to 5 μM) for 1h at 70 °C. After this time 200 μL of growing cells were centrifuged for 5' at 6000 rpm and the pellet was washed twice with an equal volume of PBS 1X. *sYFP* fluorescence intensity was measured adding 50 μl of resuspended culture into a 96-well plate and using excitation and emission wavelengths of 458 nm and 540 nm respectively. Measurements were taken using a Synergy™ HTX Multi-Mode Microplate Reader (BioTek) and fluorescence intensity was normalized for the optical density at 600nm in each well. Each experiment was performed in technical and biological triplicates; statistical analysis was performed using a Student's t-Test; significant differences are indicated as: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

References:

1. Hughes, M. F., Beck, B. D., Chen, Y., Lewis, A. S., and Thomas, D. J. "Arsenic exposure and toxicology: a historical perspective." *Toxicological Sciences* 123.2 (2011): 305-332.
2. Mukhopadhyay, Rita, and Barry P. Rosen. "Arsenate reductases in prokaryotes and eukaryotes." *Environmental health perspectives* 110.suppl 5 (2002): 745-748.

3. Rosenberg, H., Gerdes, R. G., and Chegwidden, K. A. Y. E. "Two systems for the uptake of phosphate in Escherichia coli." *Journal of bacteriology* 131.2 (1977): 505-511.
4. Meng, Y. L., Liu, Z., and Rosen, B. P. "As (III) and Sb (III) uptake by GlpF and efflux by ArsB in Escherichia coli." *Journal of Biological Chemistry* 279.18 (2004): 18334-18341.
5. Liu, Z., Rensing, C., and Rosen, B. P. "Toxicity: resistance pathways for metalloids and toxic metals." *Encyclopedia of Inorganic and Bioinorganic Chemistry* (2011): 1-13.
6. Qin, J., Rosen, B. P., Zhang, Y., Wang, G., Franke, S., and Rensing, C. "Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase." *Proceedings of the National Academy of Sciences* 103.7 (2006): 2075-2080.
7. Dalhoff, C., Hüben, M., Lenz, T., Poot, P., Nordhoff, E., Köster, H., and Weinhold, E. "Synthesis of S-Adenosyl-L-homocysteine Capture Compounds for Selective Photoinduced Isolation of Methyltransferases." *ChemBioChem* 11.2 (2010): 256-265.
8. Cantoni, G. L. "S-Adenosylmethionine; a new intermediate formed enzymatically from L-methionine and adenosinetriphosphate." *Journal of Biological Chemistry* 204.1 (1953): 403-416.
9. Ho, D. K., Wu, J. C., Santi, D. V., and Floss, H. G. "Stereochemical studies of the C-methylation of deoxycytidine catalyzed by Hha I methylase and the N-methylation of deoxyadenosine catalyzed by EcoRI methylase." *Archives of biochemistry and biophysics* 284.2 (1991): 264-269.
10. Begley, Tadhg P. *Wiley encyclopedia of chemical biology*. Vol. 4. Wiley, 2009.
11. Páez-Espino, A. David, Gonzalo Durante-Rodríguez, and Víctor de Lorenzo. "Functional coexistence of twin arsenic resistance systems in P seudomonas putida KT 2440." *Environmental microbiology* 17.1 (2015): 229-238.
12. Immacolata A., Gallo G., Limauro D., Contursi P., Ribeiro A.L., Blesa A., Berenguer J., Bartolucci S. and Fiorentino G. "An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in Thermus thermophilus HB27." *Microbial biotechnology* 10.6 (2017): 1690-1701.
13. Del Giudice, I., Limauro, D., Pedone, E., Bartolucci, S., and Fiorentino, G. "A novel arsenate reductase from the bacterium

Thermus thermophilus HB27: its role in arsenic detoxification." *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics* 1834.10 (2013): 2071-2079.

14. Antonucci, I., Gallo, G., Limauro, D., Contursi, P., Ribeiro, A. L., Blesa, A., Berenguer J., Bartolucci S. and Fiorentino, G. "Characterization of a promiscuous cadmium and arsenic resistance mechanism in Thermus thermophilus HB27 and potential application of a novel bioreporter system." *Microbial cell factories* 17.1 (2018): 78.
15. Gallo G., Antonucci I., Pirone L., Amoresano A., Contursi P., Limauro D., Pedone E., Bartolucci S. and Fiorentino G. "A physicochemical investigation on the metal binding properties of *Tt*SmtB, a thermophilic member of the ArsR/SmtB transcription factor family." *International journal of biological macromolecules* 138 (2019): 1056-1063.
16. Busenlehner, L. S., Pennella, M. A., and Giedroc, D. P. "The SmtB/ArsR family of metalloregulatory transcriptional repressors: structural insights into prokaryotic metal resistance." *FEMS microbiology reviews* 27.2-3 (2003): 131-143.
17. Schubert, H. L., Blumenthal, R. M., and Cheng, X. "Many paths to methyltransfer: a chronicle of convergence." *Trends in biochemical sciences* 28.6 (2003): 329-335.
18. Rao, S. T., and Rossmann, M. G. "Comparison of super-secondary structures in proteins." *Journal of molecular biology* 76.2 (1973): 241-256.
19. Huang, K., Xu, Y., Packianathan, C., Gao, F., Chen, C., Zhang, J., Shen, Q., Rosen, B.P. and Zhao, F.-J. "Arsenic methylation by a novel ArsM As (III) S-adenosylmethionine methyltransferase that requires only two conserved cysteine residues." *Molecular microbiology* 107.2 (2018): 265-276.
20. Wang, P. P., Bao, P., and Sun, G. X. "Identification and catalytic residues of the arsenite methyltransferase from a sulfate-reducing bacterium, Clostridium sp. BXM." *FEMS Microbiol Lett* 362 (2015): 1-8.
21. Jie Qin, B.P. Rosen, Yang Zhang, Gejiao Wang, Sylvia Franke, Christopher Rensing"Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase." *Proceedings of the National Academy of Sciences* 103.7 (2006): 2075-2080.

22. Wang, P. P., Sun, G. X., and Zhu, Y. G. "Identification and characterization of arsenite methyltransferase from an archaeon, Methanosarcina acetivorans C2A." *Environmental science & technology* 48.21 (2014): 12706-12713.
23. Wang, G., Kennedy, S. P., Fasiludeen, S., Rensing, C., and DasSarma, S. "Arsenic resistance in Halobacterium sp. strain NRC-1 examined by using an improved gene knockout system." *Journal of bacteriology* 186.10 (2004): 3187-3194.
24. Zhang, J., Cao, T., Tang, Z., Shen, Q., Rosen, B. P., and Zhao, F. J. "Arsenic methylation and volatilization by arsenite S-adenosylmethionine methyltransferase in Pseudomonas alcaligenes NBRC14159." *Appl. Environ. Microbiol.* 81.8 (2015): 2852-2860.
25. Macindoe, G., Mavridis, L., Venkatraman, V., Devignes, M. D., and Ritchie, D. W. "HexServer: an FFT-based protein docking server powered by graphics processors." *Nucleic acids research* 38. suppl_2 (2010): W445-W449.
26. Carr, J. F., Danziger, M. E., Huang, A. L., Dahlberg, A. E., and Gregory, S. T. "Engineering the genome of Thermus thermophilus using a counterselectable marker." *Journal of bacteriology* 197.6 (2015): 1135-1144.
27. Mougiakos, I., Mohanraju, P., Bosma, E.F. et al "Characterizing a thermostable Cas9 for bacterial genome editing and silencing." *Nature communications* 8.1 (2017): 1-11.
28. Cava, F., Laptenko, O., Borukhov, S., Chahlafi, Z., Blas-Galindo, E., Gómez-Puertas, P., and Berenguer, J. "Control of the respiratory metabolism of Thermus thermophilus by the nitrate respiration conjugative element NCE." *Molecular microbiology* 64.3 (2007): 630-646.
29. Cava, F., De Pedro, M. A., Blas-Galindo, E., Waldo, G. S., Westblade, L. F., and Berenguer, J. "Expression and use of superfolder green fluorescent protein at high temperatures in vivo: a tool to study extreme thermophile biology." *Environmental microbiology* 10.3 (2008): 605-613.
30. Mougiakos, I., Bosma, E. F., Weenink, K., Vossen, E., Goijvaerts, K., van der Oost, J., and van Kranenburg, R. "Efficient genome editing of a facultative thermophile using mesophilic spCas9." *ACS synthetic biology* 6.5 (2017): 849-861.
31. Zhang, Yang. "I-TASSER server for protein 3D structure prediction." *BMC bioinformatics* 9.1 (2008): 40.

32. Vajda, S., Yueh, C., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., and Kozakov, D. "New additions to the C lus P ro server motivated by CAPRI." *Proteins: Structure, Function, and Bioinformatics* 85.3 (2017): 435-444.
33. Kozakov, D., Hall, D. R., Xia, B., Porter, K. A., Padhorny, D., Yueh, C., and Vajda, S. "The ClusPro web server for protein–protein docking." *Nature protocols* 12.2 (2017): 255.
34. Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., and Vajda, S. "How good is automated protein docking?." *Proteins: Structure, Function, and Bioinformatics* 81.12 (2013): 2159-2166.
35. Aulitto, M., Fusco, S., Fiorentino, G., Limauro, D., Pedone, E., Bartolucci, S., and Contursi, P. "Thermus thermophilus as source of thermozymes for biotechnological applications: homologous expression and biochemical characterization of an α-galactosidase." *Microbial cell factories* 16.1 (2017): 28.
36. Fiorentino, G., Ronca, R., Cannio, R., Rossi, M., and Bartolucci, S. "MarR-like transcriptional regulator involved in detoxification of aromatic compounds in Sulfolobus solfataricus." *Journal of bacteriology* 189.20 (2007): 7351-7360.
37. Fiorentino, G., Del Giudice, I., Bartolucci, S., Durante, L., Martino, L., and Del Vecchio, P. "Identification and physicochemical characterization of BldR2 from Sulfolobus solfataricus, a novel archaeal member of the MarR transcription factor family." *Biochemistry* 50.31 (2011): 6607-6621.
38. Consolati, T., Bolivar, J. M., Petrasek, Z., Berenguer, J., Hidalgo, A., Guisán, J. M., and Nidetzky, B. "Biobased, Internally pH-Sensitive Materials: Immobilized yellow fluorescent protein as an optical sensor for spatiotemporal mapping of pH inside porous matrices." *ACS applied materials & interfaces* 10.8 (2018): 6858-6868.

Primer	Sequence	Length	Primer Name
ArsMfw	aaacatatggcggtgaaggcc	21	ArsMfw
ArsMrv	aaaaagctttcccttgtcccaaa	23	ArsMrv
ArsMp Fw	GGAAGGGGAAGCGGTGTTT	19	ArsMp Fw
ArsMp Rv	CGCCCTATACTTGGGGAGATG	21	ArsMp Rv
BG15583	gggttgatgttcccctggaccaaaatggtggtggacctgaaga	43	Forward Backbone part 1
BG15526	ggaaaacgttcttcggggcg	20	Reverse Backbone part 1
BG15527	cgccccgaagaacgttttcc	20	Forward Backbone part 2
BG15528	aagccagggttccacggtggaaaaggtgctccgggtggaagac	43	Reverse Backbone part 2
BG15529	ggaggcggtcttccacccggagcacctttttccaccgtggaaccc	44	Forward Upstream Homologue Arm
BG15580	gggccttgtttcattccaggctggctcaccgccatctccccaa	43	Reverse Upstream Homologue Arm
BG15581	cctatacttggggagatggcggtgagccagcctggaatgaaac	43	Forward Downstream Homologue Arm
BG15582	ggggaggtcttcaggtccaccaccattttggtccaggggaacat	43	Reverse Downstream Homologue Arm
BG15724	gacatcgggatcacctc	17	ThermoCas9 part 1 Forward NO Ends
BG15725	gtcaggccgtactcg	15	ThermoCas9 part 1 Reverse NO Ends
BG15726	atgaagtacaagatcggcctggacatcgggatcacctc	38	ThermoCas9 part 1 Forward With Ends
BG15727	aaacttcacgatgtccaggcccgtcgggttcagggtcaggccgtactcg	49	ThermoCas9 part 1 Reverse With Ends
BG15624	ggcctggacatcgtgaagttt	21	PF ThermoCas9 2
BG15625	tcacagggcctgatcg	17	PR ThermoCas9 2
BG15620	gaacaagaaccttcaggactgg	22	Forward Backbone part 1 for pMK-Pnqo-ThermoCas9-(HA)
BG15621	tgtccaggccgatcttgtacttcatatgcccctccttttcgtg	42	Reverse Backbone part 1 for pMK-Pnqo-ThermoCas9-(HA)
BG15626	cggcgagacgatcaggcccctgtgatgagatccggctgctaacaa	45	Forward Backbone part 2 for pMK-Pnqo-ThermoCas9-(HA)
BG15627	ccagtcctgaaggttcttgttc	22	Reverse Backbone part 2 for pMK-Pnqo-ThermoCas9-(HA)
BG15665	tcatgtaactcgccttgatcgttggttggggtagtccagcac	42	Primer Forward 16*sp*
BG15666	gagggcccacatcaccaacgccctcaacgcccttaacggg	40	Primer Reverse 16*sp* Spacer 1
BG15667	gccttaccgggaggaggaggtggtcaacgcccttaacggg	40	Primer Reverse 16*sp* Spacer 2
BG15669	ccatggggttactgcggatctagtcaacgcccttaacggg	40	Primer Reverse 16S RNA promoter NT Spacer
BG15670	ctagatccgcagtaaccccatgggtcatagttccccgag	40	Primer Forward sgRNA NT Spacer
BG15671	gggcgttggtgatgtgggccctcgtcatagttccccgag	40	Primer Forward sgRNA Spacer 1
BG15672	ccacctcctcctcccggtaaggcgtcatagttccccgag	40	Primer Forward sgRNA Spacer 2
BG15674	gtatggcttcattcagctccggttcgattggctccaatttcttg	43	Primer Reverse sgRNA
BG15675	gaaccggagctgaatgaag	19	Primer Forward Backbone 2 for pMK-Pnqo-ThermoCas9-sgRNA-(HA)
BG15676	caggccgatcttgtacttcat	21	Primer Reverse Backbone 1 for pMK-Pnqo-ThermoCas9-sgRNA-(HA)
BG16268	ctcctctttcgggatgagcc	20	Primer Forward Screening arsM
BG16269	gaggttggcgtgaaggtcc	19	Primer Reverse Screening arsM
BG16418	ctcctccaggggactacag	20	Primer Forward Inside ArsM
BG16419	tatcccttgtcccaaagggc	20	Primer Reverse Inside ArsM
BG16545	ccatgcagtgggacgacac	19	Primer Forward Screening arsM 2.0
BG16546	tcacggagcaaggagaact	19	Primer Reverse Screening arsM 2.0

Table S1 Primers used in this study for PCR reactions.

Final construct	Template	Primers	Product	Size
pMK- ThermoCas9-NT (non reparing)	pMK-Pnqo-ThermoCas9	F BG15622	*thermocas9* region	4.47 kbp
		R BG15677		
	Genome *T. thermophilus* HB27	F BG15665	16s RNA promoter and non-targeting spacer	0.365 kbp
		R BG15669		
	pThermoCas9_Δgoi	F BG15670	Non-targeting spacer and sgRNA module	0.281 kbp
		R BG15674		
	pMK-Pnqo-ThermoCas9	F BG15675	Backbone	3.1 kbp
		R BG15627		
pMK- ThermoCas9-sp1 (non reparing)	pMK-Pnqo-ThermoCas9	F BG15622	*thermocas9* region	4.47 kbp
		R BG15677		
	Genome *T. thermophilus* HB27	F BG15665	16s RNA promoter and spacer 1	0.365 kbp
		R BG15666		
	pThermoCas9_Δgoi	F BG15671	Spacer 1 and sgRNA module	0.281 kbp
		R BG15674		
	pMK-Pnqo-ThermoCas9	F BG15675	Backbone	3.1 kbp
		R BG15627		
pMK- ThermoCas9-sp2 (non reparing)	pMK-Pnqo-ThermoCas9	F BG15622	*thermocas9* region	4.47 kbp
		R BG15677		
	Genome *T. thermophilus* HB27	F BG15665	16s RNA promoter and spacer 2	0.365 kbp
		R BG15667		
	Plasmid containing sgRNA	F BG15672	Spacer 2 and sgRNA module	0.281 kbp
		R BG15674		
	pMK-Pnqo-ThermoCas9	F BG15675	Backbone	3.1 kbp
		R BG15627		
pMK- ThermoCas9-NT	pMK-Pnqo-ThermoCas9-HA	F BG15622	*thermocas9* region	4.47 kbp
		R BG15677		
	Genome *T.*	F BG15665	16s RNA promoter	0.365

	thermophilus HB27	R BG15669	and non-targeting spacer	kbp
	pThermoCas9_Δgoi	F BG15670 R BG15674	Non-targeting spacer and sgRNA module	0.281 kbp
	pMK-Pnqo-ThermoCas9-HA	F BG15675 R BG15627	Backbone	3.1 Kb
	pMK-Pnqo-ThermoCas9-HA	F BG15620 R BG15676	HR template	4.4 Kb
pMK- ThermoCas9-sp1	pMK-Pnqo-ThermoCas9-HA	F BG15622 R BG15677	*thermocas9* region	4.47 Kb
	Genome *T. thermophilus* HB27	F BG15665 R BG15666	16s RNA promoter and spacer 1	0.365 kbp
	pThermoCas9_Δgoi	F BG15671 R BG15674	Spacer 1 and sgRNA module	0.281 kbp
	pMK-Pnqo-ThermoCas9-HA	F BG15675 R BG15627	Backbone	3.1 kbp
	pMK-Pnqo-ThermoCas9-HA	F BG15620 R BG15676	HR template	4.4 kbp
pMK- ThermoCas9-sp2	pMK-Pnqo-ThermoCas9-HA	F BG15622 R BG15677	*thermocas9* region	4.47 kbp
	Genome *T. thermophilus* HB27	F BG15665 F BG15667	16s RNA promoter and spacer 2	0.365 kbp
	pThermoCas9_Δgoi	F BG15672 R BG15674	Spacer 2 and sgRNA module	0.281 kbp
	pMK-Pnqo-ThermoCas9-HA	F BG15675 R BG15627	Backbone	3.1 kbp
	pMK-Pnqo-ThermoCas9-HA	F BG15620 R BG15676	HR template	4.4 kbp

Table S2: Q5 PCR reactions for the construction of all plasmids.

Chapter VI

Identification of a new heavy metal resistant strain of *Geobacillus stearothermophilus* isolated from a hydrothermally active volcanic area in Southern Italy

Microorganisms thriving in hot springs and hydrothermally active volcanic areas are dynamically involved in heavy metal biogeochemical cycles; they have developed peculiar resistance systems to cope with such metals that nowadays can be considered among the most permanent and toxic pollutants for humans and the environment. For this reason, their exploitation is required to unravel mechanisms for the detoxification of toxic metals and to address remediation of heavy metal pollution with eco-sustainable approaches. In this work, we isolated a novel strain of the thermophilic bacterium *Geobacillus stearothermophilus* from solfataric terrains in Pisciarelli, a well-known hydrothermally active zone of the Campi Flegrei volcano located near Naples in Italy through ribotyping, 16SrRNA sequencing and mass spectrometry analyses. The Minimal Inhibitory Concentration (MIC) towards several heavy metal ions indicated that the novel *G. stearothermophilus* isolate is particularly resistant to some metal ions. Genomic, functional and morphological analyses suggest that it is endowed with metal resistance systems for arsenic and cadmium detoxification.

Chapter VI

Identification of a new heavy metal resistant strain of *Geobacillus stearothermophilus* isolated from a hydrothermally active volcanic area in Southern Italy

Rosanna Puopolo[1], Giovanni Gallo[1], Angela Mormone[2], Danila Limauro[1], Patrizia Contursi[1], Monica Piochi[2], Simonetta Bartolucci[1] and Gabriella Fiorentino[1]*

Affiliations:

1. Dipartimento di Biologia, Università degli Studi di Napoli Federico II, Napoli, Italy
2. Istituto Nazionale di Geofisica e Vulcanologia, sezione Osservatorio Vesuviano, Napoli, Italy

* Corresponding author: fiogabri@unina.it

(Paper submitted)

Introduction

Heavy metals are among the most persistent and toxic pollutants. Differently from organic xenobiotics, which can be converted into not harmful substances, they cannot be completely removed from the environment[1]. Anthropogenic sources including mining and various industrial (vehicle emissions, lead-acid batteries, paints) and agricultural activities (fertilizers, aging water supply) led to their increasing accumulation[2-4]. The prolonged exposure to heavy metal is responsible for several human diseases, as documented by the World Health Organization[5]. For example, arsenic and cadmium have been classified as carcinogenic compounds by the *International Agency for Research on Cancer* (IARC) in both humans and animals[6,7], while the exposure to lead is responsible for 3% of cerebrovascular disease worldwide [8]. Therefore, reduction of heavy metal pollution appears among the greatest challenges of the new century[9-11].

Because of massive accumulation of toxic metals into the environment, the majority of microorganisms has developed genetic resistance mechanisms[12-14] and even peculiar metabolic pathways to cope with toxic metals[15-17]. For instance, Gram-positive and Gram-negative bacteria possess arsenic resistance systems encoded by operons carried either on plasmids or on the chromosome; the operon structure most commonly found is made up of three genes (*arsRBC*)[18,19], where *arsR* encodes a trans-acting repressor of the ArsR/SmtB family involved in transcriptional

regulation[20-22], *arsB* encodes an As(III) efflux transporter (ArsB/Acr3)[23], and *arsC* encodes a cytoplasmic arsenate reductase that converts As(V) to As(III), the latter extruded outside the cell[24-26]. Other proteins can also improve the arsenic resistance, such as the arsenite methyltransferases ArsM[27]. The arsenic resistance system of some microorganisms is also responsible for Cd(II) tolerance[28-30]. Cadmium resistance systems generally consist of at least two genes, i.e., *cadC* and *cadA*, the first coding for a helix-turn-helix transcriptional regulator that regulates *cadA* the gene coding for a cadmium-translocating P-type ATPase[31].

Volcanic and geothermal springs also represent a rich source of heavy metal resistant microorganisms, since many metals, such as arsenic, are naturally present in these particular niches[32,33]; it has been reported that microorganisms living in such ecosystems can remodel extreme environments participating in metal geochemical cycles promoting or inhibiting their release from sediments[34]. The interest in the comprehension of the molecular mechanisms underlying heavy metal resistance in extreme environments is growing fast because it is fundamental to develop systems effective towards heavy metal pollution through eco-sustainable approaches[35-39].

In this work, we describe the isolation and characterization of a new thermophilic heavy metal resistant microorganism from the solfataric terrain of Pisciarelli in the Campi Flegrei volcano located near Naples in Italy. The site has extreme environmental conditions in terms of temperature, acidity and As rich geochemistry[40,41].

Results and Discussion

Geochemical characterization of the sampling site

As other similar volcanic systems worldwide, the volcano is an acidic sulfate environment determined by the hot circulation of aggressive sulfurous fluids during the dormant dynamics deriving from mixing between deeply infiltrating meteoric waters and ascending magmatic gases with associated intense rock alteration and concentration of certain elements, such as As[40, 42-46].

However, differently from the diffuse and fumarolic outgassing characterizing the Solfatara crater, the Pisciarelli site is a water dominant environment, showing the formation of boiling pools and water springs and the opening of low-energetic geyser-type vents. The site represents the shallowest portion of a widespread

geothermal system that develops in the subsurface and converts into brines in its deeper roots directly supplied by the magmatic outgas. Due to the increased hydrothermal activity since 2006, the site shows maximum temperatures of ca. 110 °C and up to 260 tons/day of CO_2[47] with abundance of H_2S and presence of minor gaseous species such as CH_4, N_2, H_2 and CO.
At the time of sampling, the bubbling mud pool was at pH 4.8 and 94.3°C and the marginal mud at pH 6.0 and 55.3°C, while surrounding soils were at temperature up to 98-99°C and very acidic pH. These values are in the range known for the area, although lower temperatures were also measured in the mud pool (c.a. 70°C). Also, the mineralogical and chemical features of the sampled materials[40] are those usually determined. Indeed, the mud was typically gray in color and essentially enclosed sulfates (i.e. K- and Al- bearing alunite), sulfides (i.e., Fe- plus S-bearing pyrite), and silica-phases; dried water samples crystallized NH_4-bearing sulfates. The mud is enriched in As (10-20 ppm) and Hg (around 40 ppm) compared to the protolith volcanic deposits, contains few wt% of Fe2O3, ca. 60 ppm of V, 10-20 ppm of Pb, <10 ppm of Co, Ni and Cr, 10-20 ppm of Cu, 1-2 ppm of Tl, and practically lacks of Cd being at <0.1 ppm. Based on Valentino and Stanzione[48], Pisciarelli waters are rich in SO4-2 (1400-7000 mg/l) and NH_4 (500-1000 mg/l), contains F (0.5-30 mg/l), Al (65-20 mg/l) and B (0.1-0.8 mg/l), lacks of carbonate species and of chlorine; the content of As, Hg, Tl, Pb is ca. 40-2000, 40-250, 2-8 and 5-30 µg/l, respectively. The general enrichment in S, NH_4, As and Hg is consistent with the volcanic setting and the magmatic /geothermal outgas support.

Isolation and identification of Geobacillus stearothermophilus
Upon incubation of mud samples taken from the marginal water-poorer portion, cell growth was observed in LB medium at pH 7 and 50°C. Single colonies were isolated by serial dilutions in the same medium. Ribotyping and 16S rRNA sequencing led to the identification of a *Geobacillus* specie. *Geobacilli* were firstly described by Nazina et al. in 2001[49]; they are Gram-positive, endospore-forming, aerobic or facultative anaerobic thermophiles, growing optimally at temperatures between 50°C and 72°C exploitable for various biotechnological applications[50,51]. The isolated strain showed an optimal growth temperature of 60°C. Due to the peculiar taxonomy of *Geobacillus*, where several species are candidates for unification, it was not possible to differentiate

unambiguously the species by 16S rRNA sequencing (see e.g. LTPs123_SSU_tree of the SILVA Living tree project); therefore a MALDI TOF MS analysis was committed to DSMZ, which identified the new isolate as *Geobacillus stearothermophilus*. To the best of our knowledge, the genomes of only three strains of *G. stearothermophilus* have been completely sequenced (Table 1): strain "10" isolated from the Yellowstone hot spring (USA); strain "DSM458" isolated from a sugar beet factory in Austria [52] and strain "B5" isolated from a rice stack in China.

Organism	Strain	Genome Size (Mb)	CG%	Gene	Protein	BioProject
Geobacillus stearothermophilus	10	3.67	52.61	3645	3312	PRJNA252389
Geobacillus stearothermophilus [52]	DSM458	3.46	52.10	3683	3165	PRJNA327158
Geobacillus stearothermophilus	B5	3.39	52.50	3426	3045	PRJNA513473

Table 1: List of *G. stearothermophilus* strains with sequenced genomes, as reported in NCBI genome databank.

Metal ion resistance and antibiotic susceptibility in *G. stearothermophilus*

MICs towards different antibiotics and heavy metals were determined to evaluate the sensitivity and the tolerance of the new *G. stearothermophilus* isolate, respectively. To this purpose, the microorganism was grown in the presence of different heavy metals (see Table 2) and antibiotics (ampicillin, kanamycin, chloramphenicol, tetracycline, hygromycin, bleomycin).

G. stearothermophilus resulted to be sensitive to all the antibiotics tested even at the lowest concentration used; actually, to the best of our knowledge, no antibiotic resistances have been documented, although the genome of *G. stearothermophilus 10* presents a sequence coding for a putative tetracycline MFS efflux protein (locus tag: GT50_RS17520). Interestingly, *G. stearothermophilus* showed high tolerance to As(V), as well as V(V), as reported in Table 2. This is not surprising, considering the similarity in structure of the vanadate ion with the arsenate ion, which is itself like the phosphate ion. Both ions could be taken up by phosphate transport systems[53]. The low tolerance to As(III) in comparison to As(V) (MIC values: 1.9

mM and 117 mM, respectively) could be traced back to the lack of active transport systems for As(III) efflux (see also below). Likewise, the legume symbiont *Sinorhizobium melitoti* is very tolerant to As(V) but is very sensitive to As(III) since it is deficient of As(III) transporter systems[53,54].

Metal ions	mM	±
As (III)	1.90	0.10
As (V)	117	3.00
Cd (II)	0.90	0.10
Co (II)	2.00	0.50
Co (III)	2.75	0.25
Cr (VI)	0.25	0.01
Cu (II)	4.10	0.10
Hg (II)	0.02	0.00
Ni (II)	1.30	0.10
V (V)	128	2.00

Table 2: *G. stearothermophilus* resistance to heavy metal ions.

Consequently, as shown in Fig. 1 the presence of As(V) in the growth medium slowed down *G. stearothermophilus* growth: the generation time changed from 30 min for cells grown in the absence of As(V) to 60 min and 125 min for those grown in the presence of As(V) 25mM and 50 mM, respectively.

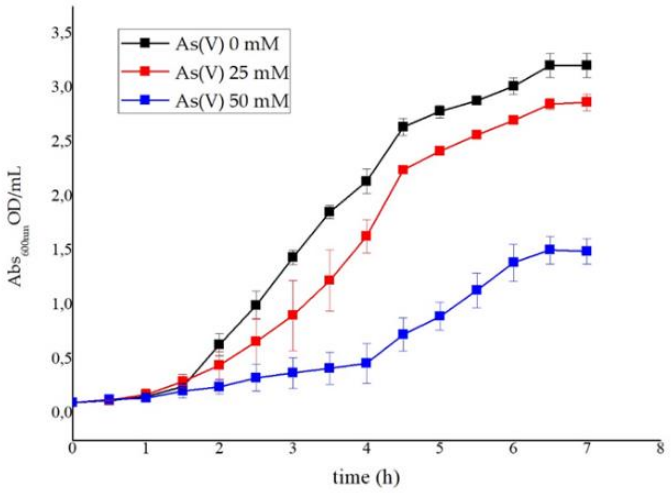


Fig. 1: Growth curves of *G. stearothermophilus* grown in the presence and absence of As(V).

In order to analyze the arsenic transformation activity of *G. stearothermophilus*, a $AgNO_3$ colorimetric method was employed[45, 55-57]. This method is based on the formation of colored precipitates upon reaction of $AgNO_3$ with arsenic; in particular the addition of $AgNO_3$ to the grown cells produces Ag_3AsO_4, a brown precipitate if $AgNO_3$ reacts with As(V) and Ag_3AsO_3, a bright yellow precipitate if $AgNO_3$ reacts with As(III). Fig. 2 shows a brown precipitate indicating that As(V) is the predominant chemical specie and suggesting the inability of *G. stearothermophilus* to extrude As(III), in line with the hypothesis of the lack or low activity of As(III) efflux systems.

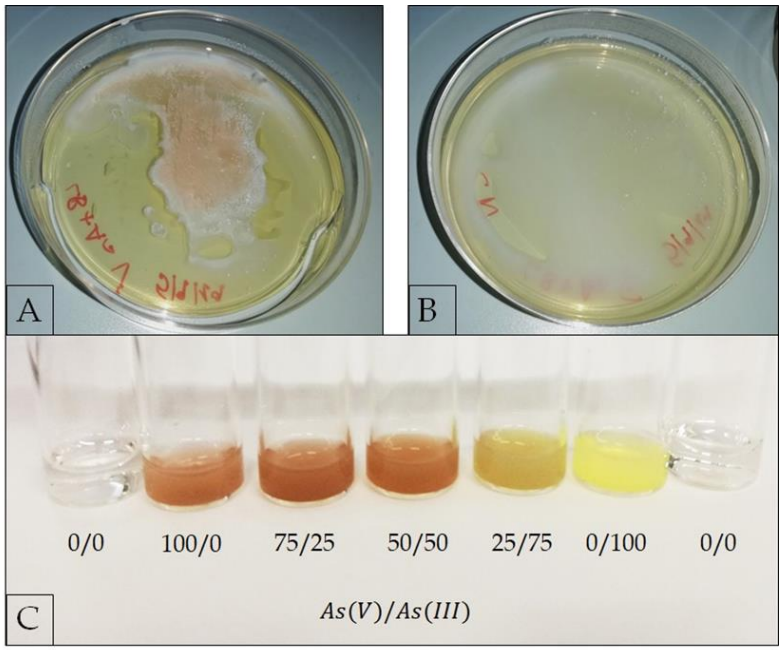


Fig. 2: (A) Silver nitrate test on *G. stearothermophilus* grown on LB agar plate supplemented with 50mM As(V); (B) LB agar plate supplemented with 50mM As(V) (control). (C) Determination of precipitate color as function of As(V)/As(III) ratio (%). The concentration of total arsenic in solution is each sample is 50 mM (i.e. for the ratio 50/50 there are in solution As(V) 25mM and As(III) 25mM).

Bioinformatic analyses

In order to verify if different isolates of *G. stearothermophilus* had arsenic and cadmium resistance systems and to understand if such systems were conserved inside the *specie*, a comparative genomic analysis was carried out on the sequences of three *G. stearothermophilus* genomes available at NCBI (Table 1). The study revealed differences in the abundance and type of putative arsenic and cadmium resistance genes in the genomes analysed (Table 3). All of them contained one conserved copy of *cadC* and *cadA*: the alignment of the corresponding proteins from the three different strains showed a high degree of identity (92%). Regarding arsenic resistance systems, a copy of ArsB/Acr3 arsenite efflux transporters was found in each genome, whereas at least a simple *ars* system encoding the arsenate reductase (*arsC*) in tandem with an ArsR/SmtB transcriptional regulator was found in the genomes of *G. stearothermophilus* 10 and B5 strains (86% of identity of both proteins). On the other hand, the strain DSM 458 encodes a unique arsenate reductase. Taken together, the *in silico* analysis of the genomes showed that the number and type of genes coding for elements involved in arsenic resistance can be variable within the same species and depends on the specific evolutionary adaptation of that particular strain[58].

Strain	Locus	Protein Name	Protein ID
***Geobacillus stearothermophilus* 10** (PRJNA252389)	GT50_RS07590	ArsR: "family transcriptional regulator"	WP_014196895.1
	GT50_RS07505	ArsC: "arsenate reductase (thioredoxin)"	WP_053414123.1
	GT50_RS07510	ArsB: "ACR3 family arsenite efflux transporter	WP_053414999.1
	GT50_RS06280	ArsC: "arsenate reductase family protein"	WP_053413998.1
	GT50_RS12465	Hypothetical CadC "helix-turn-helix transcriptional regulator"	WP_013523123.1
	GT50_RS12470	CadA: "cadmium-translocating P-type ATPase"	WP_053414489.1
***Geobacillus stearothermophilus* DSM 458** (PRJNA327158)	GS458_RS16830	ArsC: arsenate reductase (thioredoxin)"	WP_044745043.1
	GS458_RS16835	ArsB: "ACR3 family arsenite efflux transporter"	WP_095860271.1
	GS458_RS15800	ArsC: "arsenate reductase family protein"	WP_033010367.1
	GS458_RS03695	Hypothetical CadC: "helix-turn-helix transcriptional regulator"	WP_015374106.1
	GS458_RS03700	CadA: "cadmium-translocating P-type ATPase"	WP_095860189.1
***Geobacillus stearothermophilus* B5** (PRJNA513473)	EPB69_RS07030	ArsR: "ArsR/SmtB family transcription regulator"	WP_160270653.1
	EPB69_RS15665	ArsR: "ArsR/SmtB family transcription regulator"	WP_160270373.1
	EPB69_RS15730	ArsR: "ArsR/SmtB family transcription regulator"	WP_160270745.1
	EPB69_RS15655	ArsC: "arsenate reductase (thioredoxin)"	WP_011232692.1
	EPB69_RS15660	Arsenical-resistance protein	WP_011232693.1
	EPB69_RS03435	CadC: "helix-turn-helix transcriptional regulator"	WP_033005425.1
	EPB69_RS03440	CadA: "cadmium-translocating P-type ATPase"	WP_160268846.1

Table 3: List of putative genes for As and Cd(II) resistance in *Geobacillus stearothermophilus* strains.

Analysis of cellular morphology

In order to better define *G. stearothermophilus* morphology, we resolved to analyse cells through TEM. As shown in Fig. 3 cells have a typical bacillar rod shape when they are actively growing.

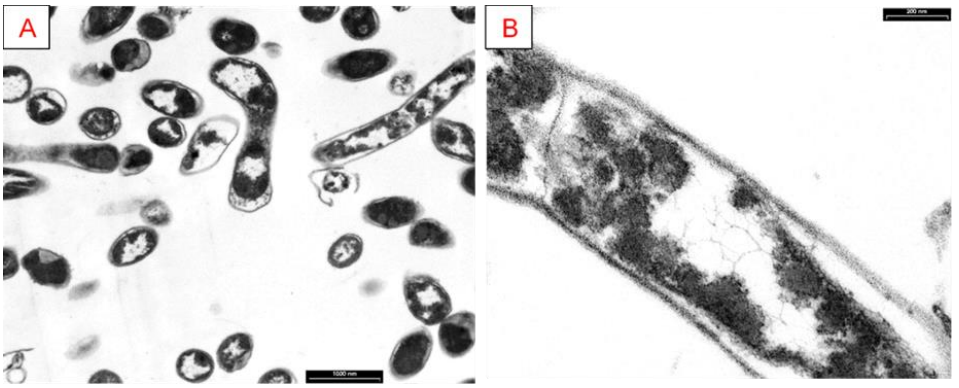


Fig. 3: TEM images of *G. stearothermophilus* in exponential phase at different scales: (**A**) 1000 nm; (**B**) 200 nm.

Moreover, with the aim to verify whether As(V) and Cd(II) had any effect on cell morphology, TEM images were also acquired on samples of *G. stearothermophilus* grown for 16 hours in the presence of As(V) and Cd(II) at concentrations corresponding to the MIC values, and they were compared to images of control cells without heavy metal treatment (Fig. 4). The sections obtained revealed the structure of the cell more clearly in the control cells (panel A of Fig. 4) than in those treated with heavy metals. However, the presence of several cells in division suggests that both As(V) and Cd(II) do not cause significant changes in the cellular structure and cell viability (panel A and B of Fig. 4). Nevertheless, it appeared that the cell wall of *G. stearothermophilus* is influenced by handling both As and Cd. In particular, the cell wall of *G. stearothermophilus* treated with As(V) (panel B of Fig. 4) exhibited abundance of ridges and grooves that can be explained assuming a reduction in cell permeability. Interestingly, Cd(II) treated cells (panel C of Fig. 4) appeared darker and this phenomenon could be ascribed to the ability of *G. stearothermophilus* to adsorb Cd(II), as also reported by Hetzer et al[59].

In conclusion, electron microscopy analyses highlighted that the cell morphology of *G. stearothermophilus* changes the in presence of

As(V) and Cd(II), but further investigations are needed to understand at molecular level their resistance mechanisms.

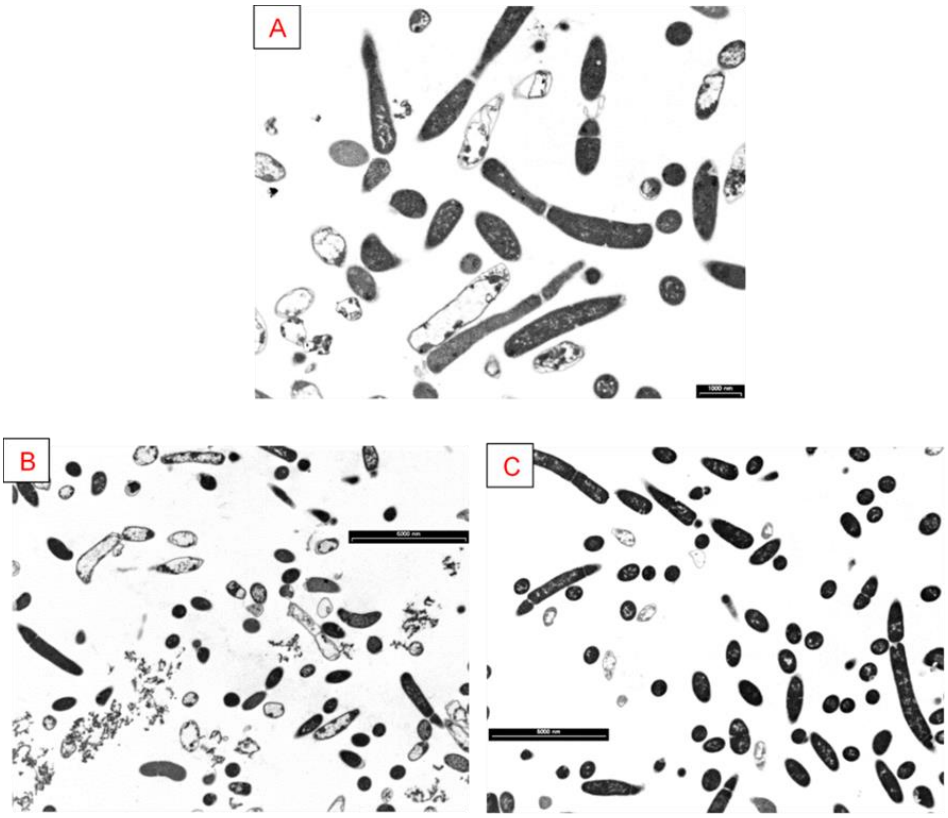


Fig. 4: TEM images of *G. stearothermophilus*: after overnight growths (**A**); in the presence of As(V) (**B**) and Cd(II) (**C**) at their respective MIC values. Scale: 1000nm in **A**, elsewhere 6000nm.

Conclusions

With the aim to characterize new thermophilic heavy metal resistant microorganisms, soil sampling was performed in a hydrothermal volcanic area near Naples in Italy, known as Pisciarelli. This is an acidic sulfate area located close to the Solfatara crater famous for an intense endogenous diffuse and fumarolic water-dominant outgassing activity; the chemical composition of mud and water samples has revealed that the main metal is iron, but, as in other geothermal settings, arsenate is an additional significant component. Since geothermal sites are very interesting sources of thermophilic organisms and Pisciarelli is an arsenic-rich area, we

hypothesized that novel thermophiles could be found able to detoxify this metal or use it for energy-yielding reactions. From a mud sample it was isolated a microorganism growing preferentially at 60°C. Subsequent molecular identification revealed homology to the species *Geobacillus stearothermophilus*. Our laboratory culturing experiments have demonstrated the ability of *G. stearothermophilus* to grow in the presence of arsenate in a range of concentrations comparable to those of bacteria classified as arsenic resistant, and in agreement with natural environmental setting as well. This study highlights the adaptation capabilities of the new isolate of *G. stearothermophilus* and its tolerance to extreme environmental conditions and points out to further molecular and physiological investigations to clarify its role in the biogeochemical cycle of arsenic as well as its potential for the management of heavy metal environmental contaminations.

Materials and Methods

Chemicals

The antibiotics and metal salts used in this work have been purchased by Sigma-Aldrich. Antibiotics are: Ampicillin (CAS Number: 7177-48-2); Bleomycin sulfate (CAS Number: 9041-93-4); Chloramphenicol (CAS Number: 56-75-7); Kanamycin sulfate (CAS Number: 70560-51-9); Hygromycin B (CAS Number: 31282-04-9); Tetracycline (CAS Number: 60-54-8). Metal salts used are the following: Sodium (meta)arsenite ($NaAsO_2$); Sodium arsenate dibasic heptahydrate ($Na_2HAsO_4 \cdot 7H_2O$); Cadmium chloride ($CdCl_2$); Cobalt chloride hexahydrate ($CoCl_2 \cdot 6H_2O$); Cobalt chloride ($CoCl_3$); Chromium oxide (CrO_3); Copper nitrate trihydrate [$Cu(NO_3)_2 \cdot 3H_2O$]; Mercury chloride, ($HgCl_2$); Nickel chloride ($NiCl_2$); Sodium orthovanadate dihydrate ($Na_3VO_4 \cdot 2H_2O$).

Strain isolation and molecular identification

Soil samples were collected in April 2016 at the hydrothermal site of Pisciarelli (Fig. 5) that lies on the Solfatara volcano, one of the various erupting inside the larger volcanic field of the Campi Flegrei since ca. 4100 years[60].

The materials were aseptically sampled from the main mud pool; both pH and temperature were measured contextually to sampling by indicator papers and a Fluke digital thermocouple probe, respectively. The water was collected from the bubbling mud pool,

while the mud was taken from its marginal water-poorer portion (Fig. 5). Temperature and pH values were 94.3°C and 4.8 in the bubbling mud pool respectively, while in the marginal water-poorer portion were 55.3°C and 6.0 respectively. The local levels of arsenic were in the 10-20 ppm range[40].

Fig. 5: Pisciarelli sampling site (left) and collected mud (right).

Enrichment cultures were set in 50 ml tubes with 20 ml of LB medium and inoculated with 1 g of soil sample at 37°C, 50°C and 75°C for 24 h. Then serial dilutions of those culture samples were stocked on LB agar plates for incubation at 37°C, 50°C and 75°C for 24 h. Bacterial colonies were found in the mud sample grown at 50°C. The isolation of a pure strain was carried out by repeated streak plating on solid medium prepared with the LB/agar incubated at 50°C.
To identify the new isolate, different approaches were used using standard procedures such as automated ribotyping on digested chromosomal DNA, fatty acids analysis and 16S based identification [61]. For 16S rDNA analysis, genomic DNA was extracted, the 16S rDNA gene amplified through PCR and the purified PCR products sequenced at Eurofins. The resulting sequencing data (1500 nt) were analysed using different databases and compared to sequences present in the EMBL database (https://www.ebi.ac.uk/ena).

Since the identification was not sure, a MALDI-TOF MS analysis was also commissioned to DSMZ: sample preparation and instrumental conditions are described elsewhere[62]. The identification report was generated by the Biotyper software and the strain was identified as *Geobacillus stearothermophilus* with a score corresponding to high probable strain identification (higher than 2.3).

Determination of optimal growth conditions

A frozen (-80°C) glycerol-stock of *G. stearothermophilus* was streaked on a LB plate (solidified by the addition of 1.5% Agar) and incubated overnight at 50°. Single colonies that appeared on the plate were inoculated into different liquid LB media at pH 3, pH 5, pH 7 and shaken at 50°C and 60°C. Growth was only observed at pH 7.0, whereas the optimal growth temperature turned out to be 60°C. The generation time (G) was calculated with the following formula: G=t/n where *t* is the time interval and *n* the number of generations (*t* is considered between 3h and 4h, in the exponential phase). All the experiments have been repeated in triplicate.

Antibiotic Susceptibility

For the determination of Minimal Inhibitory Concentrations (MIC) towards antibiotics, a frozen (-80°C) stock of *G. stearothermophilus* was streaked on a LB/agar plate and incubated at 60°C overnight. A single colony was inoculated into liquid LB medium and shaken at 60°C up to the exponential phase (1.5 OD_{600nm}). Then the bacterial culture was diluted up to 0.1 OD_{600nm} in LB medium supplemented with antibiotics (Ampicillin, Kanamycin, Chloramphenicol, Tetracycline, Hygromycin, Bleomycin) at final concentration ranging between 5 and 50 µg/mL and incubated for 16 h. The MIC values were determined as the lowest concentration of antibiotics that completely inhibited the growth of the strain as evaluated by OD_{600nm} measurements after incubation for 16 h. All the experiments have been repeated in triplicate.

Heavy metal resistance

For the determination of MIC towards heavy metal ions (As(V), As(III), Cd(II), Co(III), Cr(VI), Cu(II), Hg(II), Ni(II), V(V)), cell cultures were grown and diluted as described above. The heavy metals were added at increasing concentration ranging from 0.1 mM to 120 mM. The MIC values were determined as described above. The values reported are the average of three independent experiments.

Evaluation of As(V) biotransformation

The As-transformation capacity of *G. stearothermophilus* was evaluated using a colorimetric assay based on the formation of precipitates upon reaction of $AgNO_3$ with arsenic[63]. A single colony was cultured in LB liquid medium at 60 °C up to 1.0 OD_{600nm}; then an aliquot of the cell suspension was streaked on LB agar plates containing 50 mM sodium arsenate (Na_2HAsO_4). The LB agar plates were incubated at 60°C for 18 h and then flooded with 0.1M $AgNO_3$ to assess the transforming ability in the presence of arsenic. The colour of the precipitate on the plate was compared to a color scale, which could easily distinguish by eye different ratios of As(V)/As(III). In detail, the color scale was developed by mixing defined ratios of As(V) and As(III) being 50 mM the final concentration of total arsenic in each sample. All the experiments have been repeated in triplicate.

Bioinformatic analysis

Bioinformatic analyses were performed to evaluate the presence of arsenic and/or cadmium resistance genes in the genomes of the following three sequenced strains of *G. stearothermophilus*: a) strain n° 10 (Accession BioProject PRJNA252389); b) strain DSM 458 (Accession BioProject PRJNA327158); c) strain B5 (Accession Bioproject PRJNA513473). *Loci* containing sequences coding for putative arsenic resistance proteins were identified on the NCBI database and the corresponding translated sequences were aligned with the multiple sequence alignment program Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/).

Transmission electron microscopy (TEM)

G. stearothermophilus was grown in LB medium and in LB supplemented with As(V) 117 mM or Cd(II) 0.9 mM for 16 hours; a control grown in the absence of heavy metal ions was harvested at 1.5 OD_{600nm} corresponding to a mid-exponential growth phase. Cells were pelleted by centrifugation, washed twice with buffer phosphate saline (PBS 1%) and fixed as reported by Pinho et al.[64]. Resin sections were prepared with the ultramicrotome (LKB SuperNova) and the sections were serially stained with uranyl acetate and lead citrate. The sections were then studied on a Philips EM 208s Transmission Electron Microscope. Control cells without metal treatment were compared against heavy metal treated samples to check for possible heavy metal accumulation.

References:

1. Mathew, B.B.; Singh, H.; Biju, G.; Krishnamurthy, N.B. Classification , Source , and Effect of Environmental Pollutants and Their Biodegradation. **2017**, *36*, 55–71.
2. Rensing, C.; Rosen, B.P. Heavy Metals Cycle (Arsenic, Mercury, Selenium, others). In *Encyclopedia of Microbiology*; **2009** ISBN 9780123739445.
3. Kadirvelu, K.; Thamaraiselvi, K.; Namasivayam, C. Removal of heavy metals from industrial wastewaters by adsorption onto activated carbon prepared from an agricultural solid waste. *Bioresour. Technol.* **2001**, *76*, 63–65.
4. 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.
5. Organization, W.H. WHO human health risk assessment toolkit: chemical hazards. **2010**.
6. International Agency for Research on Cancer International Agency for Research on Cancer Iarc Monographs on the Evaluation of Carcinogenic Risks To Humans. *Iarc Monogr. Eval. Carcinog. Risks To Humans* **2002**, *96*, i-ix+1-390.
7. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans *Arsenic, metals, fibres, and dusts. A Review of human carcinogens*; **2012**.
8. Mamtani, R.; Stern, P.; Dawood, I.; Cheema, S. Metals and disease: A global primary health care perspective. *J. Toxicol.* 2011, **2011**.
9. Gillispie, E.C.; Sowers, T.D.; Duckworth, O.W.; Polizzotto, M.L. Soil Pollution Due to Irrigation with Arsenic-Contaminated Groundwater: Current State of Science. *Curr. Pollut. Reports* **2015**.
10. Waseem, A.; Arshad, J.; Iqbal, F.; Sajjad, A.; Mehmood, Z.; Murtaza, G. Pollution Status of Pakistan: A Retrospective Review on Heavy Metal Contamination of Water, Soil, and Vegetables. *Biomed Res. Int.* **2014**.
11. Gunduz, O.; Simsek, C.; Hasozbek, A. Arsenic pollution in the groundwater of Simav Plain, Turkey: Its impact on water quality and human health. *Water. Air. Soil Pollut.* **2010**.
12. Bartolucci, S.; Contursi, P.; Fiorentino, G.; Limauro, D.;

Pedone, E. Responding to toxic compounds: A genomic and functional overview of Archaea. *Front. Biosci.* **2013**.

13. Oremland, R.S.; Stolz, J.F. The ecology of arsenic. *Science (80-.).* **2003**.
14. Kulp, T.R.; Hoeft, S.E.; Asao, M.; Madigan, M.T.; Hollibaugh, J.T.; Fisher, J.C.; Stolz, J.F.; Culbertson, C.W.; Miller, L.G.; Oremland, R.S. Arsenic(III) fuels anoxygenic photosynthesis in hot spring biofilms from Mono Lake, California. *Science (80-.).* **2008**.
15. Pedone, E.; Bartolucci, S.; Fiorentino, G. Sensing and adapting to environmental stress: the archaeal tactic. *Front. Biosci.* **2004**.
16. Mukhopadhyay, R.; Rosen, B.P.; Phung, L.T.; Silver, S. Microbial arsenic: From geocycles to genes and enzymes. *FEMS Microbiol. Rev.* **2002**.
17. Rosen, B.P. Bacterial resistance to heavy metals and metalloids. *J. Biol. Inorg. Chem.* **1996**.
18. Saltikov, C.W.; Olson, B.H. 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.* **2002**.
19. Ji, G.; Silver, S. Regulation and expression of the arsenic resistance operon from Staphylococcus aureus plasmid pI258. *J. Bacteriol.* **1992**.
20. 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**, *10*, 1690–1701.
21. 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**.
22. Fernández, M.; Morel, B.; Ramos, J.L.; Krell, T. Paralogous regulators ArsR1 and ArsR2 of Pseudomonas putida KT2440 as a basis for arsenic biosensor development. *Appl. Environ. Microbiol.* **2016**.
23. Patel, P.C.; Goulhen, F.; Boothman, C.; Gault, A.G.; Charnock, J.M.; Kalia, K.; Lloyd, J.R. Arsenate detoxification in a Pseudomonad hypertolerant to arsenic. *Arch. Microbiol.*

2007.

24. 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 Proteomics* **2013**, *1834*, 2071–2079.
25. 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*.
26. 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.
27. Wang, P.P.; Bao, P.; Sun, G.X. Identification and catalytic residues of the arsenite methyltransferase from a sulfate-reducing bacterium, Clostridium sp. BXM. *FEMS Microbiol. Lett.* **2015**.
28. 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*, 1–9.
29. Dhankher, O.P.; Shasti, N.A.; Rosen, B.P.; Fuhrmann, M.; Meagher, R.B. Increased cadmium tolerance and accumulation by plants expressing bacterial arsenate reductase. *New Phytol.* **2003**.
30. Xu, C.; Rosen, B.P. Metalloregulation of Soft Metal Resistance Pumps. In *Metals and Genetics*; **1999**.
31. Abbas, S.Z.; Rafatullah, M.; Hossain, K.; Ismail, N.; Tajarudin, H.A.; Abdul Khalil, H.P.S. A review on mechanism and future perspectives of cadmium-resistant bacteria. *Int. J. Environ. Sci. Technol.* **2018**.
32. Sabadell, J.E.; Axtmann, R.C. Heavy metal contamination from geothermal sources. *Environ. Health Perspect.* **1975**.
33. Nriagu, J.O. A global assessment of natural sources of atmospheric trace metals. *Nature* **1989**.
34. 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.
35. Özdemir, S.; Kilinc, E.; Poli, A.; Nicolaus, B.; Güven, K. Biosorption of Cd, Cu, Ni, Mn and Zn from aqueous solutions by thermophilic bacteria, Geobacillus toebii sub.sp. decanicus and Geobacillus thermoleovorans sub.sp. stromboliensis: Equilibrium, kinetic and thermodynamic studies. *Chem. Eng. J.* **2009**, *152*, 195–206.
36. 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.
37. 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.
38. 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**.
39. 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.
40. Piochi, M.; Mormone, A.; Strauss, H.; Balassone, G. The acid sulfate zone and the mineral alteration styles of the Roman Puteoli (Neapolitan area, Italy): Clues on fluid fracturing progression at the Campi Flegrei volcano. *Solid Earth* **2019**, *10*, 1809–1831.
41. Valentino, G.M.; Cortecci, G.; Franco, E.; Stanzione, D. Chemical and isotopic compositions of minerals and waters from the Campi Flegrei volcanic system, Naples, Italy. *J. Volcanol. Geotherm. Res.* **1999**.
42. Giggenbach, W.F. Geothermal solute equilibria. Derivation of Na-K-Mg-Ca geoindicators. *Geochim. Cosmochim. Acta* **1988**.
43. Hedenquist, J.W.; Lowenstern, J.B. The role of magmas in the formation of hydrothermal ore deposits. *Nature* **1994**.
44. Ellis, A.J.; Mahon, W.A.J. Chemistry and geothermal systems. **1977**.
45. Webster, J.G.; Nordstrom, D.K. Chapter 4: Geothermal Arsenic. The source, transport and fate of arsenic in

geothermal systems. In *Arsenic in Ground Water*; **2003** ISBN 978-0-306-47956-4.
46. Farnfield, H.R.; Marcilla, A.L.; Ward, N.I. Arsenic speciation and trace element analysis of the volcanic río Agrio and the geothermal waters of Copahue, Argentina. *Sci. Total Environ.* **2012**, *433*, 371–378.
47. Cardellini, C.; Chiodini, G.; Frondini, F.; Avino, R.; Bagnato, E.; Caliro, S.; Lelli, M.; Rosiello, A. Monitoring diffuse volcanic degassing during volcanic unrests: The case of Campi Flegrei (Italy). *Sci. Rep.* **2017**.
48. Valentino, G.M.; Stanzione, D. Source processes of the thermal waters from the Phlegraean Fields (Naples, Italy) by means of the study of selected minor and trace elements distribution. *Chem. Geol.* **2003**.
49. Nazina, T.N.; Tourova, T.P.; Poltaraus, A.B.; Novikova, E. V.; Grigoryan, A.A.; Ivanova, A.E.; Lysenko, A.M.; Petrunyaka, V. V.; Osipov, G.A.; Belyaev, S.S.; et al. Taxonomic study of aerobic thermophilic bacilli: Descriptions of Geobacillus subterraneus gen. nov., sp. nov. and Geobacillus uzenensis sp. nov. from petroleum reservoirs and transfer of Bacillus stearothermophilus, Bacillus thermocatenulatus, Bacillus th. *Int. J. Syst. Evol. Microbiol.* **2001**.
50. Kananavičiute, R.; Čitavičius, D. Genetic engineering of Geobacillus spp. *J. Microbiol. Methods* **2015**.
51. Burgess, S.A.; Flint, S.H.; Lindsay, D.; Cox, M.P.; Biggs, P.J. Insights into the Geobacillus stearothermophilus species based on phylogenomic principles. *BMC Microbiol.* **2017**.
52. Egan, K.; Kelleher, P.; Field, D.; Rea, M.C.; Paul Ross, R.; Cotter, P.D.; Hilla, C. Genome sequence of Geobacillus stearothermophilus DSM 458, an antimicrobial-producing thermophilic bacterium, isolated from a sugar beet factory. *Genome Announc.* **2017**.
53. Yang, H.C.; Fu, H.L.; Lin, Y.F.; Rosen, B.P. Pathways of Arsenic Uptake and Efflux. In *Current Topics in Membranes*; **2012**.
54. Yang, H.C.; Cheng, J.; Finan, T.M.; Rosen, B.P.; Bhattacharjee, H. Novel pathway for arsenic detoxification in the legume symbiont Sinorhizobium meliloti. *J. Bacteriol.* **2005**.
55. Wu, D.; Zhang, Z.; Gao, Q.; Ma, Y. Isolation and characterization of aerobic, culturable, arsenic-tolerant

bacteria from lead–zinc mine tailing in southern China. *World J. Microbiol. Biotechnol.* **2018**.

56. Krumova, K.; Nikolovska, M.; Groudeva, V. Isolation and identification of arsenic-transforming bacteria from arsenic contaminated sites in Bulgaria. *Biotechnol. Biotechnol. Equip.* **2008**.
57. Krumova, K.; Nikolovska, M.; Groudeva, V. Characterization of arsenic-transforming bacteria from arsenic contaminated sites in Bulgaria. *Biotechnol. Biotechnol. Equip.* **2008**.
58. Chang, J.S.; Kim, Y.H.; Kim, K.W. The ars genotype characterization of arsenic-resistant bacteria from arsenic-contaminated gold-silver mines in the Republic of Korea. *Appl. Microbiol. Biotechnol.* **2008**.
59. 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.
60. Piochi, M.; Mormone, A.; Balassone, G.; Strauss, H.; Troise, C.; De Natale, G. Native sulfur, sulfates and sulfides from the active Campi Flegrei volcano (southern Italy): Genetic environments and degassing dynamics revealed by mineralogy and isotope geochemistry. *J. Volcanol. Geotherm. Res.* **2015**.
61. Schumann, P.; Pukall, R. The discriminatory power of ribotyping as automatable technique for differentiation of bacteria. *Syst. Appl. Microbiol.* **2013**.
62. Schumann, P.; Maier, T. MALDI-TOF mass spectrometry applied to classification and identification of bacteria. In *Methods in Microbiology*; **2014**.
63. Simeonova, D.D.; Lièvremont, D.; Lagarde, F.; Muller, D.A.E.; Groudeva, V.I.; Lett, M.C. Microplate screening assay for the detection of arsenite-oxidizing and arsenate-reducing bacteria. *FEMS Microbiol. Lett.* **2004**.
64. Pinho, S.C.; Nunes, O.C.; Lobo-da-Cunha, A.; Almeida, M.F. Inactivation of Geobacillus stearothermophilus spores by alkaline hydrolysis applied to medical waste treatment. *J. Environ. Manage.* **2015**.

General conclusion

The increase in public health problems due to heavy metal pollution requires a thorough study for the development of easier-to-use systems for the of detection and subsequent remediation of areas polluted by heavy metals.
This PhD project was aimed to the development of biosensors for the monitoring of environmental pollutants by exploiting a thermophilic system, due to its higher resistance of the latter in comparison to corresponding mesophilic counterparts.
The study of the resistance mechanisms to heavy metals in thermophilic microorganisms is advantageous, compared to the mesophilic counterpart, because these microorganisms very have a higher resistance to these contaminants; due to the adaptation to the high concentrations of heavy metals present in the hydrothermal vent, and from an application point of view; this type of biosensor could be used downstream of industrial processes that require high temperatures or presence of chemical agents to which these microorganisms are resistant.
The work performed during my PhD studies helped in increasing knowledge of *T. thermophilus* HB27 arsenic/cadmium resistance system at molecular level. Using biochemical techniques, the structural-functional characterization of the proteins was assessed *in vitro;* on the other hand, *in vivo* function was proved using a genetic editing system based on the CRISPR-Cas9 system adapted to work in *T. thermophilus*.
In this work new components of the arsenic resistance system in *T. thermophilus* HB27 have been characterized; in particular, they are: a)the transcription factor *Tt*SmtB (chapter III) that can be considered the playmaker of the resistance system; b) *Tt*ArsX, a P_{1B}-type membrane ATPase responsible for the efflux of arsenic, which we demonstrated to be also responsible for the efflux of the cadmium ions (chapter IV); c) *Tt*ArsM an arsenite methyltransferase which detoxify arsenite by methylation (chapter V). Moreover, the thesis reports on the creation of two different *whole-cell* biosensors.
In the chapter IV we obtained the first *whole-cell* biosensor able to detect around 10 µM of As (III), As (V) and Cd (II). This whole-cell biosensor is based on the plasmid pMHp*nqo*, which contains a

cassette composed by the promoter recognized by *Tt*SmtB, downstream of the reporter gene coding for a thermostable β-galactosidase, and an expression cassette for the metal responsive factor *Tt*SmtB under the control of the constitutive promoter p*nqo*.
The second type of *whole-cell* biosensor was described in the chapter V and is based on the integration of the reporter gene (*sYFP*) into the *T. thermophilus* HB27 genome. This biosensor is sensitive to much lower concentrations of arsenic and cadmium [1 µM As(III) and As(V), 0.5 µM Cd(II)] with an increase in sensitivity of approximately 10 times and is expected to be more stable and not dependent on growth with antibiotics. To obtain this reporter system, we developed a genetic editing tool based on CRISPR-Cas9.
Another goal of this research was the identification of new heavy metal resistant microorganisms for future application in bioremediation; in the last chapter (chapter VI) we have identified and characterized a new bacterial strain of *Geobacillus stearothermophilus* endowed with high arsenate and vanadate resistance characteristics.
In conclusion, this PhD project allowed to increase the knowledge on the comprehension of microbial resistance systems to heavy metals; to realize two different types of *whole-cells* biosensors for arsenic and to set up a new genomic editing system for *T. thermophilus* HB27.
A fine tuning of the biosensor's properties, through further genetic engineering will be attempted to make these systems even more suitable for applications. For example; it can be possible to use the CRISPR-Cas9 system that we developed in this thesis, to improve the performance of the biosensor by acting, for example, on the promoter sequences of the sensor protein *Tt*SmtB or by replacing all the genes involved in resistance with the sYFP, increasing drastically the amount of signal produced and making it possible to identify lower concentrations of metal.

Acknowledgements

I thank Prof. John van der Oost, from the Wageningen University, for supervising my research activity during my stay in the Laboratory of Microbiology, as visiting PhD student.

Acknowledgements go also to grants from: i) Società Italiana di Biochimica e Biologia Molecolare; ii) Università degli Studi di Napoli Federico II that financed the original research project: "Immobilization of ENzymes on hydrophobin-functionalized Nanomaterials"; iii) ERA-NET Cofund MarTERA who funded the project. "FLAshMoB: Functional Amyloid Chimera for Marine Biosensing"; iv) GoodbyWaste: Obtain GOOD products – exploit BY-products – reduce WASTE.

Appendix I: List of the Publication

- G. Gallo, I. Antonucci, L. Pirone, A. Amoresano, P. Contursi, D. Limauro, E. Pedone, S. Bartolucci and G. Fiorentino. **A physicochemical investigation on the metal binding properties of *Tt*SmtB, a thermophilic member of the ArsR/SmtB transcription factor family.** July 2019, International Journal of Biological Macromolecules.
- G. Gallo, R. Puopolo, D. Limauro, S. Bartolucci and G. Fiorentino. **Metal-Tolerant Thermophiles: From the Analysis of Resistance Mechanisms to their Biotechnological Exploitation.** 2018, The Open Biochemistry Journal
- Pane, K., Cafaro, V., Avitabile, A., Torres, M.D.T., Vollaro, A., De Gregorio, E., Catania, M.R., Di Maro, A., Bosso, A., Gallo, G. and Zanfardino, A., **Identification of Novel Cryptic Multifunctional Antimicrobial Peptides from the Human Stomach Enabled by a Computational–Experimental Platform.** 2018, ACS Synthetic Biology.
- Antonucci, I., Gallo, G., Limauro, D., Contursi, P., Ribeiro, A.L., Blesa, A., Berenguer, J., Bartolucci, S. and Fiorentino, G., 2018. **Characterization of a promiscuous cadmium and arsenic resistance mechanism in *Thermus thermophilus* HB27 and potential application of a novel bioreporter system**. Microbial cell factories, 17(1), p.78.
- Zanfardino, A., Bosso, A., Gallo, G., Pistorio, V., Di Napoli, M., Gaglione, R., Dell'Olmo, E., Varcamonti, M., Notomista, E., Arciello, A. and Pizzo, E., 2018. **Human apolipoprotein E as a reservoir of cryptic bioactive peptides: The case of ApoE 133-167.** Journal of Peptide Science, p.e3095.
- 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, 10 (6), 1690-1701.

- Member of the organising committee of the II Industrial Biotechnology Congress: **BioID&A** (Biotechnology Identity and Application) held in Naples on October 28th
- G. Gallo, I. Mougiakos, A. Scarica, M. Bianco, A. Carpentieri, P. Pucci, S. Bartolucci, J. van der Oost and G. Fiorentino. **Identification and Characterization of a new Arsenite Methyltransferase from *Thermus thermophilus* HB27: a new piece of the puzzle of the arsenic resistance (system).** Enzyme Discovery and Engineering for Biotechnological Applications, December 3-5, 2019 (Selected for short talk)
- G. Gallo, I. Mougiakos, A. Carpentieri, P. Pucci, S. Bartolucci, J. van der Oost and G. Fiorentino. **Arsenite methyltransferase from *Thermus thermophilus* HB27: a new piece of the puzzle of the arsenic resistance (system).** 60th SIB Congress Lecce, September 18 – 20, 2019
- G. Gallo, G. Fiorentino, P. Contursi, S. Bartolucci, D. Limauro. **Biochemical characterization of a peroxiredoxin of *Thermus thermophilus* HB27: a new player in the antioxidant defense**. 60th SIB Congress Lecce, September 18 – 20, 2019
- G. Gallo, I. Antonucci, L. Pirone, A. Amoresano, P. Contursi, D. Limauro, E. Pedone, S. Bartolucci and G. Fiorentino. ***Tt*SmtB; a thermophilic member of the ArsR/SmtB transcription factor: insights into stability and metal binding properties.** 2° Workshop BIO/10 Docenti e Ricercatori di Biochimica della Campania May 17th (Selected for short talk)
- G. Gallo, R. Puopolo, I. Antonucci, D. Limauro, S. Bartolucci and G. Fiorentino. **A new microbial isolate of *Geobacillus stearothermophilus* to address toxic metal pollution.** 12th edition of the International Congress on Extremophiles (Extremophiles2018) September 16-20, 2018 in Ischia (Naples, Italy).
- Antonucci, G. Gallo, D. Limauro, P. Contursi, A. L. Ribeiro, A. Blesa, J. Berenguer, S. Bartolucci and G. Fiorentino. **Cadmium and arsenic resistance in *Thermus thermophilus* HB27: molecular mechanisms and application in biosensing.** 12th edition of the International Congress on Extremophiles (Extremophiles2018) September 16-20, 2018 in Ischia (Naples, Italy).
- Immacolata Antonucci, G. Gallo, Josè Berenguer, Simonetta Bartolucci, Gabriella Fiorentino. ***Tt*SmtB, an arsenic sensing**

transcriptional repressor, mediates cadmium resistance in *Thermus thermophilus* HB27. Italian Society of Biochemistry and Molecular Biology, 59th Congress, Caserta, September 20 – 22, 2017.

- Anna Zanfardino, G. Gallo, Valeria Pistorio, Katia Pane, Michela Di Napoli, Andrea Bosso, Eugenio Notomista, Mario Varcamonti and Elio Pizzo. **Two new cryptic antimicrobial peptides identified in the human Apolipoprotein**. FISV - Federazione Italiana Scienze della Vita, 13th Congress. Sapienza University of Rome, Italy September 20 - 23 2016

Appendix III: Experiences in foreign laboratories

Visiting period at Bacterial Genetics group within the Laboratory of Microbiology, at the Department of Agrotechnology and Food Sciences of Wageningen University as a visiting PhD student for the period from November 1, 2018 to April 30, 2019.

Laboratory of Microbiology
Department Agrotechnology & Food Sciences
Wageningen University

Wageningen, May 7, 2019

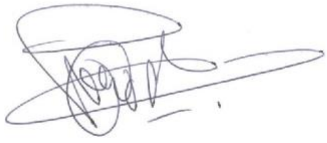
Confirmation letter

Dear Giovanni Gallo,

herewith I certify, that you stayed at Bacterial Genetics group within the Laboratory of Microbiology, at the Department of Agrotechnology and Food Sciences of Wageningen University as a visiting PhD student for the period from November 1, 2018 to April 30, 2019. Your research program was focused on the development of CRISPR-Cas based tools, to perform gene editing of the thermophilic bacterium *Thermus thermophilus* HB27. You successfully used the ThermoCas9 enzyme to generate specific mutants, and we look forward to publish these results in a joint publication in the near future.

It was a great pleasure having you here, to revive the collaboration with our friends in Napoli !

Sincerely,



John van der Oost

Prof. Dr. J. van der Oost,
Molecular Microbiology and Biochemistry
Laboratory of Microbiology, Wageningen University
Stippeneng 4, 6708 WE Wageningen, The Netherlands
Tel. +31-317-483108
e-mail john.vanderoost@wur.nl

ACS SyntheticBiology
Cite This: ACS Synth. Biol. XXXX, XXX, XXX−XXX
Research Article
pubs.acs.org/synthbio

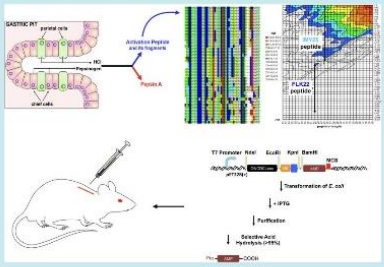
Identification of Novel Cryptic Multifunctional Antimicrobial Peptides from the Human Stomach Enabled by a Computational−Experimental Platform

Katia Pane,[@,▲] Valeria Cafaro,[†,▲] Angela Avitabile,[†] Marcelo Der Torossian Torres,[‡,§,∥,⊥,#,▽] Adriana Vollaro,[⊗] Eliana De Gregorio,[⊗] Maria Rosaria Catania,[⊗] Antimo Di Maro,[¶] Andrea Bosso,[†] Giovanni Gallo,[†] Anna Zanfardino,[†] Mario Varcamonti,[†] Elio Pizzo,[†] Alberto Di Donato,[†] Timothy K. Lu,[‡,§,∥,⊥,#] Cesar de la Fuente-Nunez,[*,‡,§,∥,⊥,#] and Eugenio Notomista[*,†]

[@]IRCCS SDN, Via E. Gianturco, 113, 80143 Naples, Italy
[†]Department of Biology, University of Naples Federico II, Naples 80126, Italy
[‡]Synthetic Biology Group, MIT Synthetic Biology Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[§]Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[∥]Department of Biological Engineering, and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States
[⊥]Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02139, United States
[#]The Center for Microbiome Informatics and Therapeutics, Cambridge, Massachusetts 02139, United States
[▽]Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André 09210580, Brazil
[⊗]Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II, Naples 80131, Italy
[¶]Department of Environmental, Biological and Pharmaceutical Sciences and Technologies, University of Naples, Vanvitelli, Caserta 81100, Italy

Supporting Information

ABSTRACT: Novel approaches are needed to combat antibiotic resistance. Here, we describe a computational−experimental framework for the discovery of novel cryptic antimicrobial peptides (AMPs). The computational platform, based on previously validated antimicrobial scoring functions, indicated the activation peptide of pepsin A, the main human stomach protease, and its N- and C-terminal halves as antimicrobial peptides. The three peptides from pepsinogen A3 isoform were prepared in a recombinant form using a fusion carrier specifically developed to express toxic peptides in *Escherichia coli*. Recombinant pepsinogen A3-derived peptides proved to be wide-spectrum antimicrobial agents with MIC values in the range 1.56−50 μM (1.56−12.5 μM for the whole activation peptide). Moreover, the activation peptide was bactericidal at pH 3.5 for relevant foodborne pathogens, suggesting that this new class of previously unexplored AMPs may contribute to microbial surveillance within the human stomach. The peptides showed no toxicity toward human cells and exhibited anti-infective activity *in vivo*, reducing by up to 4 orders of magnitude the bacterial load in a mouse skin infection model. These peptides thus represent a promising new class of antibiotics. We envision that computationally guided data mining approaches such as the one described here will lead to the discovery of antibiotics from previously unexplored sources.



Antibiotic resistance dramatically reduces the effectiveness of existing treatments for bacterial infections. Drug-resistant organisms can lead to recalcitrant infections; such infections are currently responsible for 23 000 deaths per year in the U.S. and have been predicted to kill 10 million people annually by the year 2050.[1] Therefore, new treatment options to combat antibiotic resistance are urgently needed.

Antimicrobial peptides (AMPs) constitute a promising alternative to conventional antibiotic therapies for the treatment of infections. These agents are key components of innate immunity in virtually all living organisms, acting as a

Received: February 22, 2018

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

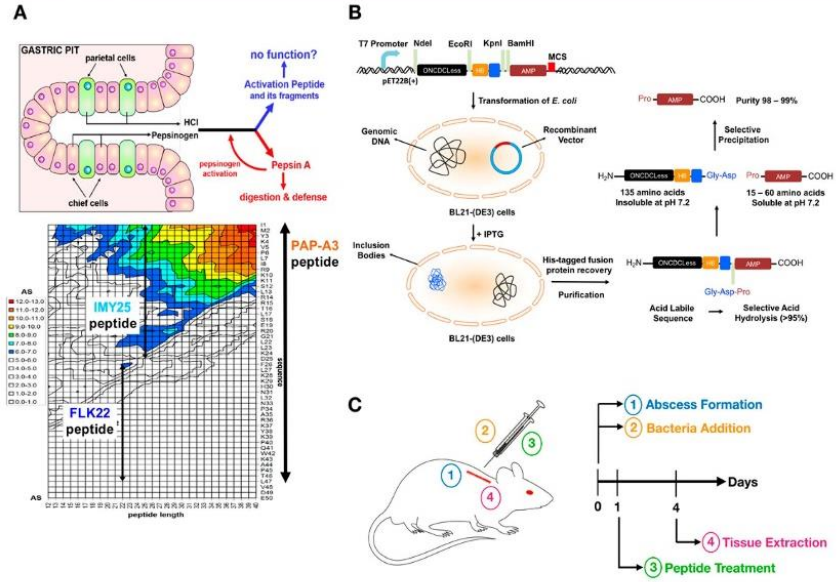


Figure 1. Overview, validation, and application of our computational−experimental platform. (A) Outline of the computational strategy for the discovery of new antimicrobial peptides. Top panel shows the structure of a "gastric pit" of gastric mucosa. The scheme shows parietal cells and chief cells, responsible for the secretion of HCl and pepsinogen, respectively. Mucus cells, secreting the protective mucus, are not shown. The activation routes of pepsinogen and the physiological roles of its products are shown on the right. The lower panel depicts our computational approach, which was leveraged to identify novel cryptic antimicrobial peptides. (B) Left panel shows recombinant platform for the production of peptides identified *in silico*. Right panel displays biochemical strategy used for peptide purification. (C) Schematic of mouse model used to assess the activity of computationally identified peptides.

first line of defense against microbial, fungal, parasitic, and viral invasions.[2] Generally, AMPs have from 12 to 50 residues, are positively charged, and possess a high content of hydrophobic residues (around 50%). In contrast to conventional antibiotics, AMPs mainly operate by binding to microbial membranes through electrostatic and hydrophobic interactions. This binding alters the membrane, eventually leading to cell death.[3]

Numerous AMPs also display other biological functions such as immunomodulation,[4] promotion of angiogenesis,[5] wound healing,[6] and anticancer activity.[7] In fact, AMPs have recently been renamed host defense peptides (HDPs) to account for their immune modulatory functions. Furthermore, some AMPs are able to eradicate bacterial biofilms, a serious threat to human health.[8−10] Examples include human and bovine AMPs,[11,12] a peptide derived from chicken cathelicidin CATH-2,[13] synthetic variants of natural AMPs[14] and D-enantiomeric peptides.[15]

In addition to well-known HDPs, like defensins and cathelicidins,[4] other HDPs present within the human body are the so-called "cryptic HDPs/AMPs".[16] Cryptic HDPs are derived from the proteolytic processing of large precursors not necessarily related to host defense. Both host and bacterial proteases can be involved in their release. Well-studied examples are peptides derived from the milk protein lactoferrin.[16,17] In the stomach of infants, lactoferrin, cleaved by pepsin A, releases broad-spectrum AMPs from its N-terminus that are more potent than lactoferrin itself.[17,18] After weaning, the action of lactoferrin-derived HDPs is no longer necessary because an effective (bio)chemical barrier against microbial invasions is set up by the combined action of stomach acidity—pH 2.5−4.0 during digestion[19]—and proteolytic activity of the aspartic proteases—pepsin A, gastricsin and several minor proteases.[20]

Here, through a computational−experimental approach, we show that the role in host defense of pepsin A may be even more complex. Pepsin A is secreted as an inactive zymogen, pepsinogen A, containing at the N-terminus a prosegment also known as activation peptide (47 residues long in the case of the human protein).[21,22] Activation peptide inhibits the potentially dangerous proteolytic activity of pepsin A by spanning its active site. After secretion, the acidic pH of gastric juice promotes the release of activation peptide from pepsinogen A (Figure 1). This process can be autocatalytic (*i.e.*, intramolecular) or mediated by an already functional pepsin molecule (intermolecular). Moreover, activation can occur in a single step, releasing the 47-residue peptide, or in two steps, with an intermediate cleavage in the middle of the activation peptide (between either Leu23 and Lys24 or Asp25 and Phe26 in the case of human pepsinogen A[22]). Low pH values (below 2.0) favor the intramolecular process, whereas high concentration of pepsin A and high pH values favor the intermolecular process and single-step activation.[22]

Interestingly, the average concentration of pepsin A in gastric juice is in the range of 0.5−1 mg/mL, but after

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

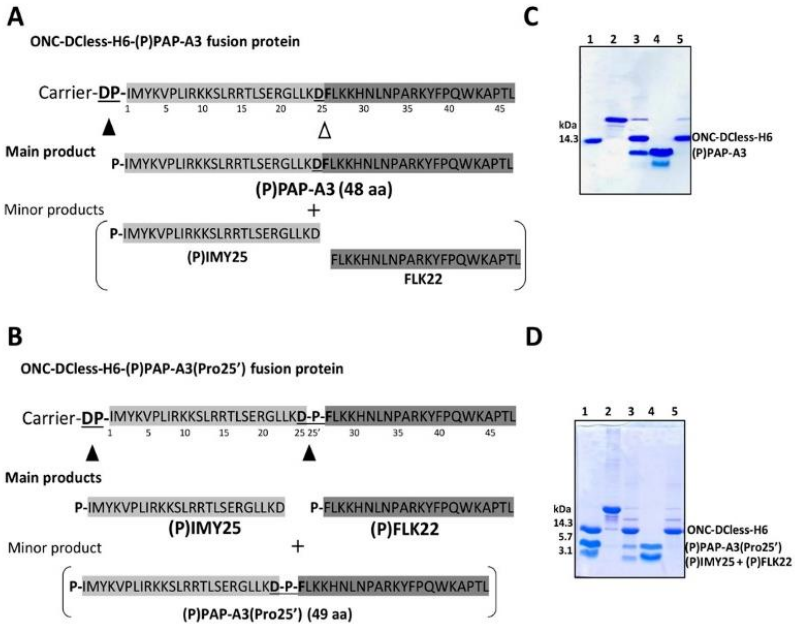


Figure 2. Major and minor products derived from the acid-catalyzed hydrolysis of ONC-DCless-H6-(P)PAP-A3 and ONC-DCless-H6-(P)PAP-A3(Pro25′). (A) and (B) Sequences of ONC-DCless-H6-(P)PAP-A3 and ONC-DCless-H6-(P)PAP-A3(Pro25′), respectively, and of their hydrolysis products. In both panels, the sequences of pepsinogen regions 1−25 and 26−47 are highlighted in light and dark gray, respectively; acid labile sites are shown with bold and underlined characters; filled and empty arrowheads indicate major and minor hydrolysis sites, respectively. (C) and (D) SDS-PAGE analysis of the hydrolysis products of ONC-DCless-H6-(P)PAP-A3 and ONC-DCless-H6-(P)PAP-A3(Pro25′), respectively, before and after the selective precipitation of the ONC carrier at pH 7. In both panels, lane 1, *Gallus gallus* lysozyme (14.3 kDa), purified (P)PAP-A3 (5.792 kDa) and (P)IMY25 (3.111 kDa) [smaller peptides are included only in panel (D)]; lane 2, purified fusion protein by IMAC; lane 3, acid cleavage pattern; lanes 4, soluble fraction after neutralization at pH 7.0; lane 5, insoluble fraction after neutralization at pH 7.0 (carrier protein and uncleaved fusion protein). Gels (20%) were stained with Coomassie Brilliant Blue R-250 in the presence of 4% formaldehyde.

stimulation by pentagastrin and insulin, this concentration can increase up to about 2 mg/mL.[23,24] As activation peptide is released in a stoichiometric amount by the activation process, these values correspond to an average estimate of the activation peptide and its fragments of 15−30 μM and a maximum concentration of about 60 μM. These surprisingly high physiological concentrations prompted investigation into whether, after their release, the activation peptide and its fragments might play other biological roles. In 1998, Minn and co-workers demonstrated that the activation peptides of pepsinogen A and progastricsin from the bullfrog *Rana catesbeiana* possess potent and broad-spectrum antimicrobial activity.[25] They also tested the fragments deriving from the activation peptides of human and monkey progastricsins (in the case of gastricsin, activation is prevalently or exclusively a two-step process), but they found only very modest activity.[25] After these findings were reported, the potential role of mammalian activation peptides was not investigated further.

We recently developed a novel computational tool for the detection of cryptic AMPs,[26] which has allowed the identification of several bioactive peptides derived from human, plant, and even archaeal proteins.[27−34] Using this tool, we have analyzed the zymogens of aspartic proteases and found that antimicrobial determinants are present in the activation peptide of pepsinogen A isoforms.

Here we report the recombinant production of the activation peptide of human pepsinogen isoform A3, named (P)PAP-A3, and its two complementary fragments 1−25 and 26−47, termed (P)IMY25 and (P)FLK22, respectively. We have characterized the bioactivity of these peptides and determined their antimicrobial activity at pH 7.0 and 3.5 against planktonic bacteria, including foodborne pathogens such as *Escherichia coli*, *Salmonella* spp. and *Enterococcus faecalis*. Using an interdisciplinary approach combining computational biology, biochemistry, and microbiology, we have identified and characterized the biological activity of previously unexplored AMPs, which may be developed as next-generation antimicrobials. Our findings provide novel insights into cryptic multifunctional AMPs that may be involved in microbial surveillance within the human gastrointestinal tract.

RESULTS

Identification of Cryptic Antimicrobial Peptides *via* Computational Analysis. Our computational tool for the identification of cryptic AMPs is based on the linear correlation of the antimicrobial potency of a given AMP, expressed as log (1000/MIC), to a score, called the "absolute score" (AS),

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

which depends on both the amino acid composition of the AMP and two strain-dependent variables describing the contributions of charge and hydrophobicity to antimicrobial activity.[26] Figure S1 shows the AS values for fragments 12 to 40 residues long at the N-terminus of human pepsinogen isoform A3, isoforms A4/A5 (whose activation peptides are identical) and progastricsin (pepsinogen C) as a function of fragment length and position. According to the parameter set used, AS values between 6.0 and 7.6 would correspond to MIC values in the range of 160−10 μM, whereas AS values higher than 7.6 would correspond to MIC values lower than 10 μM. For AS values close to or higher than 9.0, the linear correlation is no longer valid, so that further increase of the AS values is unlikely to correspond to a decrease in MIC values. As shown in Figure S1, only pepsinogen A activation peptides contain high-scoring fragments with AS values higher than 9.0 (*e.g.*, AS = 12.7 for fragment 1−40). The isometric plots also show that fragment 1−24 from pepsinogen A3, A4, and A5 corresponds to a local maximum with an AS of 8.9. Fragments 1−23 and 1−25, derived from (auto)proteolytic processing, have slightly lower scores (AS = 8.0 and 7.4, respectively), thus suggesting significant antimicrobial activity. Fragment 26−47 of pepsinogen A3 has an AS value slightly higher than the lowest threshold (AS = 6.1), suggesting the possibility of a measurable, even if low, antimicrobial activity. In order to verify that these features are limited to human pepsinogens, we analyzed a panel of 38 other pepsinogens A sequences (the multiple alignment is shown in Figure S2). With few exceptions, the activation peptides showed the presence of at least one fragment with an AS > 9.0 (Figure S3A). It is worth noting that human PAP-A3 had the highest AS value among the analyzed sequences. Only three sequences showed maximum AS values lower than the intermediate threshold (namely the mouse, cat, and dog pepsinogens A). When we limited the analysis to fragments in the region corresponding to residues 1−25 of the human pepsinogen, only about 50% of the sequences showed AS values close to or higher than 7.6.

These scores were calculated using the theoretical net charge at pH 7.0. In the pH range 2.5−4.0, a physiological range for gastric juice, the complete protonation of histidine residues and the partial protonation of glutamate residues increase the net positive charge, thus providing significantly higher AS values for almost all the analyzed sequences (Figure S3B). However, no threshold can be defined in this case, as the calibrations previously reported[26] were all derived from antimicrobial activity values measured at pH 7.0.

Recombinant Production of Pepsinogen A3-Derived Peptides. On the basis of the computational analysis reported in the previous section, we focused subsequent work on the activation peptide of human pepsinogen A3 (PAP-A3) and its main fragments, 1−25 (IMY25) and 26−47 (FLK22). Given the length of PAP-A3, we decided to produce a recombinant activation peptide in *E. coli* as a fusion protein with a new carrier derived from a modified ribonuclease (Figure S4), previously used to produce four recombinant AMPs.[27−29] The engineered carrier enables the production of AMPs in *E. coli* as insoluble fusion proteins in which the peptides, located at the C-terminus, are joined by a flexible linker region having an acid-labile Asp-Pro site, to allow the release of the peptide from the carrier. This well characterized cleavage method[35] leads to the production of peptides with an additional proline residue at the N-terminus. Fusion protein ONC-DCless-H6-(P)PAP-A3 (Figures 2A and S4A) was used to produce recombinant PAP-A3, herein named (P)PAP-A3 due to the additional N-terminal proline residue. The two shorter peptides, IMY25 and FLK22, were simultaneously produced from the same fusion protein ONC-DCless-H6-(P)PAP-A3(Pro25′) (Figures 2B and S4B). This recombinant protein is identical to ONC-DCless-H6-(P)PAP-A3 protein, except for the insertion of a proline residue (indicated as Pro25′) between residues Asp25-Phe26 in order to create a second acid-labile sequence (Figure 2B). The final product yielded the two desired recombinant peptides, named (P)IMY25 and (P)FLK22, each having an additional N-terminal proline (Figure 2B). It should be noted that several cationic antimicrobial peptides have been prepared in a recombinant active form with an additional proline at the N-terminus. Examples are human P-LL-37,[6] (P)GKY20,[27] a cryptic CAMP deriving from the N-terminus of human thrombin, (P)ApoB$_S$ and (P)ApoB$_L$,[29] two cryptic peptides deriving from an internal region of Apolipoprotein B, and a cryptic peptide deriving from an internal region of Apolipoprotein E and named "rApoE$_{PM}$ (133−150)",[28] as it contains two additional residues at the N-terminus (proline and methionine). The recombinant peptides P-LL-37, (P)GKY20 and rApoE$_{PM}$ (133−150) have demonstrated bioactivities (antimicrobial and/or immunomodulatory) indistinguishable from those of the corresponding synthetic peptides that lack the additional residue(s).

Fusion proteins were expressed at high levels (150−180 mg of fusion protein/L of culture) as inclusion bodies in *E. coli* BL21(DE3) (Figure S5A) and purified by IMAC (Figure 2C,D, respectively) as described elsewhere.[27] The release of peptides from recombinant proteins was obtained by chemical cleavage of the Asp-Pro sequences in acidic conditions (pH 2.0). As shown by SDS-PAGE analysis (Figure 2C,D), after 24 h of incubation at 60 °C, the Asp-Pro bond located in the linker region of both fusion proteins (Figure 2A,B, respectively) was efficiently cleaved (about 95% efficiency). Peptides were purified from the carrier by selective precipitation of the latter at pH 7.2 (Figure 2C,D), followed by repeated cycles of centrifugation. As shown by SDS-PAGE analysis (Figure 2C,D), peptides were recovered in the soluble fraction at neutral pH, whereas carriers were found in the insoluble fractions at pH 7.0. SDS-PAGE also showed that (P)PAP-A3 was contaminated by peptides with migration similar to those of (P)IMY25 and (P)FLK22, and the mixture (P)IMY25/(P)FLK22 was contaminated by a peptide with migration similar to that of (P)PAP-A3. RP-HPLC analyses of peptide mixtures derived from acid cleavage of both fusion proteins (Figure S5B,C, respectively) highlighted three main peaks for each mixture, whose identities were determined by mass spectrometry analyses (Supplementary Table S1). We estimated that about 5% of the ONC-DCless-H6-(P)PAP-A3 fusion protein was cleaved at the Asp25-Phe26 bond in (P)PAP-A3 (Figures 2A, 2C, and S5B), releasing the two shorter peptides, (P)IMY25 and FLK22, as minor products. This finding is in agreement with previous data[27] reporting low efficiency cleavage of Asp-X sequences when "X" is not a proline.

On the other hand, RP-HPLC (Figure S5C) and mass spectrometry (Supplementary Table S1) analyses of fusion protein ONC-DCless-H6-(P)PAP-A3(Pro25′) (Figure 2B) revealed that only 80% of the Asp25-Pro25′ cleavage site was cleaved (Figure S5C). The lower cleavage efficiency of this site compared to that of the Asp-Pro sequence in the flexible

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Table 1. Antibacterial Activity of Recombinant Pepsinogen A3 Derived Peptides

	MIC (μM)[a]					stomach infections
bacterial strains	(P)PAP-A3	(P)IMY25	(P)FLK22	mix[b]	(P)GKY20	references
Gram-negative						
Escherichia coli ATCC25922	6.25	25	>50	25	6.25	63, 64
Escherichia coli ATCC35218	6.25	25	>50	25	6.25	63, 64
Salmonella typhimurium ATCC14028	6.25	25	>50	12.5	6.25	65, 66
Salmonella enteriditis 706 RIVM[c]	6.25	25	>50	50	12.5	64
Pseudomonas aeruginosa RP73[e]	6.25	6.25	12.5	12.5	3.12	66
Pseudomonas aeruginosa PAO1	6.25	25	25	–[f]	25	66
Pseudomonas aeruginosa ATCC27853	12.5	–[f]	–[f]	–[f]	12.5	66
Acinetobacter baumanii ATCC17878	1.56	–[f]	–[f]	–[f]	6.25	67
Klebsiella pneumoniae ATCC700603	6.25	–[f]	–[f]	–[f]	12.5	63
Gram-positive						
Enterococcus faecalis ATCC29212	12.5	50	>50	50	12.5	63, 68
Listeria monocytogenes[d]	>50	>50	>50	>50	25	64
Staphylococcus aureus WKZ2 (MRSA)[c]	6.25	25	>50	25	12.5	69
Staphylococcus aureus ATCC6538P	1.56	1.56	3.12	–[f]	3.12	63
Bacillus spizizenii ATCC6633	12.5	–[f]	–[f]	–[f]	12.5	–[g]

[a]Assays were carried out using the broth dilution method in Nutrient Broth 0.5× except in the case of *Listeria monocytogenes* which was assayed in Heart Infusion Broth 0.5×. MIC values shown are the highest obtained from three independent experiments. [b]Mixture composed of (P)IMY25 and (P)FLK22 at equimolar concentrations was assayed. [c]*Salmonella enteriditis* 706 RIVM and methicillin-resistant *Staphylococcus aureus* WKZ2 (MRSA) were kindly provided by Prof. E. Veldhuizen. [d]*Listeria monocytogenes* clinical isolate. [e]*Pseudomonas aeruginosa* clinical isolate from cystic fibrosis patient, kindly provided by Prof. A. Bragonzi. [f]Not determined for these bacterial strains. [g]Not found in the literature to infect the human stomach.

linker region (about 95%) could be due to the presence of the bulky residues flanking the labile site (Figure 2B).

Peptide (P)PAP-A3 and the two shorter peptides, (P)IMY25 and (P)FLK22, were purified by RP-HPLC with a recovery of about 70%. The purity of the recombinant peptides was >95% as measured by RP-HPLC (Figure S6A–C). (P)PAP-A3 was produced to a final yield of about 18–20 mg/L of culture, whereas a yield of 7–8 mg/L was obtained both for (P)IMY25 and (P)FLK22.

Human Pepsinogen A3-Derived Peptides Exhibit Antimicrobial Properties. The antimicrobial activity of (P)PAP-A3, (P)IMY25, and (P)FLK22 was assayed against a panel of representative Gram-negative and Gram-positive bacteria, including gut pathogens and clinical multidrug-resistant strains. Recombinant GKY20, (P)GKY20,[27] a cryptic AMP derived from the C-terminus of human thrombin,[36] was used as positive control.

As shown in Table 1, (P)PAP-A3 exhibited broad-spectrum antimicrobial activity against all strains tested (MIC values $\leq$12.5 μM) with the exception of the *Listeria monocytogenes* clinical isolate. These MICs were comparable to those of the control peptide, (P)GKY20. As for the two smaller (P)PAP-A3 fragments, peptide (P)IMY25 inhibited the growth of most strains, but did so at MIC values slightly higher than those measured for (P)PAP-A3, in agreement with the prediction of our *in silico* method (Figure S1). On the other hand, (P)FLK22, as predicted, exhibited weaker antimicrobial activity and was active only against a few strains (Table 1).

Because IMY25 and FLK22 are released from pepsinogen A *in vivo* in stoichiometric amounts, we also investigated the antibacterial properties of an equimolar mixture of (P)IMY25 and (P)FLK22. This peptide mixture showed MIC values comparable to those of peptide (P)IMY25 alone (Table 1).

Direct Killing Activity of (P)PAP-A3. The standard MIC assay described in the previous section does not allow discrimination between bactericidal and bacteriostatic activity. Moreover, it cannot be performed at low pH values, which may inhibit bacterial growth. To address these points, we performed a plate viable count assays to test the antimicrobial effect of our lead peptide, (P)PAP-A3 (Figure S1), against common commensal and pathogenic bacteria (Figure 3).

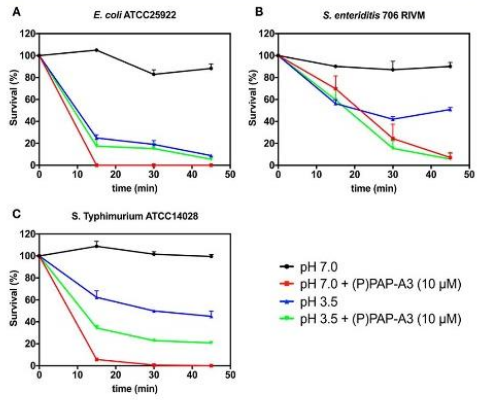


Figure 3. Time-kill assay of (P)PAP-A3. Bacterial survival was calculated as the ratio between CFU/mL in the incubated samples and CFU/mL in the untreated, not incubated cultures (t_0). Values are the mean (± SD) of two independent experiments.

Salmonella enteriditis, *Salmonella enterica* serovar *typhimurium* and *E. coli* were diluted in nutrient broth at pH 7.0 (essentially a meat broth). Part of the cultures were acidified by HCl addition to pH 3.5, a value close to the average pH of stomach contents during digestion,[24] and (P)PAP-A3 was added at a dose corresponding to its physiological concentration (10

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX–XXX

μM). At appropriate times, samples were diluted with nutrient broth containing 5 mM bicarbonate to increase the pH to 7 before plating, and the number of viable cells was determined. Bicarbonate was chosen as buffering agent because, *in vivo*, the duodenum secretes bicarbonate that neutralizes the hydrochloric acid released by the stomach.[37,38] As shown in Figure 3, at pH 7.0, (P)PAP-A3 significantly reduced the number of viable cells of the three selected strains within 45 min, a time comparable to or shorter than the time food usually remains in the stomach. At pH 3.5, *E. coli* quickly lost viability regardless of the presence of the peptide, whereas the two *Salmonella* species, more resistant to acidic pH, were effectively killed by the peptide at pH 3.5. Indeed, *S. typhimurium* survival was reduced by ~40% within minutes upon treatment with (P)PAP-A3 at pH 3.5 (Figure 3C). Therefore, it can be concluded that (P)PAP-A3 possesses bactericidal activity at both neutral and acidic pH values.

Effect of PAP-A3 and Its Fragments on Biofilms. As bacterial biofilms play a major role in infections (including gastrointestinal infections), protecting bacteria and increasing bacterial resistance both to the host immune system and to antibiotics,[39] we tested the efficacy of (P)PAP-A3 and its fragments on mature biofilms of two strains considered to be good models with which to study the resistance of biofilms to antimicrobials: *Pseudomonas aeruginosa* PAO1, a Gram-negative strain, and *Staphylococcus aureus* ATCC 6538P, a Gram-positive strain.

Total biofilm biomass and cell viability were measured in multiwell plates by crystal violet staining and the XTT assay, respectively. Confocal laser scanning microscopy (CLSM) coupled with Live/Dead bacterial cell staining was used to analyze the effect of the peptides on the structure and viability of biofilms obtained in static growth chambers.

(P)IMY25 significantly reduced *P. aeruginosa* PAO1 biofilm biomass (Figure S7A) as well as the viability of biofilm-embedded bacteria (Figure S7B). In contrast, (P)PAP-A3 and (P)FLK22 did not show antibiofilm activity (Figure S7A,B). CLSM analysis using Live/Dead staining confirmed that only (P)IMY25 disrupted the preformed *P. aeruginosa* biofilm (Figure S7C). In addition, (P)IMY25 reduced cell viability of pre-established *S. aureus* biofilms after 24 h of treatment in a concentration-dependent manner (Figure S8). The greatest levels of cell death mediated by (P)IMY25 were observed at a concentration higher than the MIC (Table 1). This finding is not unexpected, as the biofilm matrix acts as a filter that can reduce the accessibility of the embedded cells to relatively large molecules like peptides.[40,41] The filter effect could also help explain why (P)PAP-A3, the peptide most effective *vs* planktonic cells (Table 1), was not active against biofilms (Figures S7 and S8). We reasoned that this behavior could be the result of a strong interaction between the peptides and one or more components of the *P. aeruginosa* biofilm matrix, such as alginate, an acidic polysaccharide.[42] Several cationic AMPs bind to alginate, often changing their conformation upon binding. In particular, Deber and co-workers demonstrated that alginate can act as an AMP-scavenger, protecting matrix-embedded cells.[43−45] Therefore, we studied the ability of pepsinogen A-derived peptides to bind to alginate by CD spectroscopy (Figure S9 and Table S2) and by molecular modeling (Figures S10−S14) using a Monte Carlo-based strategy, which has proved useful for the modeling of several biological complexes of various natures and sizes.[46−51] This analysis, reported in detail in the Supplemental Results section of Supporting Information, showed that the three pepsinogen A-derived peptides interact with alginate, but the binding of (P)IMY25 is weaker than that of (P)PAP-A3 or (P)FLK22 and based on nonspecific electrostatic interactions. These results support the hypothesis that, of the three peptides tested, only (P)IMY25 may be able to diffuse inside the biofilm matrix and retain activity under these conditions.

Hemolytic Activity and Cytotoxicity. Assessing the cytotoxicity profile of AMPs is a key step to ensure their translatability. The toxicity of the pepsinogen A-derived AMPs toward eukaryotic cells in culture may not be particularly relevant, as under physiological conditions, gastric and duodenal mucosa are shielded by a mucous layer that defends epithelial cells from HCl and the gastric proteases.[37,52] It is very likely that this mucous layer also protects cells from membrane-active peptides. However, because PAP-A3 can also be considered a promising candidate for the development of novel topical antimicrobial agents, we evaluated its hemolytic activity and toxicity toward human keratinocytes (HaCat cells). As shown in Figure 4, (P)PAP-A3, at concentrations up to 40 μM and for incubation times up to 72 h, had negligible effects on murine erythrocytes or HaCat cells.

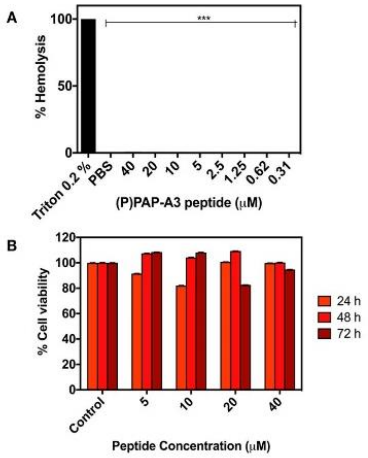


Figure 4. Cytotoxicity of recombinant peptide (P)PAP-A3 toward eukaryotic cells. (A) Peptide lytic effects on murine erythrocytes upon incubation for 1 h at 37 °C with (P)PAP-A3 or Triton X-100 0.2% used as positive control. Hemolysis was normalized to the A_{450} of Triton X-100-treated erythrocyte samples and reported as a percentage. (B) Viability (as percentage of total HaCaT cells) of HaCaT cell treated with (P)PAP-A3 determined by the MTT assay. Values are the mean (± SD) of three independent replicates.

***In Vivo* Activity of PAP-A3 and Its Fragments.** To further investigate the potential of (P)PAP-A3 and its derivatives as novel antibiotics, we tested their anti-infective potential in a murine abscess infection model. In line with our previous results (Table 1), (P)PAP-A3 displayed superior activity to (P)IMY25 and (P)FLK22 (Figure 5). Treatment with (P)PAP-A3 reduced *P. aeruginosa* CFU counts by around 4 orders of magnitude, whereas (P)IMY25 and (P)FLK22 each reduced bacterial loads by ~2 orders of magnitude (Figure 5).

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

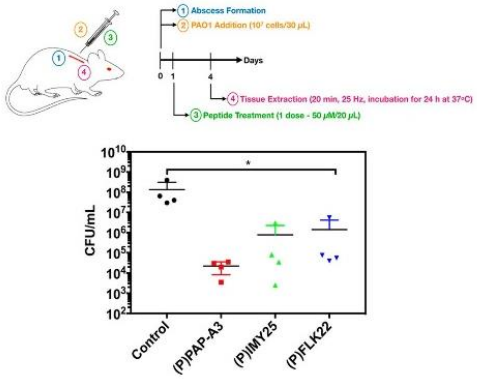


Figure 5. *In vivo* activity of cryptic peptides. The anti-infective potential of computationally identified cryptic peptides was evaluated in an abscess infection model following the outlined schematic. Two independent experiments were performed with 4 mice per group in each condition. Statistical significance was assessed using a one-way ANOVA.

Collectively, our data show that all peptides, but particularly (P)PAP-A3, exhibit antimicrobial activity *in vivo*. (P)PAP-A3 thus represents a novel peptide antibiotic that may be exploited for the treatment of bacterial infections.

■ DISCUSSION

We have demonstrated that human pepsinogen A activation peptide, at physiological concentrations, has previously undisclosed antimicrobial activity under conditions of both acidic and neutral pH. In the stomach, its bactericidal activity may complement the bacteriostatic/bactericidal activity of low pH values and the bactericidal activity of gastric aspartic proteases, thus helping to decrease the bacterial load in ingested food. When the food moves from the stomach to the duodenum, where bicarbonate secretion causes aspartic proteases to lose their catalytic activity and hence their bactericidal activity, the activation peptide can still exert its bactericidal role. From this point of view, the activation peptide could act as a substitute for lactoferrin-derived peptides in individuals whose diet does not include milk.

These findings could also help explain an intriguing peculiarity of pepsinogen physiology. It has long been known that a small but not negligible amount of pepsinogen is secreted in the bloodstream, filtered by the kidneys, and released unchanged in the urine.[53] This pepsinogen fraction, also known as "uropepsinogen", when purified from urine and exposed to pH 2.0, undergoes autocatalytic processing to fully active pepsin.[54,55] However, under physiological conditions, autocatalytic activation in the urine is highly unlikely. Moreover, even if a protease that is able to release the activation peptide exists in the urine, pepsin is inactive at pH higher than 5.0. On the other hand, the antimicrobial activity of pepsinogen A-derived AMPs can be exerted at neutral pH; therefore, uropepsinogen could have a physiological function as an AMP precursor. Obviously, more targeted studies are required to verify this hypothesis.

Our (re)evaluation of the antimicrobial activity of vertebrate and, in particular, of human pepsinogens activation peptides was prompted by an *in silico* analysis of about two thousand human secreted proteins. For this analysis, we used the method we have recently developed for the identification of cryptic antimicrobial peptides.[26] This method has already allowed the identification of cryptic antimicrobial regions in three human proteins, namely apolipoproteins E[28] and B[29] and 11-hydroxysteroid dehydrogenase-1 β-like,[30] in two plant type I ribosome-inactivating proteins from *Phytolacca dioica* L., respectively PD-L4[31] and PD-L1/2,[33] and even in the transcription factor Stf76 of the archaeal microorganism *Sulfolobus solfataricus*.[32] The fact that our computational method has correctly predicted both the high antimicrobial activity of the human pepsinogen A3 activation peptide and the presence of the antimicrobial determinants in the N-terminal half of the peptide provides a further strong confirmation of its efficacy and reliability.

Our *in silico* analysis also suggests that several other, even if not all, mammalian pepsinogens A may have antimicrobial properties. It is particularly intriguing that the determinants of antimicrobial and antibiofilm activities are located in the N-terminal half of the activation peptides, which is the most highly conserved among the mammalian sequences. We envision that our analysis will pave the way to the characterization of a new family of mammalian antimicrobial peptides likely involved in the microbial surveillance of the stomach and perhaps of the duodenum.

In spite of its length, PAP-A3 was produced in high yields and purity in *E. coli* by using the fusion systems previously developed by our group. The same fusion systems also allowed preparation of both PAP-A3 fragments from the processing of a single fusion protein. This demonstrates that the carrier ONC-DCless-H6 we have developed is suited not only for the preparation of single recombinant AMPs but also for the production of mixtures of bioactive peptides.

Finally, the three recombinant peptides, particularly (P)PAP-A3, displayed potent anti-infective activity *in vivo* (Figure 5). This peptide reduced by around 4 orders of magnitude the bacterial load in a mouse skin infection model. On the whole, our findings suggest the feasibility of developing topical antimicrobial agents based on PAP-A3 and its fragments and demonstrate that the computational–experimental platform we have developed could lead to the discovery and exploitation of bioactive peptides from previously unexplored sources.

■ EXPERIMENTAL PROCEDURES

Recombinant Vectors. ONC-DCless-H6-(P)PAP-A3, ONC-DCless-H6-(P)PAP-A3(Pro25′) and ONC-DCless-H6-(P)GKY20 fusion proteins were produced by using *pET22b-(+)/ONC-DCless-H6*-(P)PAP-A3, *pET22b(+)/ONC-DCless-H6*-(P)PAP-A3(Pro25′) and *pET22b(+)/ONC-DCless-H6*-(P)GKY20 plasmids, respectively. Details are reported into Supporting Information.

Protein Expression and Purification of Recombinant Peptides. Recombinant proteins were expressed in *E. coli* BL21(DE3) (Novagen, San Diego, CA, USA) and purified from inclusion bodies as previously described for ONC-DCless-H6-(P)GKY20 fusion protein.[27] Recombinant proteins, purified by IMAC (Ni Sepharose 6 Fast Flow was from GE Healthcare, Uppsala, Sweden), were cleaved in mild acid conditions (incubation for 24 h at 60 °C in 0.1 M acetic acid, 18 mM HCl, pH 2.0, under nitrogen atmosphere). Peptides were purified by carrier selective precipitation at neutral pH.[27]

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Cleavage efficiency of the Asp-Pro (DP) sequences into the linker region was determined by densitometric analysis of 20% SDS-PAGE[56] (ChemiDoc detection system, Bio-Rad; Quantity One software). Degree of fragmentation at internal cleavage sites was calculated by normalization of the RP-HPLC peak areas of peptides mixtures at neutral pH, with extinction coefficient calculated by the ProtParam tool. The identity of peptides was assessed by mass spectrometry analyses. Details are reported into Supporting Information.

MIC Assays. The MIC values of human pepsinogen A3-derived peptides against planktonic bacteria were examined by a broth microdilution method previously described for antimicrobial peptides[57] with minor modifications as described elsewhere.[27,28] The MIC values were determined as the lowest concentrations of different compounds at which no visible bacterial growth was observed upon overnight incubation at 37 °C. At least three independent experiments were carried out.

Time-Kill Assays. Bactericidal activity was tested by a plate viable-count method performed as previously described[20] with minor modifications. Briefly, overnight cultures were diluted in Nutrient Broth 0.5× (Becton-Dickenson, Franklin Lakes, NJ) to give final concentrations of ~5×10^6 CFU/mL (200 μL). Bactericidal activity of (P)PAP-A3 was tested at 10 μM final concentration. Assays were carried out at pH 7.0 and 3.5 (acidified by diluted HCl addition). Cell viability was also tested by control curves in the absence of peptide to evaluate any effects of pH on the growth. Samples were incubated at 37 °C for 45 min, and 10 μL of each mixture was then taken at 0, 15, 30, and 45 min and neutralized with 990 μL of Nutrient Broth 0.5× containing 5 mM sodium hydrogen carbonate. Bacterial suspensions (100 μL) were plated on LB agar plates, and viable counts were determined after incubation at 37 °C for 18 h. Experiments were carried out twice independently for each bacterial strain.

Biofilm Cultivation in Static Conditions and Confocal Microscopy. The antibiofilm activity of peptides against *S. aureus* ATCC6538P and *P. aeruginosa* PAO1 mature biofilm was determined by quantifying the biofilm biomass using the crystal violet staining method.[58] The metabolic activity of cells within mature biofilms of *P. aeruginosa* PAO-1 was assessed by using a XTT [2,3-bis(2-methyloxy-4-nitro-5-sulfophenyl)-2*H*-tetrazolium-5-carboxanilide] assay (Roche Diagnostics, Corporation, Indianapolis, IN, USA), according to the manufacturer's instructions. Experiments were carried out three times. Details are reported in the Supporting Information. For confocal microscopy experiments, bacteria were grown in a chambered cover glass (μ Slide 4 well; Ibidi GmbH, Munich, Germany), in static conditions for 48 h. Compounds were added to a 1-day-old biofilm at concentrations ranging from 25 to 6.25 μM. Untreated bacterial suspensions were used as a positive control. Following 24 h, biofilms were rinsed twice with PBS and stained with LIVE/DEAD BacLight Bacteria Viability stains (Life Technologies, Monza, Italy). The images were recorded using an LSM 700 inverted confocal laser-scanning microscope (Zeiss). The total wells were observed to see the biofilm global structure. Three different areas of each well were scanned using a 10× objective lens with the green channel for SYTO 9 (excitation 488 nm, emission 500−525 nm) and the red channel for PI (excitation 500−550 nm, emission 610−650 nm). All experiments were carried out in triplicate.

Molecular Docking. All calculations were performed using the ZMM-MVM molecular modeling package (ZMM Software Inc., http://www.zmmsoft.com), essentially as previously described.[46] All energy calculations included a hydration component as previously described.[59,60] Details are reported in the Supporting Information.

Hemolytic Assays. Erythrocytes were isolated from EDTA anticoagulated mouse blood by centrifugation followed by three PBS washes. Experiments were carried out three times as previously described.[61] Details are reported in Supporting Information.

Cytotoxicity toward Mammalian Cells. MTT assay[62] was performed on HaCat cells (1×10^4 cells/well) grown for 24, 48, and 72 h in the absence or presence of increasing concentrations (5, 10, 20, and 40 μM) of peptide (P)PAP-A3. Cell survival values were expressed as the percentage of viable cells with respect to control untreated samples. Experiments were performed three times independently.

Scarification Skin Infection Mouse Model. *P. aeruginosa* strain PAO1 was grown in tryptic soy broth medium overnight at 37 °C (250 rpm). A subculture was prepared until it reached an OD_{600} of 0.5. Subsequently cells were washed twice with sterile PBS (pH 7.4, 13 000 rpm for 1 min) and resuspended to a final concentration of 5.5×10^7 CFU/20 μL. To generate skin infection, female CD-1 mice (5−6 weeks old) were anesthetized with isoflurane and had their backs shaved. A superficial linear skin abrasion was made with a needle in order to damage the stratum corneum and upper-layer of the epidermis. Five minutes after wounding, an aliquot of 20 μL containing 5.5×10^7 CFU of bacteria in PBS was inoculated over each defined area containing the scratch with a pipet tip. One day after the infection, peptide solutions (50 μM/20 μL) were administered to the infected area. Animals were euthanized at day 4 and the area of scarified skin was excised 4 days postinfection, homogenized using a bead beater for 20 min (25 Hz), and serially diluted for CFU quantification. Two independent experiments were performed with 4 mice per group. Statistical significance was assessed using a one-way ANOVA.

Statistical Analysis. Graphs of the time-kill assay, CV and XTT assay, hemolysis and MTT assay, are presented as means and SDs from at least three independent experiments, expect for the time-kill assay (two independent experiments). We assessed differences in these assays using a standard *t* test and one-way ANOVA with Dunnett's multiple comparison tests. Data were analyzed using GraphPad Prism (Graph-Pad Software, Inc.). A *p* value less than 0.05 was considered to be statistically significant ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

Additional experimental procedures are provided in SI Material and Methods.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acssynbio.8b00084.

Supporting figures, tables, and experimental procedures (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: cfuente@mit.edu (Lead Contact).

*E-mail: notomist@unina.it.

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

ORCID

Katia Pane: 0000-0002-7594-7931
Timothy K. Lu: 0000-0002-3918-8923
Cesar de la Fuente-Nunez: 0000-0002-2005-5629
Eugenio Notomista: 0000-0003-0097-6487

Author Contributions

▲K.P. and V.C. contributed equally. K.P., V.C., E.N., and C.F.-N. conceived the project. K.P., V.C., A.A., M.D.T.T., A.V., E.D.G., A.D.M., A.B., G.G., A.Z., M.V., and C.F.-N. performed the experiments. K.P., E.D.G., M.R.C., V.C., and C.F.-N. performed data analysis. E.N. performed docking and modeling studies. K.P., V.C., E.P., E.N., A.D.D., T.K.L., and C. F.-N wrote the paper. All authors discussed the results and commented on the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants FFC#12/2014, FFC#20/2014 and FFC#16/2017 from the Italian Cystic Fibrosis Research Foundation (http://www.fibrosicisticaricerca.it), by Università di Napoli Federico II (CDA N. 19, 28/12/2016) and by the Ramon Areces Foundation (to C.F.-N.). Marcelo Der Torossian Torres was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (#2016/24413-0), and Timothy K. Lu acknowledges funding from DTRA (DTRA HDTRA1-15-1-0050).

■ REFERENCES

(1) European Comission (2017) *A European One Health Action Plan against Antimicrobial Resistance (AMR)*, pp 1−24.
(2) Zasloff, M. (2002) Antimicrobial Peptides of Multicellular Organisms. *Nature 415* (0028−0836), 389−395.
(3) Hancock, R. E. W., and Rozek, A. (2002) Role of Membranes in the Activities of Antimicrobial Cationic Peptides. *FEMS Microbiol. Lett. 206* (2), 143−149.
(4) Hilchie, A. L., Wuerth, K., and Hancock, R. E. W. (2013) Immune Modulation by Multifaceted Cationic Host Defense (Antimicrobial) Peptides. *Nat. Chem. Biol. 9* (12), 761−768.
(5) Koczulla, R., Degenfeld, G. V., Kupatt, C., Krötz, F., Zahler, S., Gloe, T., Issbrücker, K., Unterberger, P., Zaiou, M., and Lebherz, C. (2003) An Angiogenic Role for the Human Peptide Antibiotic LL-37/HCAP-18. *J. Clin. Invest. 111* (11), 1665−1672.
(6) Ramos, R., Pedro, J., Cristina, A., Costa, R., Guardão, L., Schmitt, F., Soares, R., Vilanova, M., Domingues, L., and Gama, M. (2011) Peptides Wound Healing Activity of the Human Antimicrobial Peptide LL37. *Peptides 32* (7), 1469−1476.
(7) Felício, M. R., Silva, O. N., Gonçalves, S., Santos, N. C., and Franco, O. L. (2017) Peptides with Dual Antimicrobial and Anticancer Activities. *Front. Chem.*, DOI: 10.3389/fchem.2017.00005.
(8) Costerton, W., Veeh, R., Shirtliff, M., Pasmore, M., Post, C., and Ehrlich, G. (2003) The Application of Biofilm Science to the Study and Control of Chronic Bacterial Infections. *J. Clin. Invest. 112* (10), 1466−1477.
(9) de la Fuente-Núñez, C., Reffuveille, F., Fernández, L., and Hancock, R. E. W. (2013) Bacterial Biofilm Development as a Multicellular Adaptation: Antibiotic Resistance and New Therapeutic Strategies. *Curr. Opin. Microbiol. 16* (5), 580−589.
(10) Miquel, S., Lagrafeuille, R., Souweine, B., and Forestier, C. (2016) Anti-Biofilm Activity as a Health Issue. *Front. Microbiol. 7*, 592.
(11) Jacobsen, A. S., and Jenssen, H. (2012) Human Cathelicidin LL-37 Prevents Bacterial Biofilm Formation. *Future Med. Chem. 4* (12), 1587−1599.
(12) Singh, P. K., Parsek, M. R., Greenberg, E. P., and Welsh, M. J. (2002) A Component of Innate Immunity Prevents Bacterial Biofilm Development. *Nature 417* (6888), 552−555.
(13) Molhoek, E. M., van Dijk, A., Veldhuizen, E. J. A., Haagsman, H. P., and Bikker, F. J. (2011) A Cathelicidin-2-Derived Peptide Effectively Impairs Staphylococcus Epidermidis Biofilms. *Int. J. Antimicrob. Agents 37* (5), 476−479.
(14) Pletzer, D., and Hancock, R. E. W. (2016) Antibiofilm Peptides: Potential as Broad-Spectrum Agents. *J. Bacteriol. 198* (19), 2572−2578.
(15) de la Fuente-Núñez, C., Cardoso, M. H., de Souza Cândido, E., Franco, O. L., and Hancock, R. E. W. (2016) Synthetic Antibiofilm Peptides. *Biochim. Biophys. Acta, Biomembr. 1858* (5), 1061−1069.
(16) Pizzo, E., Cafaro, V., Di Donato, A., and Notomista, E. (2018) Cryptic Antimicrobial Peptides: Identification Methods and Current Knowledge of Their Immunomodulatory Properties. *Curr. Pharm. Des. 24*, 1054.
(17) Kuwata, H., Yip, T. T., Tomita, M., and Hutchens, T. W. (1998) Direct Evidence of the Generation in Human Stomach of an Antimicrobial Peptide Domain (Lactoferricin) from Ingested Lactoferrin. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol. 1429* (1), 129−141.
(18) Nibbering, P. H., Ravensbergen, E., Welling, M. M., van Berkel, L. A., van Berkel, P. H., Pauwels, E. K., and Nuijens, J. H. (2001) Human Lactoferrin and Peptides Derived from Its N Terminus Are Highly Effective against Infections with Antibiotic-Resistant Bacteria. *Infect. Immun. 69* (3), 1469−1476.
(19) Kong, F., and Singh, R. P. (2008) Disintegration of Solid Foods in Human Stomach. *J. Food Sci. 73* (5), R67−R80.
(20) Zhu, H., Hart, C. A., Sales, D., and Roberts, N. B. (2006) Bacterial Killing in Gastric Juice−Effect of PH and Pepsin on Escherichia Coli and Helicobacter Pylori. *J. Med. Microbiol. 55* (9), 1265−1270.
(21) Kageyama, T. (2002) Pepsinogens, Progastricsins, and Prochymosins: Structure, Function, Evolution, and Development. *Cell. Mol. Life Sci. 59* (2), 288−306.
(22) Richter, C., Tanaka, T., and Yada, R. Y. (1998) Mechanism of Activation of the Gastric Aspartic Proteinases: Pepsinogen, Progastricsin and Prochymosin. *Biochem. J. 335*, 481−490.
(23) Roberts, N. B. (2006) Human Pepsins − Their Multiplicity, Function and Role in Reflux Disease. Aliment Pharmacol Ther Aliment. Pharmacol. Ther. Author J. *Aliment. Pharmacol. Ther. 24* (2), 2−9.
(24) Roberts, N. B., Sheers, R., and Taylor, W. H. (2007) Secretion of Total Pepsin and Pepsin 1 in Healthy Volunteers in Response to Pentagastrin and to Insulin-Induced Hypoglycaemia. *Scand. J. Gastroenterol. 42* (5), 555−561.
(25) Minn, I., Kim, H. S., and Kim, S. C. (1998) Antimicrobial Peptides Derived from Pepsinogens in the Stomach of the Bullfrog, Rana Catesbeiana. *Biochim. Biophys. Acta, Mol. Basis Dis. 1407* (1), 31−39.
(26) Pane, K., Durante, L., Crescenzi, O., Cafaro, V., Pizzo, E., Varcamonti, M., Zanfardino, A., Izzo, V., Di Donato, A., and Notomista, E. (2017) Antimicrobial Potency of Cationic Antimicrobial Peptides Can Be Predicted from Their Amino Acid Composition: Application to the Detection of "Cryptic" Antimicrobial Peptides. *J. Theor. Biol. 419*, 254−265.
(27) Pane, K., Durante, L., Pizzo, E., Varcamonti, M., Zanfardino, A., Sgambati, V., Di Maro, A., Carpentieri, A., Izzo, V., Di Donato, A., et al. (2016) Rational Design of a Carrier Protein for the Production of Recombinant Toxic Peptides in Escherichia Coli. *PLoS One 11* (1), e0146552.
(28) Pane, K., Sgambati, V., Zanfardino, A., Smaldone, G., Cafaro, V., Angrisano, T., Pedone, E., Di Gaetano, S., Capasso, D., Haney, E. F., et al. (2016) A New Cryptic Cationic Antimicrobial Peptide from Human Apolipoprotein E with Antibacterial Activity and Immunomodulatory Effects on Human Cells. *FEBS J. 283* (11), 2115−2131.
(29) Gaglione, R., Dell'Olmo, E., Bosso, A., Chino, M., Pane, K., Ascione, F., Itri, F., Caserta, S., Amoresano, A., Lombardi, A., et al.

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(2017) Novel Human Bioactive Peptides Identified in Apolipoprotein B: Evaluation of Their Therapeutic Potential. *Biochem. Pharmacol. 130*, 34−50.
(30) Bosso, A., Pirone, L., Gaglione, R., Pane, K., Del Gatto, A., Zaccaro, L., Di Gaetano, S., Diana, D., Fattorusso, R., and Pedone, E. (2016) A New Cryptic Host Defense Peptide Identified in Human 11-Hydroxysteroid Dehydrogenase-1 β-like: From in Silico Identification to Experimental Evidence. *Biochim. Biophys. Acta, Gen. Subj. 1861*, 2342.
(31) Pizzo, E., Zanfardino, A., Di Giuseppe, A. M. A., Bosso, A., Landi, N., Ragucci, S., Varcamonti, M., Notomista, E., and Di Maro, A. (2015) A New Active Antimicrobial Peptide from PD-L4, a Type 1 Ribosome Inactivating Protein of Phytolacca Dioica L.: A New Function of RIPs for Plant Defence? *FEBS Lett. 589* (19), 2812−2818.
(32) Notomista, E., Falanga, A., Fusco, S., Pirone, L., Zanfardino, A., Galdiero, S., Varcamonti, M., Pedone, E., and Contursi, P. (2015) The Identification of a Novel Sulfolobus Islandicus CAMP-like Peptide Points to Archaeal Microorganisms as Cell Factories for the Production of Antimicrobial Molecules. *Microb. Cell Fact. 14* (1), 126.
(33) Pizzo, E., Pane, K., Bosso, A., Landi, N., Ragucci, S., Russo, R., Gaglione, R., Torres, M. D. T., de la Fuente-Nunez, C., Arciello, A., et al. (2018) Novel Bioactive Peptides from PD-L1/2, a Type 1 Ribosome Inactivating Protein from Phytolacca Dioica L. Evaluation of Their Antimicrobial Properties and Anti-Biofilm Activities. *Biochim. Biophys. Acta, Biomembr. 1860* (7), 1425−1435.
(34) Zanfardino, A., Bosso, A., Gallo, G., Pistorio, V., Di Napoli, M., Gaglione, R., Dell'Olmo, E., Varcamonti, M., Notomista, E., Arciello, A., and Pizzo, E. (2018) Human Apolipoprotein E as a Reservoir of Cryptic Bioactive Peptides: The Case of ApoE 133−167. *J. Pept. Sci.*, e3095.
(35) Landon, M. (1977) Cleavage at Aspartyl-Prolyl Bonds. *Methods Enzymol. 47*, 145−149.
(36) Kasetty, G., Papareddy, P., Kalle, M., RydengÅrd, V., Mörgelin, M., Albiger, B., Malmsten, M., and Schmidtchen, A. (2011) Structure-Activity Studies and Therapeutic Potential of Host Defense Peptides of Human Thrombin. *Antimicrob. Agents Chemother. 55* (6), 2880−2890.
(37) Konturek, S. J., Konturek, P. C., Pawlik, T., Sliwowski, Z., Ochmański, W., and Hahn, E. G. (2004) Duodenal Mucosal Protection by Bicarbonate Secretion and Its Mechanisms. *J. Physiol. Pharmacol. 55* (2), 5−17.
(38) Seidler, U. E. (2013) Gastrointestinal HCO3- Transport and Epithelial Protection in the Gut: New Techniques, Transport Pathways and Regulatory Pathways. *Curr. Opin. Pharmacol. 13* (6), 900−908.
(39) Costerton, J. W., Stewart, P. S., and Greenberg, E. P. (1999) Bacterial Biofilms: A Common Cause of Persistent Infections. *Science 284* (5418), 1318−1322.
(40) Mah, T. F., and O'Toole, G. A. (2001) Mechanisms of Biofilm Resistance to Antimicrobial Agents. *Trends Microbiol. 9* (1), 34−39.
(41) Van Acker, H., Van Dijck, P., and Coenye, T. (2014) Molecular Mechanisms of Antimicrobial Tolerance and Resistance in Bacterial and Fungal Biofilms. *Trends Microbiol. 22* (6), 326−333.
(42) Hentzer, M., Teitzel, G. M., Balzer, G. J., Heydorn, A., Molin, S., Givskov, M., and Parsek, M. R. (2001) Alginate Overproduction Affects Pseudomonas Aeruginosa Biofilm Structure and Function. *J. Bacteriol. 183* (18), 5395−5401.
(43) Chan, C., Burrows, L. L., and Deber, C. M. (2004) Helix Induction in Antimicrobial Peptides by Alginate in Biofilms. *J. Biol. Chem. 279* (37), 38749−38754.
(44) Chan, C., Burrows, L. L., and Deber, C. M. (2005) Alginate as an Auxiliary Bacterial Membrane: Binding of Membrane-Active Peptides by Polysaccharides*. *J. Pept. Res. 65* (3), 343−351.
(45) Kuo, H. H., Chan, C., Burrows, L. L., and Deber, C. M. (2007) Hydrophobic Interactions in Complexes of Antimicrobial Peptides with Bacterial Polysaccharides. *Chem. Biol. Drug Des. 69* (6), 405−412.
(46) Siepi, M., Morales-Narváez, E., Domingo, N., Monti, D. M., Notomista, E., and Merkoçi, A. (2017) Production of Biofunctionalized MoS_2 Flakes with Rationally Modified Lysozyme: A Biocompatible 2D Hybrid Material. *2D Mater. 4* (3), 035007.
(47) Donadio, G., Sarcinelli, C., Pizzo, E., Notomista, E., Pezzella, A., Di Cristo, C., De Lise, F., Di Donato, A., and Izzo, V. (2015) The Toluene O-Xylene Monooxygenase Enzymatic Activity for the Biosynthesis of Aromatic Antioxidants. *PLoS One 10* (4), e0124427.
(48) De Rosa, M., Zanfardino, A., Notomista, E., Wichelhaus, T. A., Saturnino, C., Varcamonti, M., and Soriente, A. (2013) Novel Promising Linezolid Analogues: Rational Design, Synthesis and Biological Evaluation. *Eur. J. Med. Chem. 69*, 779−785.
(49) Notomista, E., Cafaro, V., Bozza, G., and Di Donato, A. (2009) Molecular Determinants of the Regioselectivity of Toluene/o-Xylene Monooxygenase from Pseudomonas Sp. Strain OX1. *Appl. Environ. Microbiol. 75* (3), 823−836.
(50) Notomista, E., Scognamiglio, R., Troncone, L., Donadio, G., Pezzella, A., Di Donato, A., and Izzo, V. (2011) Tuning the Specificity of the Recombinant Multicomponent Toluene O-Xylene Monooxygenase from Pseudomonas Sp. Strain OX1 for the Biosynthesis of Tyrosol from 2-Phenylethanol. *Appl. Environ. Microbiol. 77* (15), 5428−5437.
(51) Zanfardino, A., Restaino, O. F., Notomista, E., Cimini, D., Schiraldi, C., De Rosa, M., De Felice, M., and Varcamonti, M. (2010) Isolation of an Escherichia Coli K4 KfoC Mutant Over-Producing Capsular Chondroitin. *Microb. Cell Fact. 9* (1), 34.
(52) Yandrapu, H., and Sarosiek, J. (2015) Protective Factors of the Gastric and Duodenal Mucosa: An Overview. *Curr. Gastroenterol. Rep. 17* (6), 24.
(53) Kate, R. W., Pals, G., Eriksson, A. W., Donker, A. J., and Meuwissen, S. G. (1989) The Renal Metabolism of Pepsinogen A and C in Man. *Clin. Nephrol. 31* (2), 103−106.
(54) Minamiura, N., Ito, K., Kobayashi, M., Kobayashi, O., and Yamamoto, T. (1984) Uropepsinogen in Human Urine: Its Protein Nature, Activation and Enzymatic Properties of Activated Enzyme. *J. Biochem. 96* (4), 1061−1069.
(55) Sun, L. Q., Yasunobu, K. T., Minamiura, N., and Yamamoto, T. (1985) Human Uropepsinogen Is a New Human Pepsinogen Proisozyme. *Biochem. Int. 10* (4), 553−557.
(56) Laemmli, U. K. (1970) Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature 227* (5259), 680−685.
(57) Wiegand, I., Hilpert, K., and Hancock, R. E. W. (2008) Agar and Broth Dilution Methods to Determine the Minimal Inhibitory Concentration (MIC) of Antimicrobial Substances. *Nat. Protoc. 3* (2), 163−175.
(58) Stepanović, S., Vuković, D., Hola, V., Di Bonaventura, G., Djukić, S., Cirković, I., and Ruzicka, F. (2007) Quantification of Biofilm in Microtiter Plates: Overview of Testing Conditions and Practical Recommendations for Assessment of Biofilm Production by Staphylococci. *APMIS 115* (8), 891−899.
(59) Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) A New Force Field for Molecular Mechanical Simulation of Nucleic Acids and Proteins. *J. Am. Chem. Soc. 106* (3), 765−784.
(60) Lazaridis, T., and Karplus, M. (1999) Effective Energy Function for Proteins in Solution. *Proteins: Struct., Funct., Genet. 35* (2), 133−152.
(61) van Dijk, A., Molhoek, E. M., Veldhuizen, E. J. A., Bokhoven, J. L. M. T., Wagendorp, E., Bikker, F., and Haagsman, H. P. (2009) Identification of Chicken Cathelicidin-2 Core Elements Involved in Antibacterial and Immunomodulatory Activities. *Mol. Immunol. 46* (13), 2465−2473.
(62) Boukamp, P., Popp, S., Altmeyer, S., Hülsen, A., Fasching, C., Cremer, T., and Fusenig, N. E. (1997) Sustained Nontumorigenic Phenotype Correlates with a Largely Stable Chromosome Content during Long-Term Culture of the Human Keratinocyte Line HaCaT. *Genes, Chromosomes Cancer 19* (4), 201−214.

DOI: 10.1021/acssynbio.8b00084
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(63) Moles, L., Gómez, M., Jiménez, E., Bustos, G., de Andrés, J., Melgar, A., Escuder, D., Fernández, L., del Campo, R., and Rodríguez, J. M. (2017) Bacterial Diversity of the Gastric Content of Preterm Infants during Their First Month of Life at the Hospital. *Front. Nutr. 4*, 12.
(64) Sriphannam, W., Lumyong, S., Niumsap, P., Ashida, H., Yamamoto, K., and Khanongnuch, C. (2012) A Selected Probiotic Strain of Lactobacillus Fermentum CM33 Isolated from Breast-Fed Infants as a Potential Source of β-Galactosidase for Prebiotic Oligosaccharide Synthesis. *J. Microbiol. 50* (1), 119−126.
(65) Nastasi, A., Mammina, C., and Villafrate, M. R. (1993) Epidemiology of Salmonella Typhimurium: Ribosomal DNA Analysis of Strains from Human and Animal Sources. *Epidemiol. Infect. 110* (03), 553−565.
(66) Jeong, E.-S., Lee, K.-S., Heo, S.-H., Seo, J.-H., and Choi, Y.-K. (2011) Triplex PCR for the Simultaneous Detection of Pseudomonas Aeruginosa, Helicobacter Hepaticus, and Salmonella Typhimurium. *Exp. Anim. 60* (1), 65−70.
(67) Asahara, T., Takahashi, A., Yuki, N., Kaji, R., Takahashi, T., and Nomoto, K. (2016) Protective Effect of a Synbiotic against Multidrug-Resistant Acinetobacter Baumannii in a Murine Infection Model. *Antimicrob. Agents Chemother. 60* (5), 3041−3050.
(68) Perez, M., Calles-Enríquez, M., Nes, I., Martin, M. C., Fernandez, M., Ladero, V., and Alvarez, M. A. (2015) Tyramine Biosynthesis Is Transcriptionally Induced at Low PH and Improves the Fitness of Enterococcus Faecalis in Acidic Environments. *Appl. Microbiol. Biotechnol. 99* (8), 3547−3558.
(69) Fierobe, L., Decre, D., Muller, C., Lucet, J.-C., Marmuse, J.-P., Mantz, J., and Desmonts, J.-M. (1999) Methicillin-Resistant Staphylococcus Aureus as a Causative Agent of Postoperative Intra-Abdominal Infection: Relation to Nasal Colonization. *Clin. Infect. Dis. 29* (5), 1231−1238.

Received: 14 January 2018 | Revised: 5 April 2018 | Accepted: 30 April 2018
DOI: 10.1002/psc.3095

SPECIAL ISSUE ARTICLE

WILEY Journal of PeptideScience

Human apolipoprotein E as a reservoir of cryptic bioactive peptides: The case of ApoE 133-167

Anna Zanfardino[1] | Andrea Bosso[1] | Giovanni Gallo[1] | Valeria Pistorio[2] | Michela Di Napoli[1] | Rosa Gaglione[3] | Eliana Dell'Olmo[3] | Mario Varcamonti[1] | Eugenio Notomista[1] | Angela Arciello[3,4] | Elio Pizzo[1]

[1] Department of Biology, University of Naples Federico II, 80126 Naples, Italy
[2] Department of Molecular Medicine and Medical Biotechnologies, University of Naples Federico II, Naples, Italy
[3] Department of Chemical Sciences, University of Naples Federico II, 80126 Naples, Italy
[4] INBB—Istituto Nazionale Biostrutture e Biosistemi, Rome, Italy

Correspondence
Elio Pizzo, Department of Biology, University of Naples Federico II, 80126 Naples, Italy.
Email: elipizzo@unina.it

Funding information
Italian Cystic Fibrosis Foundation, Grant/Award Number: 16/2017

Bioactive peptides derived from the receptor-binding region of human apolipoprotein E have previously been reported. All these peptides, encompassing fragments of this region or designed on the basis of short repeated cationic sequences identified in the same region, show toxic activities against a broad spectrum of bacteria and interesting immunomodulatory effects. However, the ability of these molecules to exert antibiofilm properties has not been described so far. In the present work, we report the characterization of a novel peptide, corresponding to residues 133 to 167 of human apolipoprotein E, here named ApoE (133-167). This peptide, besides presenting interesting properties comparable with those reported for other ApoE-derived peptides, such as a direct killing activity against a broad spectrum of bacteria or the ability to downregulate lipopolysaccharide-induced cytokine release, is also endowed with significant antibiofilm properties. Indeed, the peptide is able to strongly affect the formation of the extracellular matrix and also the viability of encapsulated bacteria. Noteworthy, ApoE (133-167) is not toxic toward human and murine cell lines and is able to assume ordered conformations in the presence of membrane mimicking agents. Taken together, collected evidences about biological and structural properties of ApoE (133-167) open new perspectives in the design of therapeutic agents based on human-derived bioactive peptides.

KEYWORDS
antimicrobial peptides, apolipoprotein E, bacterial biofilm, inflammation, lipopolysaccharide

1 | INTRODUCTION

Human apolipoprotein E (ApoE) is a circulating glycoprotein with a molecular mass of ~34 kDa; it is mainly involved in the clearance of cholesterol and other lipids from the blood circulation via binding to cell surface ApoE receptors.[1] In humans, there are 3 common isoforms, with different affinities for LDL receptor,[2,3] produced by hepatocytes, macrophages, and adipocytes in the peripheral tissues and in astrocytes, microglia, and vascular mural cells in central nervous system.[4] Plasma ApoE is preferentially associated to very low density lipoprotein particles, whereas it is found in high-density lipoprotein-like particles in the CNS.[5,6] Brain ApoE is primarily derived from de novo synthesis because the blood-brain barrier limits ApoE transport into and out of the brain.[7] Apolipoprotein E consists of 2 main structural domains connected by a hinge region.[8] The N-terminal domain (residues 1-167) consists of a 4 alpha helix bundle containing the receptor binding region (residues 130-162), whereas the C-terminal domain (residues 206-299) consists of 3 alpha helices and represents the major lipid/lipoprotein binding region. Upon binding to lipids, ApoE undergoes a major conformational change and forms a molecular envelope around the surface of the phospholipid outer shell of a nascent lipoprotein particle.[9] Apolipoprotein E receptor-binding region retains the biological activities of the whole protein, as for instance COG-133 that shows pharmacological properties, such as

antiinflammatory and neuroprotective activities.[10,11] Other reports, focused on monomeric and dimeric synthetic peptides encompassing ApoE amino acids 130 to 169, revealed the importance of α-helical conformations and of positively charged amino acids.[12] Several recent reports are focused on the study of ApoE-derived peptide antimicrobial and immunomodulatory properties.[13-15] They have been identified in the receptor binding region, or they have been obtained by combining clusters of residues identified in the same region. For most of these peptides, the antimicrobial and immunomodulatory properties have been analyzed. Interestingly, the most active antimicrobial peptides were found to partially or almost completely cover the receptor binding region (residues 130-162), thus revealing that amino acidic composition could be determinant for their effective interaction with bacterial membrane.[16,17] Recently, our research group has developed a bioinformatics method that identifies antimicrobial peptides within the sequences of larger protein precursors and quantitatively predicts their antibacterial activity.[18] The method assigns an antimicrobial score to peptides based on their net charge, hydrophobicity and length, and 2 bacterial strain-dependent weight factors.[18] By means of this method, we have already reported the characterization of novel human peptides[19,20] including an exhaustive structural and functional study of a cryptic peptide identified in human ApoE (residues 133-150).[15] We deeply characterized the ability of ApoE (133-150) to assume stable conformations in the presence of membrane mimicking agents and LPS as well as its antibacterial and immunomodulatory properties.[15] Very interestingly, the antimicrobial score plot derived by the in silico analysis of ApoE (Figure S1) showed that, in addition to the local maximum corresponding to region 133 to 150, a further maximum was present (indicated by the black arrow on the right in Figure S1). This maximum, corresponding to region 133 to 167, has a score of 12.3, ie, only slightly higher than the score of region 133 to 150 (11.8). Based on these considerations, we hypothesized that ApoE (133-167) could possess antimicrobial properties comparable with those reported for other ApoE-derived peptides and that, at the same time, it could exert additional activities related to its length (35 residues) and amino acidic composition, the latter virtually able to drive the peptide to acquire a larger helical conformation. Accordingly, the aim of this study was to analyze ApoE (133-167) antimicrobial activity on planktonic and sessile bacteria, its effects on eukaryotic cells, its conformation in the presence of membrane mimicking agents, and its antiinflammatory properties.

2 | MATERIALS AND METHODS

2.1 | Bacterial strains and growth conditions

Bacterial strains used in this study were *Escherichia coli* ATCC 25922, *E. coli* DH5α, *Bacillus subtilis* PY79, *Bacillus globigii* TNO BMO13, *Pseudomonas aeruginosa* PAO1 wild type, *P. aeruginosa* KK27 (cystic fibrosis clinical isolate kindly provided by Dr Alessandra Bragonzi), *P. aeruginosa* ATCC 27853, methicillin-resistant *Staphylococcus aureus* WKZ2 (*MRSA* WKZ2), *S. aureus* ATCC 6538P, *S. aureus* ATCC 29213, and *Salmonella enteritidis* 706 RIVM. All bacterial strains were grown in Mueller Hinton Broth (MHB, Becton Dickinson Difco, Franklin Lakes, NJ) and on tryptic soy agar (Oxoid Ltd., Hampshire, UK). In all the experiments, bacteria were inoculated and grown overnight in MHB at 37°C and then transferred to a fresh MHB tube and grown to midlogarithmic phase.

2.2 | Eukaryotic cell lines

THP-1 cells, a pro-monocytic cell line, were cultured in RPMI 1640 (Sigma Aldrich, Milan, Italy), supplemented with 10% fetal bovine serum (HyClone, GE Healthcare Lifescience, Chicago, IL), 10-mM Hepes, 0.1-mM MEM nonessential amino acids, 1-mM sodium pyruvate, and 1% penicillin/streptomycin (Sigma Aldrich, Milan, Italy), in a 5% CO_2 humidified atmosphere at 37°C. THP-1 cells were differentiated with phorbol 12-myristate 13-acetate (Sigma-Aldrich, Milan, Italy) at a concentration of 10 ng/mL in RPMI 1640 supplemented with 1% fetal bovine serum (HyClone, GE Healthcare Lifescience, Chicago, IL). HaCat, HeLa, HEK-293, CaCo-2, and RAW 264.7 cells were from ATCC, Manassas, VA. Cells were cultured in Dulbecco's modified Eagle's medium (Sigma Aldrich, Milan, Italy), supplemented with 10% fetal bovine serum (HyClone, GE Healthcare Lifescience, Chicago, IL) and antibiotics, in a 5% CO_2 humidified atmosphere at 37°C.

2.3 | Antimicrobial activity assays

The antimicrobial activity of ApoE-derived peptides against planktonic bacteria was evaluated by broth microdilution method, performed as previously described for antimicrobial peptides[21] with minor modifications. Briefly, a single colony of each strain was resuspended in 5 mL of TY medium (Difco, Detroit, MI) and incubated overnight at 37°C. When the culture reached OD_{600nm} of 1 unit, it was diluted 1:100 (vol/vol) in 2-mM phosphate buffered saline (PBS) pH 7.0. Then, to test the effects of peptides on bacterial viability, samples (500 μL final volume) were prepared by diluting bacterial cells (20 μL) in growth medium containing the molecules under test at the desired concentration. The samples were then incubated at 37°C under stirring at 150 rpm for 4 hours. Subsequently, serial dilutions of all the samples were plated on TY-agar and incubated overnight at 37°C. The percentage of viable cells was finally evaluated by colony counting.[22] All the experiments were carried out in triplicate. To determine the minimal inhibitory concentration values, assays were performed as previously described elsewhere.[19,20] Briefly, bacteria were grown to midlogarithmic phase at 37°C and then diluted to 1×10^6 CFU/mL in Difco 0.5× Nutrient Broth (Becton-Dickenson, Franklin Lakes, NJ) containing increasing amounts of ApoE (133-167) or ApoE (133-150; 0-40 μM). Starting from a peptide stock solution, 2-fold serial dilutions were sequentially carried out, accordingly to broth microdilution method.[23] Following overnight incubation, MIC_{100} values were determined as the lowest peptide concentration responsible for no visible bacterial growth.

2.4 | ATP leakage measurements

MRSA WKZ-2 and *E. coli* ATCC 25922 were grown to midlogarithmic phase in MHB at 37°C. Bacteria were then centrifuged and diluted to 2×10^7 CFU/mL in MHB medium diluted 1:100 (vol/vol) in PBS at pH 7.0. In the case of each diluted sample, 60 μL of bacterial

suspension was incubated with 60 μL of peptide solution (0.5 or 2 μM for ApoE-derived peptides, 0.5 μM for positive control PMAP-36 peptide[21]) for 20 minutes at 37°C. The samples were then centrifuged, and the supernatant was stored at 4°C until further use. The bacterial pellet was suspended in lysis buffer (Roche Diagnostics Nederland B. V., Almere, the Netherlands) and further incubated at 100°C following the manufacturer instructions. Cell lysates were then centrifuged, and supernatants were kept on ice. Subsequently, both intracellular and extracellular ATP levels were determined by using the Roche ATP bioluminescence kit HS II, according to the manufacturer's instructions (Roche Diagnostics Nederland B.V., Almere, the Netherlands).

2.5 | Cytotoxicity on mammalian cells

Toxicity toward undifferentiated and differentiated THP-1, HaCat, HeLa, HEK-293, CaCo-2, and RAW 264.7 cells was assessed by performing the 3-(4,5-dimethylthiazol-2-yl)-2,5 diphenyltetrazolium bromide (MTT) reduction inhibition assay.[24] Cytotoxicity experiments were performed at least 4 times independently. In all the cases, cells were grown for 24 hours in the absence or in the presence of increasing concentrations of the peptide under test. Cell survival values are expressed as percentage of viable cells with respect to control untreated samples.

2.6 | Hemolytic assay

The hemolytic assay was performed as previously described.[19,20] Briefly, ethylenediaminetetraacetic acid anticoagulated mouse blood was centrifuged for 10 minutes at 800 g at 20°C to obtain red blood cells, which were washed 3 times, and 200-fold diluted in PBS. Subsequently, 75-μL aliquots of red blood cells were added to 75-μL peptide solutions (final concentration ranging from 0 to 80 μM) in 96-well microtiter plates, and the mixture was incubated for 1 hour at 37°C. A solution of 0.2% (vol/vol) Triton X-100 served as a control for complete lysis. Supernatants, collected after centrifugation, were transferred into polystyrene 96-wells plates, and absorbance was measured at 405 nm by using an automatic plate reader (FLUOstar Omega, BMG LABTECH, Germany). Hemolysis (%) was calculated as follows: $[(Abs_{405\ nm}\ \text{peptide} - Abs_{405\ nm}\ \text{blank})/(Abs_{405\ nm}\ 0.2\%\ \text{Triton} - Abs_{405\ nm}\ \text{blank})] \times 100$.

2.7 | Circular dichroism analyses

Circular dichroism (CD) experiments were performed on a Jasco J-810 CD spectropolarimeter. The cell path length was 0.1 cm. Circular dichroism spectra were collected at 20°C in the 200 to 260-nm (far-UV) interval, with a 10-nm/min scan rate, 2-nm bandwidth, and a 4-second response. Spectra are reported in mean residue ellipticity, calculated by dividing the total molar ellipticity by the number of amino acids in the molecule. Lyophilized peptides were dissolved in 50-mM sodium PBS at pH 7.4 at a concentration of 50 μM. The spectra were signal-averaged over at least 3 scans, and the baseline was corrected by subtracting the background and reported without further signal processing. Circular dichroism spectra of the peptide were collected in the absence or in the presence of increasing concentrations of sodium dodecyl sulfate (SDS; 0-20mM, Sigma Aldrich, Milan, Italy), trifluoroethanol (TFE, 0-30%, Sigma Aldrich, Milan, Italy), and lipopolysaccharide (LPS) from *P. aeruginosa* strain P10 or from a cystic fibrosis isolate strain KK27 tested at a concentration ranging from 0 to 0.8 mg/mL. Circular dichroism spectra were corrected by subtracting every time the contribution of the compound under test at any given concentration. Circular dichroism spectrum deconvolution was performed according to the Selcon method by using Dichroweb,[25] to estimate secondary structure contents.

2.8 | Real-time quantitative polymerase chain reaction

THP-1 undifferentiated cells were plated in 6-well plates at a density of 1×10^6 cells/well for 12 hours. The cells were then incubated with ApoE (133-150) or ApoE (133-167) at a concentration of 1 μM (2.5 and 4.4 μg/mL, respectively). Peptides were tested alone or in combination with LPS from *E. coli* O111:B4 (0.1 μg/mL). After an incubation of 1 hour, the cells were collected and washed with cold PBS. Total RNA was extracted from cell pellet by using QIAGEN Rneasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. One microgram of total RNA was retrotranscribed by using Maxima Reverse Transcriptase (Thermo-Fisher Scientific, Waltham, MA) according to the manufacturer's instructions. Quantitative real-time polymerase chain reaction (qRT-PCR) amplifications were performed by using the Power SYBR Green PCR Master Mix (Thermo-Fisher Scientific, Waltham, MA) in Applied Biosystems 7500 real-time PCR Systems. The qRT-PCR conditions were 95°C for 15 minutes followed by 40 cycles of 95°C for 15 seconds, 59°C for 30 seconds, and 72°C for 30 seconds. The following primers were used for cDNA amplification: COX-2 5′-TCACGCATCAGTTTTTCAAGA-3′ (forward) and 5′-TCACCGTAAATATGATTTAAGTCCAC-3′ (reverse); IL-8 5′-GGCACAAACTTTCAGAGACAG-3′ (forward) and 5′-ACACAGAGCTGCAGAAATCAGG-3′ (reverse); and as the housekeeping genes for the qRT-PCR reaction actin: 5′-ATTGCCGACAGGATGCAGAA-3′ (forward) and 5′-GCTGATCCACATCTGCTGGAA-3′ (reverse). Results were expressed as relative fold induction of the target genes relative to the reference gene. Calculations of relative expression levels were performed by using the $2^{-\Delta\Delta Ct}$ method and averaging the values of at least 3 independent experiments.[15]

2.9 | Antibiofilm activity

To test ApoE-derived peptide antibiofilm activity, bacteria were grown overnight in MHB (Becton Dickinson Difco, Franklin Lakes, NJ) and then diluted to 1×10^8 CFU/mL in BM2 medium[26] containing the peptide under test. Incubations were carried out for 24 hours to test peptide effects on biofilm formation.[20] At the end of the incubation, the crystal violet assay was performed. To do this, the planktonic culture was removed from the wells, which were washed 3 times with sterile PBS prior to staining with 0.04% crystal violet (Sigma Aldrich, Milan, Italy) for 20 minutes. The colorant excess was eliminated by 3 successive washes with sterile PBS. Finally, the crystal violet was solubilized with 33% acetic acid and samples optical absorbance

values were determined at 600 nm by using an automatic plate reader (MicrobetaWallac 1420, Perkin Elmer, Waltham, MA, USA).

2.10 | Analysis of static biofilm growth by confocal microscopy

Bacterial biofilm was grown on glass cover slips in 24-well plates in BM2 medium in static conditions at 37°C for 24 hours. To do this, *P. aeruginosa* PAO1 and *MRSA* WKZ-2 bacterial cells from an overnight culture were diluted to about 2×10^8 CFU/mL and then seeded into wells together with ApoE (133-167) or ApoE (133-150) at a concentration of 10 μM. Following incubation (24 h at 37°C), nonadherent bacteria were removed by gently washing with sterile PBS. Glass cover slips were then mounted in PBS. Viability of cells embedded into biofilm structure was determined by sample staining with LIVE/DEAD® BacLight™ Bacterial Viability kit (Molecular Probes ThermoFisher Scientific, Waltham, MA, USA). Staining was performed accordingly to manufacturer instructions. Biofilm images were captured by using a confocal laser scanning microscopy (Zeiss LSM 710, Zeiss, Germany) and a 63× objective oil immersion system. Biofilm architecture was analyzed by using the Zen Lite 2.3 software package. Each experiment was performed in triplicate. All images were taken under identical conditions.

2.11 | Statistical analysis

Statistical analysis was performed by using a Student's *t* test. Significant differences were indicated as $^*P < .05$, $^{**}P < .01$, or $^{***}P < .001$.

3 | RESULTS AND DISCUSSION

3.1 | Recombinant production of ApoE (133-167)

Recombinant ApoE (133-167) has been produced by using a previously set up experimental procedure.[27] Briefly, ApoE (133-167) coding sequence was cloned into the expression vector pET22b$^{(+)}$ downstream to ONC-DCless-H6 coding sequence.[27] The resulting fusion protein, composed by ApoE (133-167) fused to a mutated version of amphibian ribonuclease onconase, contains a His tag sequence, suitable for an easy purification of the fusion protein, between the onconase moiety and the ApoE moiety, a flexible linker (Gly-Thr-Gly), and a dipeptide (Asp-Pro), that allows to separate the carrier from the peptide by mild acidic cleavage. After protein construct expression induction in *E. coli* cells strain BL21 (DE3), inclusion bodies were collected and washed. Protein construct was then purified by affinity chromatography. Pooled fractions were extensively dialyzed against 0.1 M acetic acid pH 3.0 at 4°C and finally clarified. Following the removal of any insoluble material by centrifugation and filtration, the sample containing the fusion construct was acidified to pH 2.0 with HCl to cleave the Asp-Pro linker peptide. Sample was then purged with N_2, and incubated at 60°C for 24 hours in a water bath. Following cleavage, the pH of the sample was increased to 7.0–7.2 by adding NH_3, and the sample was then incubated overnight at 28°C to selectively precipitate the carrier onconase. The released peptide was then isolated from the insoluble components through repeated cycles of centrifugation and finally lyophilized. The purity of the peptide was checked by SDS polyacrylamide gel electrophoresis and by mass spectrometry analyses. Only peptides purified to the >95% level were subjected to further analyses. The same experimental procedure was used to express and purify ApoE (133-150) peptide.[15]

3.2 | Antimicrobial activity of recombinant ApoE (133-167)

To evaluate the antibacterial activity of ApoE (133-167), antimicrobial assays were carried out on a panel of Gram-negative and Gram-positive bacterial strains (Table 1).

Apolipoprotein E (133-167) antimicrobial activity was analyzed accordingly to broth microdilution method.[23] By performing this assay, minimal inhibitory concentration values were found to be comprised between 1.2 and 40 μM against all the tested strains. As expected from their very similar AS values (12.3 and 11.8), the toxicity of ApoE (133-167) was found to be comparable with that exerted by the previously described ApoE (133-150) peptide,[15] used as control peptide. This was also confirmed when the killing rate of the 2 ApoE-derived peptides was analyzed on 2 Gram-positive and 2 Gram-negative bacteria. Indeed, the toxicity curves obtained for the 2 peptides were found to be identical (see figure S2A-D).

To further investigate about the antimicrobial activity of ApoE (133-167), it has been also performed an alternative assay to define whether the peptide was lytic. Antimicrobial peptides generally kill bacterial cells by forming pores into membranes, thus determining the leakage of small molecules, such as ATP, which is synthesized on the inner side of the membrane. To verify if ApoE (133-167) was able

TABLE 1 Minimum inhibitory concentration (MIC_{100}, μM) values of apolipoprotein E (ApoE)-derived peptides against a panel of Gram-positive and Gram-negative bacteria

Gram-Positive Strains	ApoE (133-150) MIC_{100} Values (μM)	ApoE (133-167)	Gram-Negative Strains	ApoE (133-150) MIC_{100} Values (μM)	ApoE (133-167)
MRSA WKZ-2	10	5	*Escherichia coli* ATCC 25922	10	10
S. aureus ATCC 29213	5	5	*Pseudomonas aeruginosa* ATCC 27853	10	5
B. globigii TNO BMO13	1.2	1.2	*S. enteritidis* 706 RIVM	20	10
			Pseudomonas aeruginosa PAO1	20	40

Values were obtained from a minimum of 3 independent trials.

to induce ATP release, we performed an assay to detect the leakage of this nucleotide. To this purpose, *MRSA* WKZ-2 and *E. coli* ATCC 25922 bacterial strains were treated with 2 different doses of both ApoE-derived peptides (see section 2). As shown in Figure 1, the presence of ATP in the culture media was clearly evident; hence, we can confidently suppose that ApoE (133-167) as well as ApoE (133-150) exert a lytic effect on both bacterial strains, with *MRSA* WKZ-2 cells being particularly sensitive to the peptide.

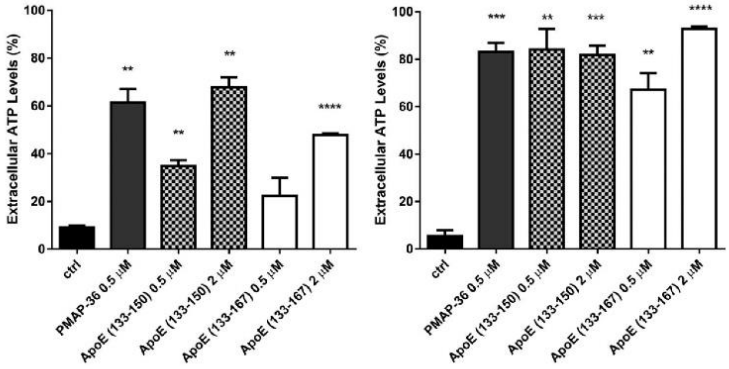


FIGURE 1 Apolipoprotein E-derived induced ATP leakage in (left) *Escherichia coli* ATCC 25922 and (right) *MRSA* WKZ-2 upon treatment with 0.5 or 2-μM peptide. PMAP-36, a canonical pore forming peptide, was used as positive control.[21] The assays were performed in 3 independent experiments

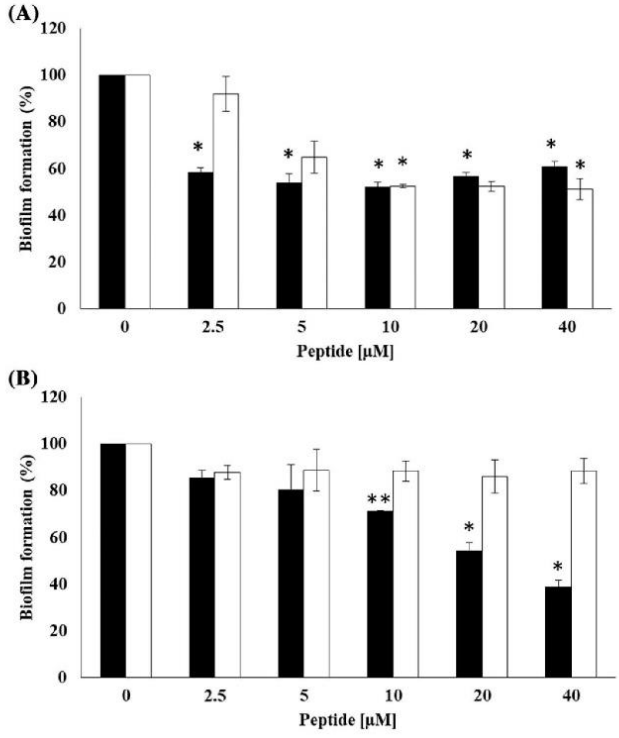


FIGURE 2 Antibiofilm activity of ApoE (133-167), black bars, and ApoE (133-150), white bars, on (A) *MRSA* WKZ-2 and (B) *Pseudomonas aeruginosa* PAO1 strains in BM2 medium. The effects of increasing concentrations of peptides were evaluated on biofilm formation. Biofilm was stained with crystal violet and measured at 600 nm. Data represent the mean (±standard deviation, SD) of at least three independent experiments, each one carried out with triplicate determinations. *$P < .05$ or **$P < .01$ was obtained for control versus treated samples

3.3 | Antibiofilm properties of ApoE (133-167)

Antibiofilm peptides represent a very promising approach to treat biofilm-related infections generally recalcitrant to conventional antibiotics.[28] To verify whether ApoE (133-167) peptide exerts antibiofilm activity, we tested its effects on biofilm formation. To do this, we incubated *MRSA* WKZ-2 and *P. aeruginosa* PAO1 cells with ApoE (133-167) for 24 hours at 37°C. The same experimental procedure was also applied to ApoE (133-150) because its antibiofilm activity has not been characterized yet. By crystal violet assay, both peptides were found to strongly inhibit (about 40% inhibition) *MRSA* WKZ-2 biofilm formation (Figure 2A). However, it has to be noticed that ApoE (133-150) is effective at a concentration (10 μM) corresponding to its MIC_{100} value on planktonic cells. Based on this, we may not exclude that peptide effects on bacterial biofilm are indeed related to its ability to directly kill planktonic cells. In the case of *P. aeruginosa* PAO1, instead, only ApoE (133-167) was found to exert a significant effect on biofilm formation (about 60% inhibition at 20 μM), whereas the shorter peptide was found to be almost inactive (Figure 2B). We also analyzed the effects of the 2 peptides on biofilm formation by confocal microscopy. To this purpose, bacterial cells were incubated with peptides (10 μM) for 24 hours at 37°C in static conditions. Ciprofloxacin antibiotic was used as a positive control (Figures 3 and 4). Following incubation, samples were double stained (Syto9/PI mixture) to discriminate between live and dead bacterial cells embedded into biofilm structure. As shown in Figure 3, a dramatic decrease of cell cohesion with respect to control sample was observed upon treatment with both peptides in the case of *MRSA* WKZ-2. This indicates a drastic inhibition of biofilm formation, in agreement with crystal violet observations. Moreover, in both cases, a significant increase in cell death was observed (Figure 3). As previously reported, for ApoE (133-150) peptide, we may not exclude that its effects on bacterial biofilm are indeed related to its toxicity toward planktonic cells because peptide concentrations lower than 10 μM were found to be ineffective on biofilm formation. When peptide effects on *P. aeruginosa* PAO1 biofilm formation were analyzed, we observed that both peptides induce a dramatic decrease of cell cohesion, although only ApoE (133-167) peptide induces a significant increase in cell death (Figure 4). This might justify the results obtained by crystal violet assay, indicating that the shorter peptide has almost no effect on *P. aeruginosa* PAO1 biofilm formation (Figure 4). Altogether, these findings support the hypothesis that ApoE (133-167) peptide is endowed with additional properties absent or limited in the case of ApoE (133-150).

3.4 | Cytotoxicity assays of ApoE (133-167) on eukaryotic cells

The promising interest in the use of antimicrobial peptides as alternative antibiotics stems from their selective toxicity toward bacterial

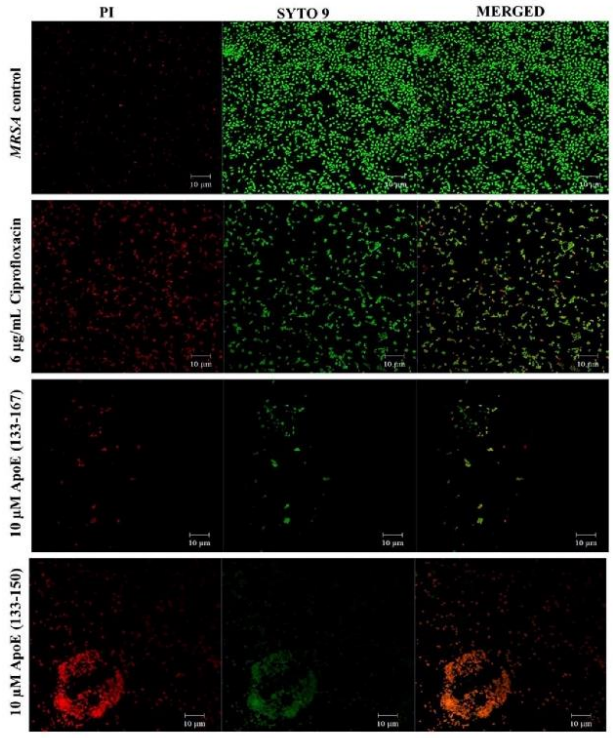


FIGURE 3 Effect of ApoE (133-167) and ApoE (133-150) on *MRSA* WKZ-2 biofilm grown in BM2 medium and visualized by controlled low strength material. Bacterial biofilm was cultivated in BM2 medium for 24 hours at 37°C together with each peptide (10 μM) or ciprofloxacin used as positive control (6 μg/mL). Biofilm images were captured by using a confocal laser scanning microscopy (Zeiss LSM 710, Zeiss, Germany) and a 63× objective oil immersion system. Biofilm architecture was analyzed by using the Zen Lite 2.3 software package. Each experiment was performed in triplicate. Scale bars correspond to 10 μm in all the cases

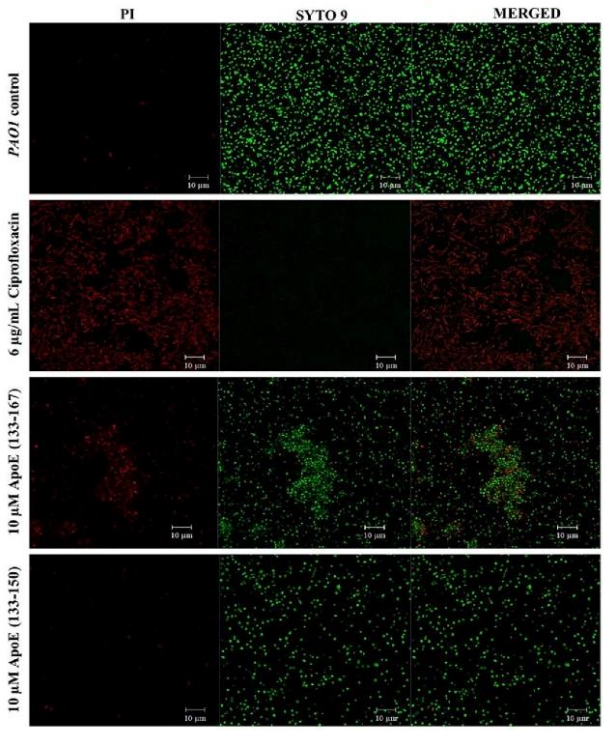


FIGURE 4 Effect of ApoE (133-167) and ApoE (133-150) on *Pseudomonas aeruginosa* PAO1 biofilm grown in BM2 medium and visualized by controlled low strength material. Bacterial biofilm was cultivated in BM2 medium for 24 hours at 37°C together with each peptide (10 µM) or ciprofloxacin used as positive control (6 µg/mL). Biofilm images were captured by using a confocal laser scanning microscopy (Zeiss LSM 710, Zeiss, Germany) and a 63× objective oil immersion system. Biofilm architecture was analyzed by using the Zen Lite 2.3 software package. Each experiment was performed in triplicate. Scale bars correspond to 10 µm in all the cases

cells.[29] We thus analyzed the cytotoxic effects and possible hemolytic effects of ApoE (133-167). As shown in figure S3A and B, the addition of peptide increasing concentrations (from 0.6 to 40 µM) to either differentiated or undifferentiated human THP-1 cells for 24 hours at 37°C did not result in any significant reduction in cell viability. Moreover, ApoE (133-167) was found to be not hemolytic for murine erythrocytes, as reported in figure S3C. This suggests that the presence of this peptide in the systemic circulation would not cause toxic effects on the blood cells, thus confirming the ability of this promising peptide to selectively recognize distinctive structural targets on bacterial cell membranes. Apolipoprotein E (133-167) peptide eventual toxicity was also tested on human keratinocytes (HaCat cells); human tumor cell line, such as HeLa, HEK-293, and CaCo-2 cells; and on murine macrophage RAW 264.7 cells. In all the cases, no significant cytotoxicity was detected (data not shown), clearly demonstrating that this peptide exerts selective toxic effects on prokaryotic cells.

3.5 | Conformational analyses of ApoE (133-167)

Apolipoprotein E (133-167) peptide conformation was analyzed by Far-UV CD. We found that the peptide is largely unstructured in PBS and that it assumes an increasingly ordered conformation in the presence of membrane mimicking agents, such as TFE and SDS micelles (Figure 5A and B). This clearly indicates that, similarly to other cationic antimicrobial peptides,[19,20] ApoE (133-167) is prone to assume a stable conformation when interacting with membrane mimicking agents. At the highest concentrations of TFE and SDS, the CD spectra showed 2 broad minima, at 208 and 222 nm, respectively, and 1 maximum at <200 nm, indicative of a putative alpha-helical conformation. It is worth noting that the maximal structured conformation was observed at a TFE concentration of about 30%, thus indicating a high propensity of ApoE (133-167) to acquire an ordered structure.

To further characterize the structural properties of ApoE (133-167), we analyzed by CD spectroscopy its binding to LPS, the main constituent of the outer membrane of Gram-negative bacteria (Figure 5C and D). When ApoE (133-167) was analyzed in the presence of increasing concentrations (0.05 to 0.8 mg/mL) of LPS purified from 2 different strains of *P. aeruginosa* (strain P10 and a cystic fibrosis isolate strain KK27), 2 minima around 208 and 222 nm appeared, clearly suggesting that the peptide is able to adopt a helical conformation upon interaction with this bacterial wall determinant (Figure 5C and D).

3.6 | ApoE (133-167) interaction with LPS

Because data obtained by CD analyses suggested an interaction between ApoE (133-167) and LPS, it seemed intriguing to provide a functional evidence for this hypothesis. To this purpose, we

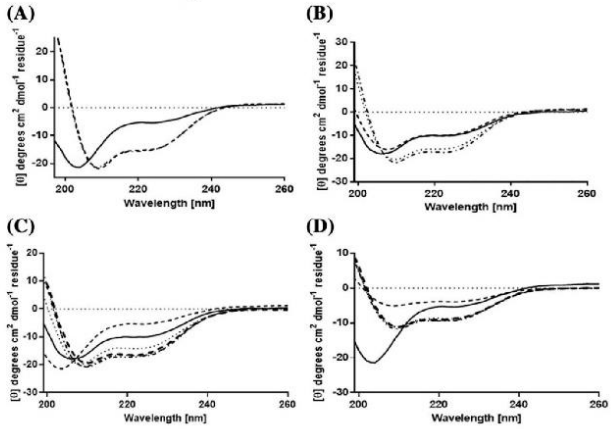


FIGURE 5 Far-UV circular dichroism spectra of ApoE (133-167); continuous line) in the presence of increasing concentrations of (A) sodium dodecyl sulfate, (B) tetrafluoroethylene, and (C and D) lipopolysaccharide

performed vitality assays by incubating LPS isolated from *E. coli* 011: B4 with 4.4 μg/mL of ApoE (133-167) for 30 minutes at 37°C. Mixtures were then added to *E. coli* ATCC 25922 cells for 4 hours. The same experimental procedure was also applied to ApoE (133-150). As reported in figure S4, bacterial vitality upon treatment with ApoE (133-167)/LPS mixture was found to be significantly lower than that observed upon administration of the peptide alone. In the case of ApoE (133-150) peptide, instead, no differences were detected when the peptide was administrated alone or in combination with LPS. This might indicate that, in our experimental conditions, ApoE (133-167) peptide is indeed sequestered by LPS, thus exerting a lower antimicrobial activity. This opens a new scenario in which ApoE (133-167) peptide biological activities go beyond its antimicrobial properties.

3.7 | Effects of ApoE (133-167) on the expression of LPS-induced inflammatory cytokines

The data described above would indicate that ApoE (133-167) is able to assume ordered conformations in the presence of LPS as well as to presumably interact with this endotoxin, thus determining a subsequent mitigation of intracellular signaling cascade induced by the recognition of the LPS by the core receptor complex composed of LPS-binding protein, CD14, Toll-like receptor 4, and MD-2.[30] As the ability of host defense peptides to bind to LPS could be associated to a mitigation of inflammatory cytokine production,[31] to confirm the hypothesis that also ApoE (133-167) might elicit antiinflammatory effects on human cells, it seemed appropriate to analyze its effects on the LPS induced IL-8 and Cox2 expression in human monocytic THP-1 cells in comparison with the effects exerted by ApoE (133-150), for which antiinflammatory properties have already been demonstrated.[15]

As shown in Figure 6, in THP-1 undifferentiated cells stimulated with LPS in the presence of ApoE (133-167), a significant decrease in mRNA expression of both IL-8 and Cox-2, with respect to cells treated with LPS alone, was observed. According to Pane et al 2016,[15] ApoE (133-150) was also able to influence mRNA expression

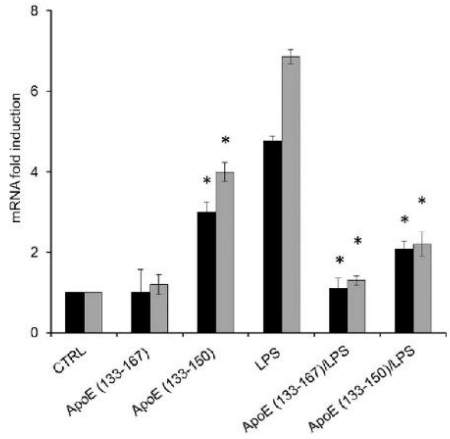


FIGURE 6 Influence of ApoE (133-167) and ApoE (133-150) on gene expression of COX-2 (black bars) and IL-8 (gray bars) in THP-1 cells treated with a mixture of 1-μM ApoE derived peptide, 2.5 μg/mL for ApoE (133–150) and 4.4 μg/mL for ApoE (133–167) and LPS (0.1 μg/mL) from *Escherichia coli* O111:B4. The results were expressed as relative fold induction respect to actin. Data were the average of 3 independent experiments ± SE

of these 2 cytokines, although to a lesser extent with respect to ApoE (133-167). This corroborates the hypothesis that ApoE (133-167) could possess in its composition and in its structure key responsive LPS scavenger determinants that are absent or limited, instead, in other ApoE-derived peptides.

4 | CONCLUSIONS

Recently, by a bioinformatics approach, we identified a cryptic peptide including the receptor binding region of ApoE, characterized by the

highest antimicrobial score among all putative or already reported ApoE-derived peptides.

Based on its length and composition, we hypothesized that this peptide, here named ApoE (133-167), could be antimicrobial but present also additional properties associated to its ability to acquire a large amphipathic conformation.

Interestingly, ApoE (133-167) was found to be able to assume a regular helical conformation in the presence of detergents and exopolysaccharides (as LPS) and to present a broad antimicrobial activity, being effective also on clinical isolates. In spite of this, our truly striking evidence is that Apo E (133-67) also showed a potent ability to inhibit biofilm formation of both Gram-positive and Gram-negative bacteria, as well as to induce a significant decrease of bacteria encapsulated in their self-produced extracellular polymeric matrix. This finding is even more significant if we consider that ApoE (133-167) revealed other intriguing properties, such as (i) the ability to interact with LPS, thus showing a potential scavenger propensity of this endotoxin; (ii) the absence of toxicity when administered to human cells; and (iii) the ability to downregulate LPS-induced proinflammatory cytokine expression.

Generally, antimicrobial peptides do not necessarily interact with a specific intracellular target that bacteria can mask or modify by genetically acquiring resistance. It should also be emphasized that this positive aspect, associated to additional biological properties of cryptic peptides, as shown for ApoE (133-167), could open an appealing future scenario in which peptides might effectively and selectively operate in conditions in which antibiotics fail. Moreover, the ability of some peptides to synergistically act in combination with conventional antibiotics[20] might allow the future design of combinatorial therapeutic approaches, which might succeed especially in those cases in which the infection is particularly difficult to eradicate due to biofilm formation.

ACKNOWLEDGEMENTS

This research was supported by the Italian Cystic Fibrosis Foundation (grant number 16/2017). We are deeply indebted to volunteers who devote many efforts in fundraising, in particular delegations "Gruppo di Sostegno FFC di San Giovanni Rotondo," "Delegazione FFC di Reggio Calabria," and Guadagnin srl.

ORCID

Elio Pizzo http://orcid.org/0000-0002-3652-8865

REFERENCES

1. Getz GS, Reardon CA. Apoprotein E as a lipid transport and signaling protein in the blood, liver, and artery wall. *J Lipid Res*. 2009;50 Suppl: S156-S161.
2. Phillips MC. Apolipoprotein E isoforms and lipoprotein metabolism. *IUBMB Life*. 2014;66(9):616-623.
3. Johnson LA, Olsen RH, Merkens LS, et al. Apolipoprotein E-low density lipoprotein receptor interaction affects spatial memory retention and brain ApoE levels in an isoform-dependent manner. *Neurobiol Dis*. 2014;64:150-162.
4. Mahley RW. Apolipoprotein E: cholesterol transport protein with expanding role in cell biology. *Science*. 1988;240(4852):622-630.
5. Mahley RW. Central nervous system lipoproteins: ApoE and regulation of cholesterol metabolism. *Arterioscler Thromb Vasc Biol*. 2016;36(7):1305-1315.
6. Wang H, Eckel RH. What are lipoproteins doing in the brain? *Trends Endocrinol Metab*. 2014;25(1):8-14.
7. Zlokovic BV. Cerebrovascular effects of apolipoprotein E: implications for Alzheimer disease. *JAMA Neurol*. 2013 Apr;70(4):440-444.
8. Giau VV, Bagyinszky E, An SS, Kim SY. Role of apolipoprotein E in neurodegenerative diseases. *Neuropsychiatr Dis Treat*. 2015;11:1723-1737.
9. Saito H, Dhanasekaran P, Baldwin F, Weisgraber KH, Lund-Katz S, Phillips MC. Lipid binding-induced conformational change in human apolipoprotein E. Evidence for two lipid-bound states on spherical particles. *J Biol Chem*. 2001;276(44):40949-40954.
10. Laskowitz DT, Thekdi AD, Thekdi SD, et al. Downregulation of microglial activation by apolipoprotein E and apoE-mimetic peptides. *Exp Neurol*. 2001;167(1):74-85.
11. Sarantseva S, Timoshenko S, Bolshakova O, et al. Apolipoprotein E-mimetics inhibit neurodegeneration and restore cognitive functions in a transgenic Drosophila model of Alzheimer's disease. *PLoS one*. 2009;4(12):e8191.
12. Clay MA, Anantharamaiah GM, Mistry MJ, Balasubramaniam A, Harmony JA. Localization of a domain in apolipoprotein E with both cytostatic and cytotoxic activity. *Biochemistry*. 1995;34(35): 11142-11151.
13. Forbes S, McBain AJ, Felton-Smith S, Jowitt TA, Birchenough HL, Dobson CB. Comparative surface antimicrobial properties of synthetic biocides and novel human apolipoprotein E derived antimicrobial peptides. *Biomaterials*. 2013;34(22):5453-5464.
14. Wang CQ, Yang CS, Yang Y, Pan F, He LY, Wang AM. An apolipoprotein E mimetic peptide with activities against multidrug-resistant bacteria and immunomodulatory effects. *J Pept Sci*. 2013;19(12):745-750.
15. Pane K, Sgambati V, Zanfardino A, et al. A new cryptic cationic antimicrobial peptide from human apolipoprotein E with antibacterial activity and immunomodulatory effects on human cells. *FEBS J*. 2016;283(11):2115-2131.
16. Wilson C, Wardell MR, Weisgraber KH, Mahley RW, Agard DA. Three-dimensional structure of the LDL receptor-binding domain of human apolipoprotein E. *Science*. 1991;252(5014):1817-1822.
17. Raussens V, Slupsky CM, Ryan RO, Sykes BD. NMR structure and dynamics of a receptoractive apolipoprotein E peptide. *J Biol Chem*. 2002;277(32):29172-29180.
18. Pane K, Durante L, Crescenzi O, et al. Antimicrobial potency of cationic antimicrobial peptides can be predicted from their amino acid composition: application to the detection of "cryptic" antimicrobial peptides. *J Theor Biol*. 2017;419:254-265.
19. Bosso A, Pirone L, Gaglione R, et al. A new cryptic host defense peptide identified in human 11-hydroxysteroid dehydrogenase-1 β-like: from in silico identification to experimental evidence. *Biochim Biophys Acta*. 2017;1861(9):2342-2353.
20. Gaglione R, Dell'Olmo E, Bosso A, et al. Novel human bioactive peptides identified in apolipoprotein B: evaluation of their therapeutic potential. *Biochem Pharmacol*. 2017;130:34-50.
21. Scocchi M, Zelezetsky I, Benincasa M, Gennaro R, Mazzoli A, Tossi A. Structural aspects and biological properties of the cathelicidin PMAP-36. *FEBS J*. 2005;272(17):4398-4406.
22. Pizzo E, Varcamonti M, Di Maro A, Zanfardino A, Giancola C, D'Alessio G. Ribonucleases with angiogenic and bactericidal activities from the Atlantic salmon. *FEBS J*. 2008;275(6):1283-1295.
23. Wiegand I, Hilpert K, Hancock RE. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat Protoc*. 2008;3(2):163-175.
24. Boukamp P, Popp S, Altmeyer S, et al. Sustained non tumorigenic phenotype correlates with a largely stable chromosome content during long-term culture of the human keratinocyte line HaCaT. *Genes Chromosom Cancer*. 1997;19(4):201-214.

25. Whitmore L, Wallace BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers*. 2008;89(5):392-400.

26. Li J, Kleintschek T, Rieder A, et al. Hydrophobic liquid-infused porous polymer surfaces for antibacterial applications. *ACS Appl Mater Interfaces*. 2013;5(14):6704-6711.

27. Pane K, Durante L, Pizzo E, et al. Rational design of a carrier protein for the production of recombinant toxic peptides in *Escherichia coli*. *PLoS one*. 2016;11(1):e0146552.

28. Pletzer D, Coleman SR, Hancock RE. Anti-biofilm peptides as a new weapon in antimicrobial warfare. *Curr Opin Microbiol*. 2016;33:35-40.

29. Takahashi D, Shukla SK, Prakash O, Zhang G. Structural determinants of host defense peptides for antimicrobial activity and target cell selectivity. *Biochimie*. 2010;92(9):1236-1241.

30. Park BS, Lee JO. Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp Mol Med*. 2013;45(12):e66.

31. Wang C, Shen M, Zhang N, et al. Reduction impairs the antibacterial activity but benefits the LPS neutralization ability of human enteric defensin 5. *Sci Rep*. 2016;6(1):22875. https://doi.org/10.1038/srep22875

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Zanfardino A, Bosso A, Gallo G, et al. Human apolipoprotein E as a reservoir of cryptic bioactive peptides: The case of ApoE 133-167. *J Pep Sci*. 2018;e3095. https://doi.org/10.1002/psc.3095

Appendix V: Autor contribution

Paper I: Metal-Tolerant Thermophiles: From the Analysis of Resistance Mechanisms to their Biotechnological Exploitation.
Contribution: GG wrote the manuscript.

Paper II: An ArsR/SmtB family member regulates arsenic resistance genes unusually arranged in *Thermus thermophilus* HB27.
Contribution: G.G performed experiments.

Paper III: A physicochemical investigation on the metal binding properties of *Tt*SmtB, a thermophilic member of the ArsR/SmtB transcription factor family.
Contribution: G.G performed experiments, analysed the data and wrote the manuscript.

Paper IV: Characterization of a promiscuous cadmium and arsenic resistance mechanism in *Thermus thermophilus* HB27 and potential application of a novel bioreporter system.
Contribution: G.G. performed experiments.

Paper V: Identification and characterization of a unique thermoactive arsenite methyltransferase.
Contribution: G.G performed the gene editing experiments, the biochemical experiments, analysed the data and wrote the manuscript.

Paper VI: Identification of a new heavy metal resistant strain of *Geobacillus stearothermophilus* isolated from a hydrothermally active volcanic area in Southern Italy.
Contribution: G.G. analysed the data

Additional paper:

Paper VII: Identification of Novel Cryptic Multifunctional Antimicrobial Peptides from the Human Stomach Enabled by a Computational–Experimental Platform.
Contribution: G.G. performed the experiments.

Paper VIII: Human apolipoprotein E as a reservoir of cryptic bioactive peptides: The case of ApoE 133-167.
Contribution: G.G. performed the experiments, analysed the data and wrote the manuscript.