Analysis of the thermal stability of mercuric reductase from the hot brine environment of Atlantis II in the Red Sea by site-directed mutagenesis: Structural interpretation of thermolabile and enhanced thermostable mutants

By

Mohamad Maged

A PhD Dissertation submitted in partial fulfillment of the requirements for the degree of

PhD in Applied Sciences

With specialization in:

Biotechnology

Under the supervision of

Dr. Hamza El Dorry
Professor of Biochemistry and Molecular Biology, Biology Department
American University in Cairo

Spring 2017
Abstract

The lower convective layer (LCL) of the Atlantis II (ATII) deep brine is a unique environment located in the central Red Sea at a depth of 2000-2200 meters. It is characterized by high salinity, temperature of 68°C and high concentrations of toxic heavy metals. Microbial community residing in the LCL of the Atlantis II deep brine have their cellular constituents evolutionarily adapted to their functions under these multiple environmental stressors. To cope with toxic heavy metals, such as mercury, and high temperature, microorganisms need a thermostable mercuric detoxification system. In this work, we analyze the properties of a mercuric reductase from this environment. The enzyme is a key component of the bacterial detoxification system of mercurial compounds. The gene for this enzyme was identified in a PCR-based MerA library which was established using DNA isolated from the microbial community of LCL-ATII environment and MerA-specific oligonucleotides as primers. Using a reverse genetics approach, the coding sequence of this gene was synthesized, cloned into an expression vector, protein expressed in E. coli and the recombinant MerA enzyme (named ATII-LCL-NH) was purified. In contrast to a homolog described earlier from the same environment (MerA ATII-LCL), the ATII-LCL-NH enzyme was found to be strongly inhibited by NaCl and it does not have the halophilic signature present in ATII-LCL (replacement of nonpolar amino acids with acidic residues and decrease of the hydrophobicity of the protein). These structural features of ATII-LCL-NH indicate that most probably the ATII-LCL-NH evolved in a microorganism that utilizes the compatible solutes strategy. As expected from an enzyme that efficiently functions in an environment characterized by high temperature, the ATII-LCL-NH was stable at 70°C. However, when compared with the halophilic and thermostable MerA ATII-LCL ortholog, the ATII-LCL-NH was much sensitive to heat treatment at 70°C, and is structurally devoid of all the acidic residues (particularly 414DDDD417) and two motifs (432KPAR435 and 465KVGKFP470) that were shown to be involved in the thermostability of the MerA ATII-LCL. Toward rational designing of a MerA enzyme with high thermal stability, the two motifs and the stretch of the four acidic residues, 414DDDD417, were used to substitute their corresponding sequences in ATII-LCL-NH. One mutant in which the
stretch of the two acidic amino acids $^{415}$DD$^{416}$ (present in the ATII-LCL) substituted the corresponding amino acids in ATII-LCL-NH was found to be more thermostable than ATII-LCL-NH. However, mutants in which the two motifs were used alone, or in conjunction with $^{414}$DDDD$^{417}$, were extremely unstable to heat treatment. Comparison of the total hydrogen bonds and salt bridges in ATII-LCL-NH and its mutants indicate that a major alteration of the hydrogen bonding occurs in all the mutants that are not stable to heat indicating that subtle structural alteration of the MerA molecule is responsible for the loss of thermostability. The work shows that although structural signature of halophilic proteins could be correlated with the decrease in the hydrophobic contact surface, the bases of thermostability of proteins, judged by the ATII-LCL-NH, are indefinable and do not associate with specific features related to functional amino acids content of the molecule.
Dedication

This Thesis is dedicated to my family to whom I am eternally grateful for their support, guidance and patience throughout my life.
Acknowledgements

I am grateful to my family who has supported me by all means. Their guidance and dedication was and will always be my driving force. They have endured moments of ups and downs throughout my PhD and personal life and always lifted me up and pushed me forward. Their enthusiasm and continuous relentless encouragement lit up my way. They had total faith in me and I will always do my best to make them proud and live up to their expectations. My parents taught me how to give without expecting any return and I hope to carry on doing so as much as could be for the rest of my life.

Many thanks to Dr. Hamza El Dorry, my thesis supervisor, from whom I have learned a lot. I am also thankful to my colleagues at the Science and technology research center of the AUC (STRC) for their encouragement, help and moral support.

Special thanks to Mr. Ahmed El Husseiny for his generous help with developing the 3D structural models’ figures. I also thank Mr. Amged Ouf for providing the Atlantis II LCL DNA.

Finally, I thank everyone who has assisted me in any way or at any time and helped me conclude my PhD dissertation.
Table of Contents

Abstract .................................................................................................................................................. II

Dedication ............................................................................................................................................... IV

Acknowledgements ............................................................................................................................ V

Table of Contents ................................................................................................................................ VI

Table of abbreviations ......................................................................................................................... IX

List of Figures .......................................................................................................................................... XI

Chapter 1: Introduction ........................................................................................................................ 1

I. The Red Sea ....................................................................................................................................... 1

II. Deep brine environment of Atlantis II ......................................................................................... 2

III. Mercury and microbial detoxification of Mercury ................................................................... 3
    a) Mercury ........................................................................................................................................... 3
    b) Microbial resistance to mercury ............................................................................................... 4
    c) Mercuric reductase (MerA) ........................................................................................................ 7

IV. Halophilic microorganisms and proteins ...................................................................................... 9

V. Thermophilic organisms and their protein adaptation ................................................................. 10
    a) Thermophiles .......................................................................................................................... 10
    b) Membrane adaptation of thermophiles ................................................................................. 10
    c) Thermozymes ....................................................................................................................... 11
    d) Molecular bases of thermal adaptation ................................................................................. 12

1. Electrostatic interactions (salt-bridge and Hydrogen bond) ..................................................... 12

2. Hydrophobic forces ..................................................................................................................... 13

3. Disulfide bridge ............................................................................................................................ 13

4. Conformational alteration .......................................................................................................... 14

5. Proline rule ................................................................................................................................... 14

VI
6. Engineering thermostable enzymes..................................................................................................15

VI. Mercuric reductase from the ATII environment (MerA ATII-LCL)...........................................17

Chapter 2: Experimental procedures ..................................................................................................20

I. Sample collection and DNA extraction .............................................................................................20

II. Establishment of mercuric reductases library from Atlantis II LCL .............................................20
    a) Designing a pair of oligonucleotides for merA amplification ..................................................20
    b) PCR amplification of Atlantis-II LCL mercuric reductases .........................................................21
    c) Library construction in pGEM vector .........................................................................................22

III. Generation of merA mutants using site-directed mutagenesis ....................................................22

IV. Expression and purification of recombinant MerA isoforms .........................................................24

V. Mercuric reductase activity assay ....................................................................................................25

VI. Heat Stability Analysis ....................................................................................................................26

VII. Salt tolerance Analysis ..................................................................................................................26

VIII. Modeling of the Three-dimensional Structure of the wild-type ATII-LCL-NH .......................27

IX. Salt Bridge prediction ......................................................................................................................27

X. Hydrogen Bond prediction ...............................................................................................................27

XI. Venn Diagram plotting ....................................................................................................................27

Chapter 3: Results and discussion .......................................................................................................29

I. Establishment of mercuric reductases library from Atlantis II LCL .................................................29
    a) Designing a pair of oligonucleotides for merA amplification ..................................................29
    b) PCR amplification of Atlantis II LCL mercuric reductases .........................................................29
    c) Library construction in pGEM vector .........................................................................................29

II. Multiple sequence alignment of non-redundant and full-length MerA sequences ........................30

III. Generation of MerA mutants using site-directed-mutagenesis .....................................................33
    a) Optimization of codon usage of the ATII-LCL-NH gene sequence ...........................................33
b) Amino acid differences between MerA ATII-LCL and ATII-LCL-NH ............................................. 34

C) Mutants generated in this work ........................................................................................................ 36

IV. Expression and Purification of ATII-LCL-NH and its mutants .................................................... 37
   a) Optimization of the induction process of the recombinant MerA ATII-LCL-NH and its mutants 37
   b) HisTag affinity column and superdex 75 size exclusion chromatography purification of MerA ATII-LCL-NH recombinant protein................................................................................................. 37

V. Effect of NaCl on the activity and kinetic parameters of ATII-LCL-NH and its mutants ......... 39
   a) Inhibition by NaCl .......................................................................................................................... 39
   b) Kinetic parameters of the ATII-LCL-NH and its mutants ......................................................... 41

VI. Thermal stability of ATII-LCL-NH merA and its mutants ............................................................ 42

Chapter 4: Conclusion .......................................................................................................................... 53

Future prospects .................................................................................................................................. 55

References ............................................................................................................................................ 56

Supplement ........................................................................................................................................ 61
# Table of abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATII</td>
<td>Atlantis II brine pool</td>
</tr>
<tr>
<td>ATII-LCL MerA</td>
<td>Halophilic mercuric reductase from Atlantis II LCL brine</td>
</tr>
<tr>
<td>ATII-LCL-NH MerA</td>
<td>Non-halophilic mercuric reductase from Atlantis II LCL brine</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>CSAR</td>
<td>Consensus sequence of the assembled ATII-LCL reads</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FAD</td>
<td>Flavin adenine dinucleotide</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>KAUST</td>
<td>King Abdullah University for Science and Technology</td>
</tr>
<tr>
<td>Kcat</td>
<td>The number of substrate molecules each enzyme site converts to product per unit time</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani media</td>
</tr>
<tr>
<td>LCL</td>
<td>Lower convective layer</td>
</tr>
<tr>
<td>MerA</td>
<td>Mercuric ion reductase</td>
</tr>
<tr>
<td>NADPH</td>
<td>Nicotinamide adenine dinucleotide phosphate</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NH</td>
<td>Non-halophilic enzyme</td>
</tr>
<tr>
<td>NmerA</td>
<td>N-terminal domain</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>O/N</td>
<td>Overnight</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffer saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethylsulfonyl fluoride</td>
</tr>
<tr>
<td>SDM</td>
<td>Site-directed mutagenesis</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>Ta</td>
<td>Annealing temperature</td>
</tr>
<tr>
<td>Tn501</td>
<td>Transposon 501</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>$T_{50%}$</td>
<td>Temperature required for 50% enzyme inactivation</td>
</tr>
<tr>
<td>WHOI</td>
<td>Woods Hole Oceanographic Institute</td>
</tr>
<tr>
<td>Wt or WT</td>
<td>Wild-type</td>
</tr>
</tbody>
</table>
List of Figures

Figure 1: Map showing the location of different well-studied Red Sea brine pools ........ 2
Figure 2: Diagram showing the microbial mer operon ................................................. 6
Figure 3: Cartoon showing the structure of mercuric reductase dimer and the binding
and catalytic sites. ........................................................................................................ 8
Figure 4: Multiple sequence alignment of non-redundant and full-length MerA
sequences. .......................................................................................................................... 32
Figure 5: Optimization of codon usage of the ATII-LCL-NH gene. ................................. 34
Figure 6: Pairwise alignment of MerAs ATII-LCL and ATII-LCL-NH. .............................. 35
Figure 7: Illustration presenting the mutations shown in Table 3 ..................................... 36
Figure 8: Inhibition of ATII-LCL-NH and its mutants by NaCl ............................... 40
Figure 9: Salt bridge between Lys 432 of Box1 and Asp 417 in the 4D. .............................. 43
Figure 10: Thermostability of ATII-LCL-NH and its mutants ........................................ 45
Figure 11: Total hydrogen bonds of ATII-LCL-NH and its mutants ............................... 47
Figure 12: Venn diagram of hydrogen bonding in ATII-LCL-NH and its mutants .......... 49
Figure 13: Total salt bridges of ATII-LCL-NH and its mutants ....................................... 50
Figure 14: Venn diagram of salt bridges in ATII-LCL-NH and its mutants ....................... 51
Chapter 1: Introduction

I. The Red Sea

The Red Sea is the inlet of the Indian Ocean. It is located between the Arabian Peninsula and the African continent [1, 2]. It has an area of 450,000 square kilometers [1]. It is a narrow body of water extending for over 1930 kilometers from Suez, Egypt in the north to Bab el-Mandeb strait in the south. This strait opens up to the Gulf of Aden, which connects the Red Sea with the Arabian Sea. Its maximum depth is 3,040 meters and the biggest width is three hundred kilometers [3]. It separates the coasts of Djibouti, Egypt, Eritrea, Israel, Jordan Saudi Arabia, Sudan and Yemen. It is connected to the Mediterranean Sea through the Suez Canal [4] since 1869.

At its north side the sea splits into the Gulf of Aqaba and the Gulf of Suez. It fills the fault depression between the two massive land blocks of Arabia and North Africa. This rift was started by the movement of the African and Arabian blocks away from each other fifty five million years ago [3].

The Red Sea is an attractive marine environment owing to its characteristic features. It is isolated and its water is relatively warmer and more saline than other seas. The water temperature is about 26°C at the north part, while is as high as 30°C at the south side. There is only a 2 degrees variation between the winter and summer average temperatures [4]. The average salinity is 4.1% which is higher than the average of 3.5% in other seas [5]. This highly dense and saline water owes its properties to several factors, the high temperatures of the dry climate of the surrounding shores and the very low rainfall of 6 cm/year. Evaporation is distinguished with an average of two meters/year. There are no draining rivers and the water circulation is low. Also it hosts 3.8% of the coral reefs in the world [6].

There are 25 brine-filled deeps discovered so far in the Red Sea [7-10] (Figure 1). These brine pools were formed by the shift of the Arabian and African tectonic plates [7, 11].
II. Deep brine environment of Atlantis II
A brine pool is defined as a deep seafloor depression filled with brine. Continuous rifting of the Red Sea floor combined with volcanic activity has caused the formation of these hot and hypersaline mounds. The Atlantis II brine pool is the biggest and hottest of these geo-thermally distinct hypersaline basins [8]. It is located 21°20.760′N, 38°04.68′E [12] between Saudi Arabia and Sudan [13]. ATII brine pool is considered the largest ore deposit of any known hydrothermal system [14].
This deep is a harsh environment harboring diverse microbial communities adapted to its extreme conditions [15]. It was discovered in the 1960s [7, 16] and has been extensively studied since then. It is located at a depth of about 2200 m below sea surface[17]. The brine pool’s temperature is 68˚C and has gradually increased over the past few decades owing to the mounting volcanic activity in the vicinity [7]. The salinity is estimated to be 25% [13] reaching near saturation of NaCl [7, 8]. According to oxygen content, salinity and temperature, the Atlantis II depression is divided into 4 layers: upper convective layers 1-3 (UCL-1-3) and lower convective layer (LCL) [5].

The lower convective layer is the deepest with salinity seven and half times as that of sea water which is 3.5% [5]. It endures extreme hypoxic conditions, has low pH of 5.3 and temperature of 68˚C. Organisms adapted to multiple abiotic stressors present in such environment could be described as polyextremophiles [18]. Owing to technical difficulties, studying of microbial community of these deep-sea brines has not attracted sufficient attention until recently [9, 15]. Moreover, pyrosequencing of the 16S rRNA of microorganisms residing in the water column overlaying the ATII, and its sediments were reported recently [17, 19, 20]. In 2016, mobile genetic elements in the Red Sea Atlantis II basin were described as well [21].

The lower brine has a volume of approximately four cubic kilometers [14]. It has a thickness of 135 meters. Recently, several genes were identified from Atlantis II LCL sequences, cloned, expressed and were shown to be adapted to their extreme environment [13, 22, 23], namely a nitrilase, a mercuric reductase and an esterase. In addition, two antibiotic resistance enzymes were isolated and characterized from the ATII brine pool [24]. Underneath the brine pool lays the sediments layer which in turn is composed of heavy metalliferous deposits [25]. The sediments harbor a remarkable consortium of microbial community which was studied and described recently [17, 20, 26] including mobile genetic elements [27].

III. Mercury and microbial detoxification of Mercury

a) Mercury
Environmental toxins are classified as organic or inorganic [28]. These poisons pose a serious ecological hazard. Microorganisms over time have adapted to the presence of such chemicals by evolving intricate defense mechanisms of detoxification.

Mercury is one of the most toxic heavy metals on earth [29]. It is ranked the 6th amongst the top 10 most toxic elements in the universe [30]. The basic anthropogenic (man-made) sources of environmental mercury contamination are burning of petroleum and coal products, the use of mercurial fungicides and the use of mercury in mines to extract gold, such that one Kg of gold produced requires the use of 1.32 Kg of mercury [29].

The toxicity of mercury to organisms lacking mercury-resistance mechanisms stems from its affinity to sulfahydryl groups (-SH) of proteins and enzymes [29, 30] altering their structure and function [31]. Mercury and sulfur are soft base and soft acid and have higher affinity towards each other. The most commonly encountered form of toxic mercury in aquatic and land systems is the organic methyl mercury [32]. Although it is more toxic than inorganic mercury, it is also more volatile. Thus the methylation and accumulation of methyl mercury in the marine organisms could have been a mechanism to detoxify mercury. In nature, the cinnabar (HgS) is the most common ore and source of mercury [33].

In brine pools the mercury originates from volcanic activity and hydrothermal/geothermal sources [30, 34-36]. The exact concentration of mercury was not measured in the ATII brine pool. In contrast, the heavy metal concentrations that were determined in the sediments [37]. The presence of high concentrations of heavy metals has an impact on the evolution of dwelling microbes and their ability to adapt to heavy metal toxicity.

b) Microbial resistance to mercury

The first report of bacteria resistant to mercury was in 1960 [38]. Microorganisms can generally detoxify mercury using several approaches [28, 30, 31, 39]. The first line of defense is blocking of entry into the cell; this is achieved via the permeability barrier that prevents the mercury from getting inside the cell. Chelation is another mechanism by
which the cell produces surface polysaccharides that form a complex with the mercury ion and render it less reactive. In addition, some microbial cells produce excess thiols; to bind mercury and minimize the exposure of the cell machinery’s sulfahydryl group to Hg [39]. Sequestration inside the cell happens through the production of metal-binding proteins that move the toxic metal outside the cell. Some microbes use metal ion pumps for transport outside the cell. This is achieved by the production of specific efflux systems on the cell membrane that export the heavy metal out of the cell [40].

The most efficient and common mechanism of resistance is the conversion of mercury into a less toxic form. The presence of minimal and subtoxic concentration of mercury initiates the expression of the inducible mer-operon. This operon is composed of multiple components. The main components are mercuric reductase merA and merB genes. The MerA enzyme detoxifies inorganic Hg. It reduces the divalent mercuric ion into volatile Hg\(^0\). The MerB (organomercurial lyase) enzyme is used to detoxify organic mercury. It cleaves the carbon-Hg bond; however, its exact mechanism of action has not been intensively studied.

Genes responsible for mercury resistance are often organized in inducible operons. These operons are spread over both the bacterial and archaeal domains of life [32, 41]. The operons could be part of transposons, plasmids or sometimes located on the chromosomes [29].

The most common genes in such operons are merR that codes for a DNA-binding and Hg(II)-sensory regulatory protein. MerR protein binds the operator/promoter region in the absence of Hg(II) to repress transcription (Figure 2). When mercury is present, it associates the MerR protein and the complex binds the operator/promoter locus and unwinds the DNA. This allows the RNA polymerase to induce transcription of the mer operon’s genes.

Genes, merT, merP and merC code for transport proteins that move mercury into the bacterial cell cytoplasm. Moreover, MerP and MerT were shown to facilitate the uptake of arylmercury compounds [41].
The *protein* MerG is predicted to inhibit organomercury compounds’ entry into the cell. Its deletion does not affect resistance to mercury but rather increased membrane permeability to organomercurials namely phenyl mercury.

MerD protein plays a role in rapid repression of mer operon transcription once most of the mercury is volatilized. This is particularly important because mercuric reductase belongs to the oxidoreductase family and has oxidizing properties. So in the absence of the substrate it could form hydrogen peroxide that is toxic to the cell [41].

![Diagram showing the microbial mer operon.](image)

**Figure 2: Diagram showing the microbial mer operon.**
This diagram shows different genes of the microbial mer operon. The regulator merR, the repressor merD, the transporters merT and merP are in grey letters. The mercuric reductase; merA is in blue letters.

The gene *merA*, the mercuric-ion reductase gene plays the keystone role. It codes for MerA which converts Hg(II) into the less toxic and volatile Hg(0). Operons having these elements impart narrow spectrum resistance to mercury, while operons having the previous elements plus *merB*, confer broad-spectrum resistance.
The MerB is an organomercurial lyase that cleaves the C-Hg bond [28] and organisms possessing \textit{merB} gene have the capacity to resist organic mercury poisoning [29]. The mer operon is versatile and several additional units were discovered.

MerF is a transmembrane protein that shows similarity to MerC. MerF, was shown to help the volatilization of mercury.

MerE is found in some Gram-negative microbes and although function is not fully described, it is proposed to facilitate transport of mercury since it possesses a transmembrane module [41].

c) Mercuric reductase (MerA)

The mercuric reductase is a homodimer that belongs to the family of flavin-dinucleotide oxidoreductase; the same family as glutathione reductase [42, 43].

The MerA has two active sites located at the dimer interface (Figure 3). Two pairs of cysteines at the catalytic core (C136 and C141) are redox-active, while a pair of cysteines (C558 and C559) each at the C-terminus of the two subunits bind mercury and present it to the active site to be reduced [44, 45]. There is an additional pair of cysteines at the N-terminal domain (NmerA) that is part of the metal-binding motif (GMTCXXC), they serve as disulfide ligands to present mercury in the cytoplasm to the catalytic site in case of thiol depletion [46]. Interestingly, NmerA is homologous to merP [30].
Figure 3: Cartoon showing the structure of mercuric reductase dimer and the binding and catalytic sites.
This simplified diagram represents the structure of the mercuric reductase homodimer. The six pairs of cysteines in the merA homodimer are highlighted in yellow, and the N-terminal domain is highlighted in red. Cys 11/14 pair serves to bind Hg$^{2+}$ and transfer it from ligands in the cytoplasm to the redox-active cysteines for reduction under physiological conditions in which intercellular thiols are depleted. Cys 136/141 pair is the redox-active cysteine involved in catalysis. Cys 558/559 pair binds Hg$^{2+}$ and transfers it to the active site of the other subunit for reduction.

MerA catalyzes the two-electron reduction of divalent mercury Hg(II) into elemental mercury Hg(0) according to the equation:

\[
\text{NADPH} + \text{RS-Hg-SR} + \text{H}^+ \rightarrow \text{NADP}^+ + \text{Hg(0)} + 2\text{RSH}. 
\]

The high thiol content in the cell facilitates the complexation of inorganic mercury into dithiol form [28]. The MerA utilizes the Flavin adenine dinucleotide (FAD) as a cofactor, which mediates the passage of electrons from the Nicotinamide adenine dinucleotide (NADPH) via the active-site’s cysteines to mercury.

Mercuric reductases are ancient enzymes theorized to have originated amongst thermophilic microorganisms in geothermal ecosystems [36, 45] after the great oxidation event about 2.4 billion years ago [47]. Ionic mercury is naturally found in several hydrothermal vents and hot springs [34].

Previous studies show that various MerAs have $K_m$ values for Hg(II) in the range of 9–70 μM and specific activities in the range of 1.05–50 μmol NADPH oxidized/min/mg
The enzyme is inactive prior to contact with the substrate and then gets activated when it interacts with the Hg substrate [48].

IV. Halophilic microorganisms and proteins
Halophilic microorganisms live in environments of elevated salt concentration [49] and require more than 2.5 Molar [50, 51] for optimal growth such as those found in the Atlantis II brine pool [8, 17, 22]. These microorganisms adapt to the elevated salinity in their surrounding medium through two mechanisms, the salt-in and the salt-out approaches.

In the salt-out approach, the organism uses ion pumps to extrude considerable amounts of inorganic solutes while at the same time maintains the intracellular osmotic pressure by accumulation of organic solutes. The most commonly used organic osmolytes are glycerol, betaine, ectoine and small amino acids like glycine and histidine [49].

In the salt-in strategy, the cell accumulates molar KCl concentrations. Consequently, intracellular enzymatic machinery copes by altered amino acid composition of proteins to maintain stability at such denaturing salt concentrations. The increase in acidic residues is noticeable in many halophilic proteins [52] from (salt-in) organisms as compared to their non-halophilic homologs [50], and those coming from organisms adapted to the salt-out approach. Halophilic proteins generally, display a lowered frequency of hydrophobic moieties and a decreased hydrophobic contact surface. The propensity in polar and charged surface amino acids [53] allow the formation of more salt-bridges, electrostatic interactions of surface residues with salt [49] and the formation of a hydration shell that enhances protein stability and solubility [54]. Halophilic proteins require the presence of high salt concentration to function and remain soluble, whereas their activity and stability is significantly reduced in the absence of salt. Some halophilic enzymes adapt by altering their structural conformation and enthalpic mechanism in order to optimize surface residues-solvent-interactions [55].
V. Thermophilic organisms and their protein adaptation

a) Thermophiles

Generally, thermophiles are able to grow at temperatures above 60°C [56]. They adapt their enzymatic and protein machinery to elevated temperatures using structural variations in secondary [57], and three-dimensional configurations, and show an altered amino acid composition [56] relative to thermo-labile homologs. Also they tend to accumulate extrinsic factors like compatible solutes and cofactors [58]. They use molecular chaperonins [57] and have the ability to regulate the expression of other stress-response enzymes [56]. According to optimal growth conditions microorganisms are classified as psychrophiles (-5 to 20°C), mesophiles (15–45°C), thermophiles (45–80°C) or generally > 55°C, moderate thermophiles (> 65°C), extreme thermophiles (> 75°C) [59] and hyperthermophiles ≥ 80°C [60].

With regards to temperature, both mesophilic and thermophilic proteins show maximum stability at room temperature [61]. However, thermophilic proteins—although highly thermostable under natural high temperatures—show low catalytic activity at room temperature [62], and thus may be of little practical use under lower temperatures [62].

b) Membrane adaptation of thermophiles

Both thermophilic archaea and bacteria have adapted their lipid membranes to withstand high temperatures sometimes exceeding the boiling point of water [63]. This is necessary to protect the intracellular macromolecules and structural elements required for the continuity of life. Bacterial lipid membranes are mainly composed of Glycerol-3-phosphate, fatty acid chains and ester linkages while archaea have glycerol-1-phosphate, hydrocarbon isoprenyl chains and ether bonds. Some thermophilic bacteria have an ether linkage similar to archaea [64]. Thermophilic archaea and bacteria generally have distinguished lipid composition such as the dicarBoxylic acid long chain of *Aquifex pyrophilus*, which grows at 85°C, and the cyclohexyl fatty acid lipid of *Bacillus acidocaldarius* that grows at 65°C. Some microbial species are able to regulate their lipid content and composition such as the *Methanocaldococcus jannaschii*
(methanoarchaea) that lowered its dither lipids to 20% and increased archaeol-based cyclic lipids to 40% when their growth temperature changed from 45˚C to 65˚C [64].

One characteristic feature of thermophiles is that their membrane maintains the liquid in the crystalline state required for vitality. For archaea, there is no phase-transition of the ether lipids in their physiological temperature range 0-100˚C owing to the isoprenoid structure and ether bonds. While in bacteria some adaptation takes place. For instance; alteration of the ratio of saturated to unsaturated fatty acids and the percentage of iso and anteiso fatty acids [64].

Additionally, thermophilic membranes have to maintain an efficient permeability barrier at higher growth temperatures towards protons. This barrier allows the generation of the proton motive force across the membrane, and energy production of the cell at elevated temperatures. This property is less efficient in some thermophilic bacteria and hence they rely more on sodium pumps rather than proton pumps for energy transduction [65].

**c) Thermozymes**

Thermozymes by definition are enzymes produced by thermophilic or hyperthermophilic organisms. They are classified as thermostable if they resist elevated temperatures, or thermophilic if they are optimally active at high temperatures above 60˚C [60]. The term stability comprises either kinetic or thermodynamic aspects or both. Kinetic thermal stability is defined by the half-life (t\(_{1/2}\)) of the protein at a certain temperature, while thermodynamic stability is defined by the free energy of stabilization and the melting temperature (t\(_m\)) [60]. On the other hand, enzymes produced by mesophiles are called mesozymes.

Single amino acid changes can account for enhanced thermal stability. And a group of amino acid changes in specific sites and critical locations on the polypeptide chain have an additive effect on the protein heat stability. However, there is no concept to predict such sites or amino acid residues to be substituted because each protein behaves in a distinguished manner [66].
d) Molecular bases of thermal adaptation

Variations in structural conformation and intermolecular interactions define the bases for thermal adaptation of proteins [60]. Intermolecular interactions can be classified as covalent or non-covalent forces. Electrostatic forces (salt-bridge also known as ion-pair and Hydrogen-bond), and hydrophobic interactions belong to the class of non-covalent mechanism, while disulfide bride formation belongs to the class of covalent mechanism of adaptation.

1. Electrostatic interactions (salt-bridge and Hydrogen bond)

Three types of electrostatic interactions stabilize proteins: salt-bridge (charge-charge), charge-dipole and dipole-dipole. Salt-bridge occurs between amino acid residues of opposing charges such as a salt-bridge between aspartic acid and lysine. Generally, thermozymes show a propensity of charged residues [67] noticeably on the surface as compared to mesozymes suggesting more salt-bridges.

Salt-bridge formation between oppositely charged amino acids on the protein surface is a frequently encountered mechanism of non-covalent thermal adaptation such as the bridge forming between aspartic acid/glutamic acid side chain oxygen or carboxyl oxygen with lysine/arginine/histidine nitrogen or amide group of the N-terminal. The distance between the participating atoms is by definition less than 4 Angstroms [68]. Salt bridge’s energy of stabilization is approximately three Kcal/mol [69]. A number of salt bridges can group together to form a network with a synergistic energy of stabilization. One example of such ion pair involvement in thermal stability is the Glutamate Dehydrogenase from *T. kodakarensis*. Substitution of a threonine with glutamic acid significantly improved temperature resistance due to formation of an ion pair at the subunit-subunit interface. Conversely, a substitution of a glutamic acid with a glutamine resulted in the reduction of temperature tolerance of the mutated enzyme due to the disruption of a salt bridge [59]. Hence, salt bridge formation or disruption in site-directed mutagenesis studies may improve or lower thermal stability depending on the location of the bridge and its geometry [70, 71].
Hydrogen bond contributes to additional stabilization energy of the molecule by one kcal/mol [72]. It occurs within the polypeptide backbone and also occurs between polypeptide chains and the surrounding water/aqueous medium.

The distance between donor and acceptor atoms is 1.5-2.6 Angstroms [72]. Some individual substitutions of charged amino acids on the surface of the protein increase the half-life of the protein under high temperature and its thermal stability by formation of a hydrogen bond [73] with a neighboring residue or with the aqueous solvent. Both hydrogen bonds and salt bridges could help stabilize protein molecule by optimizing secondary structure interactions and subunits association to form more stable tertiary and quaternary structures [68].

Both salt bridges and hydrogen bonds are more densely distributed along the protein surface and binding interfaces because they are more hydrophilic than the hydrophobic core. Such bonds play a major role in binding of water molecules in a particular order to stabilize the protein macromolecule [68]. The geometry of water-protein interaction is crucial to both protein folding and binding functionality.

2. Hydrophobic forces

Hydrophobic force is another type of non-covalent interactions such as the weak Van der Vaals forces. It is defined by the “ratio of the buried non-polar surface area to the total molecular non-polar surface area”. Each buried methyl group increases thermal stability by about 1.3°C [74]. Hydrophobicity could contribute to additional energy of stabilization of the molecule by 0.5 kcal/mol [72]. In a protease isolated from *Bacillus stearothermophilus* the substitution of one glycine amino acid by alanine (additional methyl group) lead to significant improvement of thermal stability by increasing hydrophobicity and stabilization of a localized secondary structure α-helix [59]. Hydrophobicity is a key contributor to protein thermal stability.

3. Disulfide bridge
Disulfide bridge is an example of covalent mechanism of temperature adaptation. It is formed through cysteine substitutions at specific sites. The bridge imparts thermal stability mainly by lowering entropy of the unfolded state of the protein.

4. **Conformational alteration of thermozymes relative to mesozymes account for thermal adaptation**

One prominent feature of thermozymes is that they are less flexible and more rigid to resist unfolding. Rigidity is produced by optimized electrostatic interactions, reduction of conformational strains and stabilization of the α-helices [60]. Some thermozymes owe their thermostability to having a highly compact configuration and a core that is more tightly packed. They also show more hydrophobic properties. This is accompanied with burying hydrophobic residues and side chains to minimize their interaction with the solvent and result in a tightly packed structure [58]. A remarkable example is the substitution of alanine by isoleucine (higher hydrophobic character) and that alone resulted in the increase of the T_m of the mutated protein Ala31Ile by 11°C [60, 75]. Another example is substitution of lysine with methionine, which resulted in the increase of T_m of the mutated protein Lys35Met by 14°C [75]. Some amino acid substitutions reduce the conformational entropy of the unfolded state. The higher the entropy of the unfolded state, the lower the thermal stability is. Glycine has the highest conformational entropy and thus mutation of glycine at favorable positions result in improving thermal stability. Stabilization of the protein by the formation of α-helix is attributed to the α-helix dipole moment. At the N-terminus there is 0.6 positive unit charge and at the C-terminus a 0.6 negative unit charge. The oppositely charged dipoles are stabilized by electrostatic interaction with positively and negatively charged residues at the corresponding terminus of the α-helix. Valine, isoleucine and threonine destabilize α-helices and their substitution at favorable alpha helix positions could enhance protein thermal stability [76]. Adjacent anti-parallel α-helical configuration imparts a stabilization energy of 7 kcal/mol [60].

5. **Proline rule**
One principle to account for differences in heat stability between thermophilic and thermolabile proteins is the proline rule [66]. It is proposed that certain proline substitutions at critical sites on the protein backbone lower conformational entropy of the unfolded state and thus stabilize the protein in its native state at elevated temperatures. It also accounts for rising rigidity and restriction of polypeptide flexibility and therefore improves protein thermal stability in a cumulative manner [77]. In some proteins the critical locations were found on first turns of α-helices, flexible loops and the second sites on β turns. In other proteins those sites were random [78]. Conversely, proline substitutions at undesirable sites result in disruption of conformation and lowered protein thermal stability [79] and thus site-specific mutation and/or replacement has to be taken into consideration when engineering a thermostable enzyme.

Conversely to glycine, proline’s pyrididine ring restricts its possible configurations and thus has the lowest conformational entropy. One example is the D-Xylose Isomerase. Substituting glycine at position 138 of the wild-type with proline led to an elevation of thermal stability by 100% of the G138P mutant. Moreover, a double mutant combining the G138P with G247D (replacing glycine 247 of the wild-type with aspartic acid) show a 2.5 times higher half-life. Proline was introduced at a random coil and is predicted to enhance the rigidity of the structure and lower the flexibility of the backbone [80]. Similarly, the hyperthermophile *T. kodakaraensis* Ribonuclease HII has numerous proline residues at the N-terminal of several α-helices. Several mutants were designed to either add additional proline at the N-terminal of the α-helices or remove the already present proline residues. Replacing proline residues at positions 46, 70 and 174 lead to reduction of heat stability while substitution of glutamic acid at position 111 and lysine at position 199 with proline lead to the increase of the thermo-stability of the Ribonuclease HII enzyme [81].

6. **Engineering thermostable enzymes**

Site-directed mutagenesis studies are carried out to study the effect of addition, deletion or substitution of amino acid residues at specific positions on the structure and function of enzymes. One of the main objectives of such studies is to engineer biocatalysts with
superior and desirable characteristics [82]. For such purpose there are three approaches: rational design, directed-evolution and *de novo* synthesis [83].

In the rational design, we have a previous and extensive knowledge of the enzyme’s mechanism of action and catalytic behavior, in addition, the three-dimensional structure is already known and thus the mutagenesis work is specific.

In directed evolution, a set of mutants is created and enzymes with favorable properties are selected.

In *de novo* synthesis, computer algorithms are used to design artificial enzymes based on prior knowledge of three-dimensional conformation and folding of naturally-found proteins [82]. There are many examples in literature that correlate single amino acid substitutions with alter thermal stability of the resulted isoforms or mutants [84]. Sometimes the pronounced effect of one amino acid mutation exceeds a local change in the secondary or tertiary structure to a wide-spread impact on the entire molecule [84].

The alkaline protease from *B. pumilus* is one example of individual amino acid replacement and its effect on protein temperature resistance. The valine at position 149 in the wild-type is replaced with isoleucine in the mutant form. The V149I mutant is twice the half-life of the wild-type at 60˚C. Interestingly, another mutant replacing arginine at position 249 with glutamic acid is decreased to only 26% compared to that of the wild-type. Thermodynamic and structural analysis revealed that the mutant V149I has higher "free energy of activation of thermal denaturation". The R249E had a totally a different mechanism of action. The arginine at position 249 forms 4 Hydrogen bonds with the C-terminal part of the protein. Replacement of this arginine moiety with glutamic acid, lead to the disruption of this interaction and could explain the lowered temperature resistance of the mutant [85].

Generally, bases of thermostability are not universal and differ from enzyme to enzyme and sometimes no apparent or striking change in sequence or structure accounts for thermal stability [86]. To sum up, simple amino acid changes and the corresponding
interactions (electrostatic and hydrophobic) as well as disulfide bridges and metal binding represent the molecular bases of thermal adaptation. While alterations in flexibility/ rigidity, quaternary configuration and compact packing represent the conformational bases of thermal resistance.

VI. Mercuric reductase from the ATII environment (MerA ATII-LCL)

In a previous work from our laboratory, a mercuric reductase from Atlantis II lower convective layer was characterized [22]. It is designated here for simplicity as ATII-LCL. This enzyme is activated by sodium chloride in a salt dependent manner. It reaches its maximum activity at 4 Molar NaCl. The enzyme is also thermostable as it retains 70% of its activity after 10 minutes of incubation at 70°C. This enzyme is 93% similar to a soil homolog that was used as control in the study. In contrast to ATII-LCL, the soil enzyme was less thermostable and is inactivated in the presence of salt. The ATII-LCL shows an abundance of acidic amino acid substitutions most notably on the surface of the molecule. It also has two distinguished short amino acid motifs of 4 and 7 amino acids named Box1 and Box2 respectively. In the vicinity of these Boxes comes a stretch of 4 aspartic acids (4D: Asp414-417), one of them, Asp 417, shows that it can potentially form a salt-bridge with the first lysine residue found in Box1. Interestingly, the substitution of Box1 with the corresponding amino acids found in the soil ortholog lowered the thermal stability. This could be due to the lost salt-bridge between lysine residue of Box1 and the aspartic acid 417. The replacement of Box2 reduced further the temperature resistance. The removal of both Boxes and replacing them with the corresponding residues in the soil homolog rendered the enzyme as thermo-labile as the soil enzyme (Table 1).

The substitution of these two Boxes did affect neither the kinetics parameters nor the salinity profile and thus were suggested to be directly involved in heat stability of the enzyme.
Table 1: Influence of Box1 and Box2 in the thermostability of MerA ATII-LCL

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Residual activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATII-LCL</td>
<td>80</td>
</tr>
<tr>
<td>Box1 mutant</td>
<td>60</td>
</tr>
<tr>
<td>Box2 mutant</td>
<td>30</td>
</tr>
<tr>
<td>Box1/Box2 mutant</td>
<td>20</td>
</tr>
</tbody>
</table>

1. The data is taken from Sayed et al. [22]
2. The residual activities of the enzymes were measured after incubation at 60°C for 10 minutes. The values (residual activity) are relative to those obtained after incubation at 25°C for 10 minutes.
3. Refer to mutants with residues that were substituted with their corresponding amino acids in the soil ortholog [22].

Both Boxes have proline residues, thus the reduction of thermal stability could have stemmed from the proline substitutions per se or the disruption of molecular interactions that stabilize the structure at high temperatures or both.

It is worth mentioning that Box1 lies very close to the dimerization domain that connects the two monomers together [87]. And Box2 forms an anti-parallel beta sheet structure and lies within the dimerization domain. Anti-parallel beta sheets are known to stabilize the protein structure [88]. This enzyme offers a model of a catalyst that adapts well to several harsh conditions characteristic of its native environment.

In this work, I would like to advance our understanding of how some critical amino acids of MerA ATII-LCL interact together—or with the surrounding amino acids—and cause the observed thermostability. To that end, we search an ATII MerA library for a sequence that is missing all the acidic amino acids and the two Boxes present in MerA ATII-LCL. This MerA sequence, from the same environment, should be ideal to introduce those critical amino acids sequentially in the place of their corresponding amino acids, analyze their interaction, and measure the thermostability of the recombinant MerA mutants. The work presented in this thesis characterizes a MerA recombinant enzyme (MerA ATII-LCL-NH) lacking all the acidic amino acids and the two Boxes present in MerA ATII-LCL. Three sets of amino acids sequences, Box1, Box2, and the 4 aspartic acids (4D:
Asp414/417) were used to generate different MerA ATII-LCL-NH mutants, and thermal stabilities of the mutants were calculated and the potential interactions of the substituted amino acids within the modeled 3D structure are presented.
Chapter 2: Experimental procedures

I. Sample collection and DNA extraction

Water samples were collected from the Atlantis II brine pool of the Red Sea at a depth of 2200 meters below sea surface. This was done in March/April 2010 during the expedition of the research vessel Aegaeo as part of the collaboration between KAUST and WHOI.

A rosette of Niskin bottles connected to a conductivity, temperature and depth device was lowered to the lower convective layer of the Atlantis II brine pool. Water samples from this layer were serially filtered through a series of 0.3 μm filter, 0.8 μm filter and on 0.1 μm cellulose-ester filter (Millipore) and stored in sucrose buffer. Collective DNA was extracted by (Mr. Amged Ouf) in our laboratory using a Marine DNA isolation kit (Epicenter) as described in details in [22].

II. Establishment of mercuric reductases library from Atlantis II LCL

a) Designing a pair of oligonucleotides for merA amplification

To design Primer for amplification we did multiple sequence alignment of agricultural soil MerA (GenBank™ accession number AEV57255.1), Tn501 MerA (GenBank™ accession number CAA77323.1) and consensus sequence of assembled reads (CSAR) using Multalin software [89]. The CSAR was obtained by screening an established metagenomics library from the lower convective layer of Atlantis II for mercuric reductase sequences.

Multiple sequence alignment of the three merA sequences (using Multalin software) show that the 5' and the 3' ends of the aligned merAs have identical nucleotide sequences. The designed pair of primers started with the ATG start codon of the forward primer and ended with the TGA stop codon of the reverse primer.

Two perfectly matching primers were synthesized accordingly;

Forward primer, MerA-F, 5' ATGACCCATCTAAAAATCACC GG CATGACTTTG 3' and reverse primer, MerA-R, 5' TCACCGGC CGCAGCAG GAAAGCTGCTTC 3'.
As for the MerA-F; length, 32, GC content, 44% and T$_m$ is 73°C. As for the MerA-R; length, 28, GC content 64% and T$_m$ is 80°C. The following sketch shows how the pair of oligonucleotides was designed. The MerA oligonucleotides used in the amplification process are shown in red letters on the top (MerA-F) and bottom (MerA-R) of the aligned sequences. The ATG start codon and the TGA termination codon are highlighted in green.

This pair of primers was used to amplify mercuric reductase genes from the DNA isolated from the Atlantis II LCL.

b) PCR amplification of Atlantis-II LCL mercuric reductases

Polymerase chain reaction was done using Phusion high-fidelity DNA polymerase (Thermo scientific, MA, USA) at a concentration of 1 Unit per 50 µl reaction as follows; 5X HF reaction buffer (containing MgCl$_2$ 7.5 mM) to a final concentration of 1X, primers 0.3 µM, dNTPs 0.3 mM, 100 ng DNA and DMSO 3%. These were the final optimized conditions for PCR amplification. If primers or dNTPs were used in other concentrations, this will result in multiple bands unlike the sharp single discreet band obtained from the optimized conditions.

30 cycles were done in a Veriti thermal cycler (Applied Biosystems, CA, USA) as follows; 30 seconds denaturing time at 95°C, 30 seconds annealing time at 70°C, 1 ½ minutes extension time at 72°C and then followed by a final extension step of 7 minutes at 72°C.
Polymerase chain reaction was performed using the ATII-LCL environmental DNA, MerA-F primer and MerA-R primer. A single discreet band of approximately 1.7 Kb, as expected from the merA gene length of 1686 base pairs, was obtained (see Supplemental Figure 1). The band was gel excised, purified using QIAquick gel extraction kit (Qiagen, MD, USA) and utilized to establish an ATII-LCL MerA library.

c) Library construction in pGEM vector

The band of interest was gel excised and purified using QIAquick PCR Purification Kit (Qiagen, MD, USA). The purified PCR fragments were ligated overnight at 15°C in pGEM vector (Promega, WI, USA) as recommended in manufacturer’s manual.

Plasmids were transformed into XL-1 Blue cells and plated on LB plates containing 100 µg/ml Ampicillin and incubated O/N at 37°C. 40 random colonies were picked and grown in 5 ml LB broth supplemented with 100 µg/ml ampicillin and shaken at 220 rpm overnight at 37°C. Glycerol stocks were made and stored at -80°C. Plasmids were extracted using PureYield™plasmid miniprep kit (Promega, WI, USA).

Sanger dideoxy sequencing form both ends was done using the T7 and SP6 pGEM universal primers (Macrogen Korea laboratory).

One of those genes, the ATII-LCL-NH, was chosen to be the template to mount all mutants present in this work. NCBI BLASTp [90] shows that its closest match is *Aquabacterium parvum* [91]. It is a gram-negative Betaproteobacteria.

This custom-made gene was synthesized at (Genscript, NJ, USA) in order to optimize codon usage and GC content and to increase expression levels in the host *Escherichia coli*. The synthesized gene was cloned into Champion™ pET SUMO (Invitrogen, CA, USA) and transformed into BL21 (DE3) cells for expression. This gene was used to generate all the mutants in this work. Therefore, it is safe to mention that the generated mutants’ genes are codon-optimized.

III. Generation of merA mutants using site-directed mutagenesis
The codon-optimized ATII-LCL-NH mercuric reductase gene was used as the template to generate the mutants described in this work. Mutagenic primers were generated using the web-based Quickchange primer design software available online at (www.agilent.com/genomics/qcpd). The Quickchange II site-directed mutagenesis kit (Agilent, CA, USA) was used to create various mutants addressed in this paper. Note that primers had to be used in excess as recommended in the instruction manual.

The cycling conditions started with an initial denaturation step of 30 seconds at 95°C. The Annealing temperature (T_a) of Box1 primer was calculated to be 77.5°C, Box2 77.2°C, 2D 75.4°C and 4D 75.1°C. Denaturation was done for 30 seconds at 95°C, annealing at the corresponding (T_a) for 30 seconds and extension for 7 minutes and 18 seconds at 68°C. 18 cycles were done as required for multiple amino acids substitutions.

Reaction tubes were put on ice immediately to bring down temperature to about 37°C. The restriction enzyme Dpn I was added to the reaction mixture at a concentration of 10 units per reaction and incubated for 1 hour at 37°C to digest the parent strand.

The mutant genes were cloned into pGEM vector (Promega, WI, USA) for propagation and further sequencing and pET SUMO vector (Invitrogen, CA, USA) for expression. The sequence of the primers used to generate each of the mutants is as follows;

4D-sense
5'-gaccaaccgcagtttgtctacgtggatgatggtaccgtgcagcaatcaacatg-3'

4D-antisense
5'-catgttgattgctgcacgggtaccatcatcatccacgtagacaactgcggttggtc-3'

2D-antisense
5'-ttgctgcacgggtaccagcatcatctgccacgtagacaactgcggttggtc-3'

2D-sense
5'-cagttttgtctacgtggcagatgctggtaccgtgcagcaatcaacatg-3'
**Box1-sense**

5'-caacatgacgggctgatgcagctaagccgccaggcaatgcggctgttggttttc-3'

**Box1-antisense**

5'-gaaaaccacagcccgcattgccctcgccggcttagctgcattcaccgcccgtcatgttg-3'

**Box2-antisense**

5'-gggcacgcgcacgattccagggaacttacacctggttgtcaatgccatcgtgatgcgc-3'

**Box2-sense**

5'-gcgcacgatgcattgaaaacaaggtcggtaagttccccctggataacgtgccgccc-3'

IV. Expression and purification of recombinant MerA isoforms

An overnight culture of transformed *E. coli* BL21 (DE3) cells harboring the recombinant plasmid (carrying merA isoform) was diluted 50-fold in fresh LB medium supplemented with 20 µM FAD and 50 µg/ml kanamycin.

Expression of recombinant enzymes was induced at an OD$_{600}$ of 0.6 by the addition of IPTG to a final concentration 0.2 mM overnight at 25°C. Cultures were centrifuged, re-suspended in lysis buffer (20 mM phosphate buffer, pH 7.4 containing 1 mM EDTA, 2 mM β-mercaptoethanol, supplemented with 50 µM FAD, protease inhibitor PMSF 1 mM, Triton X-100 1% v/v, imidazole 40 mM and lysozyme 0.2 mg/ml [34]) and incubated for 30 minutes at room temperature. Then the sample was sonicated three times for 30 seconds on ice using (Branson Sonifier 150) at output power 5 and cellular debris removed by centrifugation at 14,000 g for 45 min at 4°C. It is worth noting that higher concentration of IPTG used could result in the formation of inclusion bodies. Thus it is strongly recommended to use the indicated concentrations of IPTG to avoid this problem.

The clear supernatant was filtered on 0.2 µm syringe filter and applied to a pre-equilibrated 1 ml His-Trap HP column (GE Healthcare, IL, USA) at a flow rate of 1
ml/minute using the AKTA purifier instrument (UNICORN version 4.0, Pump P-950 and Frac-950). The wavelength detector was adjusted to 280 and 450 nm.

The column was washed with 15 column volumes (CV) of the binding buffer (20 mM sodium phosphate buffer, 40 mM Imidazole, pH 7.4). The protein was then eluted from the column using 500 mM of imidazole at a rate of 0.5 ml/min. Elution continued until no detectable reading was observed at 280 and 450 nm and fractions (250 µl) were collected.

SDS-PAGE was used to analyze the eluted fraction for the presence of the desired band on a 10% gel. The fractions containing the desired band were pooled.

The pooled yellow fractions (1.5 ml) were slowly passed on a pre-equilibrated 24-ml Superdex 75 10/30 gel filtration column (GE healthcare, IL, USA) at a flow rate of 0.5 ml/min to remove excess FAD and further purify the protein. Elution was done using 1X phosphate buffer saline (PBS) (KCl 2 mM, NaCl 137 mM, Na$_2$HPO$_4$ 10 mM, KH$_2$PO$_4$ 1.8 mM) until no detectable reading was observed at 280 and 450 nm.

Fractions, 500 µl, were collected and analyzed on 10% SDS-PAGE gel. Fractions containing the desired band were pooled (2 ml). The protein sample was concentrated on Amicon™ Ultra-4 10K centrifugal filters (Millipore, MA, USA) to a final concentration of 30 mg/ml and stored in aliquots at -80°C until further use.

Protein concentration was determined using Pierce™ BCA protein assay kit (Thermo Scientific, MA, USA). Protein purity was verified by electrophoresis on 10% SDS-polyacrylamide gels.

It is worth noting that induction at 37°C of mutants 4D, B1B2 and 4DB1B2) results in the formation of inclusion bodies. Therefore, these mutants were induced at 24°C overnight. This will be elaborated more in the results and discussion section.

V. Mercuric reductase activity assay
Routine enzyme assays were done in triplicates at 25°C in the assay mixture (phosphate buffer 80 mM, β-mercaptoethanol 1 mM, EDTA 1mM, HgCl₂ 50 μM). Enzymes were used at 100 μg in a final volume of 1 ml. The reaction was started by the addition of 100 μM NADPH.

The enzyme activity was monitored by observing the initial rate of substrate-dependent NADPH consumption at 340 nm using a Shimadzu UV-1800 spectrophotometer. The unit of activity is defined as the amount of enzyme that catalyzes Hg⁺²-dependent oxidation of 1 μmol of NADPH/min [92]. Kinetic and statistical parameters were calculated using the Graphpad prism software (Confidence Interval = 95%). All samples were replicated at least three times and the data including the mean and standard deviation were entered into the Graphpad software. The program plots the curves and shows the mean and standard deviation. Assaying the ATII-LCL-NH enzyme and other mutants in this work under different substrate mercury concentrations was done to calculate the affinity (Kₘ) and Vₘₐₓ parameters. Triplicates of each point were entered into the Graphpad Prism software including the mean and standard deviation. The Software draws the double reciprocal plot, from which the Kₘ and Vₘₐₓ are deduced. The molecular weight of the monomer was used in the Graphpad Prism software to calculate the Kₐₜ.

VI. Heat Stability Analysis
Replicate samples of each enzyme were incubated at the designated temperatures for 10 minutes in the Veriti Thermal cycler (Applied Biosystems, CA, USA) and then dipped immediately on ice. The vials were spinned down to separate any precipitated protein and clear supernatant taken. The residual activity was then measured by adding the heated enzyme to the reaction mixture as previously described. Results are expressed relative to the activity observed after incubation for 10 minutes at 25°C [22].

VII. Salt tolerance Analysis
Replicate samples of each enzyme were added to the standard reaction mixture containing the indicated concentrations of NaCl. The reaction was started by the addition of 100 μM NADPH. The initial rate of NADPH oxidation was measured. The
percentage of inactivation by NaCl was calculated relative to the enzyme activity in the absence of NaCl. The enzyme activity in the absence of NaCl is taken as the 100%.

VIII. Modeling of the Three-dimensional Structure of the wild-type ATII-LCL-NH
The amino acid sequences of the ATII-LCL-NH and other mutants were used to obtain the protein data bank (PDB) file using the free online software SWISS-MODEL.

The 3D structure of the ATII-LCL-NH MerA was built by homology modeling against structures of the N-terminal domain of Tn501 mercuric reductase (Protein Data Bank code 2kt2) [93]. And the C-terminal core (amino acids 95-561)(Protein Data Bank code 1zk7) [46] domain of the Tn501 mercuric reductase, using SWISS-MODEL [94]. Manual inspection of the output was done in PyMOL (PyMOL Molecular Graphics System, Version 1.5.0.1 Schrödinger, LLC). Rendering of the final three-dimensional models were performed using PyMOL [95]. The protein data bank file for the ATII-LCL-NH and all the mutants was further used to analyze potential electrostatic interactions.

IX. Salt Bridge prediction
The Protein Data Bank file of the ATII-LCL-NH and all the mutants was used as the input file to predict the putative salt bridges between oppositely charged residues in the entire dimer. A cutoff of 4 Angstroms was used. Potential Salt bridges were predicted using the online software ESBRI [71, 96, 97] available at: http://bioinformatica.isa.cnr.it/ESBRI/input.html.

X. Hydrogen Bond prediction
The protein databank files of the ATII-LCL-NH and all other mutants in this work were used as input to predict Hydrogen bonding. Potential Hydrogen bonds were predicted using the online software tool available at: (http://cib.cf.ocha.ac.jp/bitool/HBOND/) [98, 99], using the default parameters.

XI. Venn Diagram plotting
The Venn Diagrams in this work were drawn using the VENNY online software version 2.1 available at (http://bioinfogp.cnb.csic.es/tools/venny/) [100]. The spreadsheet
generated from the PDB file of each mutant of either salt bridges or hydrogen bonds was used as input to create the Venn diagrams.
Chapter 3: Results and discussion

I. Establishment of mercuric reductases library from Atlantis II LCL
   a) Designing a pair of oligonucleotides for merA amplification

We designed a pair of oligonucleotide primers to amplify mercuric reductases from the Atlantis II LCL DNA. The sequences of agricultural soil merA (GenBank™ accession number AEV57255.1), Tn501 merA (GenBank™ accession number CAA77323.1) and merA consensus sequence of assembled reads (CSAR) from the Atlantis II dataset were used to generate oligonucleotide primers for the PCR amplification process. Multiple sequence alignment of the three merA sequences (using Multalin software) show that the 5’ and the 3’ ends of the aligned merAs have identical nucleotide sequences. The designed pair of primers started with the ATG start codon of the forward primer and ended with the TGA stop codon of the reverse primer as shown in the sketch in the experimental procedures section).

b) PCR amplification of Atlantis II LCL mercuric reductases

Polymerase chain reaction was performed using the ATII-LCL environmental DNA, MerA-F and MerA-R as shown in the sketch in the experimental procedures section. A single discreet band of approximately 1.7 Kb, as expected from the merA gene length of 1686 base pairs, was obtained (Supplemental Figure 1).

c) Library construction in pGEM vector

The purified PCR DNA fragment was cloned into pGEM vector and the library was composed of forty clones that were randomly selected. The recombinant plasmids were isolated for further analyses of the inserted DNA fragments and subsequent DNA sequencing. PCR was carried out on ten recombinant plasmids using the merA-F and merA-R primers to verify the size of the inserted DNA fragments as shown in Supplemental Figure 2B. All the analyzed recombinant plasmids (shown in Supplemental Figure 2A) have inserts of 1.7 Kb corresponding to the expected size, 1686 base pair, that can potentially code for full length MerA of 561 amino acid residues (Supplemental Figure 1).
The Forty extracted plasmids were then sequenced from both ends at the Macrogen laboratory in Korea using the T7 and SP6 universal primers. The nucleotide sequences were of high quality and all bases have Phred value of sixteen or higher. The forward and reverse sequences from each reaction were assembled and analyzed.

II. Multiple sequence alignment of non-redundant and full-length MerA sequences

The sequencing of the inserted DNAs of the forty isolated recombinant plasmids from the merA library resulted in eight full-length non-redundant mercuric reductase sequences (Figure 4). The sequences have very few amino acids differences between them, varied between one and two. The sequence designated ATII-LCL-NH has a high similarity to the well-characterized mercuric reductase Tn501 with only one amino acid different at position 386, an alanine residue instead of a valine in Tn501. This sequence was also obtained before from a separate library. All of the eight non-redundant full-length sequences ended with glycine except 2A11 and 2A4 genes. This could have been a sequencing error, although each gene was sequenced more than once and the chromatogram showed discrete high-quality peaks.

All eight sequences have the cysteines 11/14 in the N-terminal metal binding domain, catalytic cysteine residues 136/141 and mercury binding cysteine residues 558/559.
Figure 4: Multiple sequence alignment of non-redundant and full-length MerA sequences.

Multiple sequence alignment of eight non-redundant full-length MerA sequences obtained from the ATII-LCL MerA library. The first sequence of the figure, ATII-LCL-NH, is a MerA sequence identified also in a separate library. The amino acid variants/substitutions are in red letters. The NmerA domain [93] is underlined in green. The Dimerization domain [87] is underlined in purple.
Very few variations were observed between the above sequences. The observed few substitutions were not in the redox-active or mercury-binding cysteine residues. Therefore, one of them was chosen as the representative sequence. The ATII-LCL-NH was the selected sequence to be the backbone for mounting all the mutants in this work.

III. Generation of MerA mutants using site-directed-mutagenesis

As mentioned earlier, eight full-length metagenome ATII-LCL MerA were obtained and one of them, ATII-LCL-NH, was found to be different just by one amino acid residue from the well characterized Tn501 MerA. The ATII-LCL-NH MerA sequence is missing all the acidic amino acids, including the two Boxes responsible for the thermostability of the MerA ATII-LCL [22] (Figure 6). The ATII-LCL-NH sequence therefore was selected to introduce sequences from the ATII-LCL MerA that have shown to affect, or have the potential to affect, the thermostability of the protein.

a) Optimization of codon usage of the ATII-LCL-NH gene sequence

Since ATII-LCL-NH was selected to generate recombinant mutants with the desired mutations, its nucleotide sequences was analyzed to optimize the codon usage for efficient expression in E. coli.

The ATII-LCL-NH gene sequence has a high GC content (64%). This is expected from a gene isolated from a deep hydrothermal environment [101]. In order to enhance the heterologous expression of this protein in E. coli, codon usage and the GC/AT content of the coding sequence were optimized for expression at the Genscript laboratory (NJ, USA) (Figure 5 and Table 2). The optimized gene has 57% GC content.
Table 2: GC content of the wild-type and optimized ATII-LCL-NH gene.

<table>
<thead>
<tr>
<th></th>
<th>Optimized</th>
<th>Wild-type</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>57.1%</td>
<td>63.9%</td>
</tr>
</tbody>
</table>

Figure 5: Optimization of codon usage of the ATII-LCL-NH gene.
The nucleotides sequence of ATII-LCL-NH was optimized at Genscript (NJ, USA) for heterologous expression into *E. coli*. The optimized and unoptimized bases are in blue and red colors respectively.

b) Amino acid differences between MerA ATII-LCL and ATII-LCL-NH

A multiple sequence alignment of the merA ATII-LCL and the ATII-LCL-NH was done using Multalin software [89] (Figure 6). The amino acid differences between both homologs are highlighted in red, which are mainly acidic amino acids and Box1 and Box2.

In a previous work [22], the halophilic and heat stability of MerA ATII-LCL were thoroughly studied. Site-directed-mutagenesis of the MerA ATII-LCL template revealed
the implication of two motifs (Box1) and (Box2) in the thermal inactivation of the enzyme. It was also noticed that the two Boxes are in close proximity with a short stretch of four aspartic acids (4D) located at position 414-417. Analysis of the 3-dimensional structure of the ATII-LCL MerA shows that lysine 432 in Box1 forms a salt bridge with aspartic acid 417 located in the stretch of the 4D residues. This observation suggested that the 4D stretch may be implicated in the thermostability.

Figure 6: Pairwise alignment of MerAs ATII-LCL and ATII-LCL-NH.
The ATII-LCL-NH is the upper sequence in grey. The ATII-LCL is the lower sequence in black. The amino acids different in ATII-LCL compared with ATII-LCL-NH are in red colored letters. The NmerA domain is over lined in green. The Dimerization domain is over lined in purple. The cysteine pairs 11/14 and 558/559 which are responsible for Hg$^{2+}$ binding, and the cysteine pair 136/141 involved in the catalytic site are highlighted in yellow. The positions of the amino acids that are involved in the mutations performed in this work are in black Boxes.
c) Mutants generated in this work

Six mutants were generated by site-directed-mutagenesis in this work, all involved the four aspartic acids at position 414 to 417 and the two Boxes shown in Figure 6. The substituted amino acid of each mutants and its corresponding residue in ATII-LCL-NH are shown in Table 3. To help comparing the six mutants generated in this work, a sketch showing the alteration performed in each mutant; (Figure 7).

Table 3 Mutations to substitute residues from the ATII-LCL-NH to their corresponding amino acids in the ATII-LCL enzyme.

<table>
<thead>
<tr>
<th>Mutation*</th>
<th>From</th>
<th>To</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>A415-A416 ($^{415}$AA$^{416}$)</td>
<td>D415-416 ($^{415}$DD$^{416}$)</td>
</tr>
<tr>
<td>4D</td>
<td>A414-417 ($^{414}$AAAA$^{417}$)</td>
<td>D414-417 ($^{414}$DDDD$^{417}$)</td>
</tr>
<tr>
<td>B2</td>
<td>DSRTLT$^{465}$-470 ($^{465}$DSRTLT$^{470}$)</td>
<td>KVGKFP$^{465}$-470 ($^{465}$KVGKFP$^{470}$)</td>
</tr>
<tr>
<td>B1B2</td>
<td>$^{432}$LDLT$^{435}$/$^{465}$DSRTLT$^{470}$</td>
<td>$^{432}$KPAR$^{435}$/$^{465}$KVGKFP$^{470}$</td>
</tr>
<tr>
<td>2DB1B2</td>
<td>$^{415}$AA$^{416}$/432$^{33}$LDLT$^{435}$/465$^{65}$DSRTLT$^{470}$</td>
<td>$^{432}$KPAR$^{435}$/465KVGKFP$^{470}$</td>
</tr>
<tr>
<td>4DB1B2</td>
<td>$^{414}$AAAA$^{417}$/432$^{33}$LDLT$^{435}$/465$^{65}$DSRTLT$^{470}$</td>
<td>$^{432}$DDDD$^{417}$/465$^{35}$KPAR$^{465}$KVGKFP$^{470}$</td>
</tr>
</tbody>
</table>

*Refer to Figure 7, and its legend, for the mutation names.

![Figure 7: Illustration presenting the mutations shown in Table 3.](image)

Yellow spheres are the cysteine residues involved in binding and reduction of Hg II; red spheres are aspartic acid residues; Blue Boxes are Box1 and Box2. For simplicity, the mutants ATII-LCL-NH/2D, ATII-LCL-NH/4D, ATII-LCLNH/B2,
ATII-LCL-NH/2DB1B2, and ATII-LCL-NH/4DB1B2 will be referred to in the text as 2D, 4D, B2, 2DB1B2 and 4DB1B2 respectively.

IV. Expression and Purification of ATII-LCL-NH and its mutants

a) Optimization of the induction process of the recombinant MerA ATII-LCL-NH and its mutants

ATII-LCL-NH gene was expressed in *E. coli* BL21 cells under the control of T7 promoter in pET-SUMO expression vector at 37°C. Similar levels of MerA protein yield were obtained by induction of the recombinant protein at different IPTG concentrations, 0.1, 0.2, or 0.5 mM IPTG (Supplemental Figure 3). Moreover, the maximum level of expression of the recombinant protein did not change with the timing of induction.

The protein was expressed in the soluble cellular fraction of the cell lysate. However, some of the mutant enzymes (4D, 4DB1B2 and B1B2) were only expressed in the soluble fraction after overnight induction with 0.1 or 0.2 mM IPTG at 25°C. Their expression for two hours at 37°C resulted in the formation of inclusion bodies. This could be explained in the context of their thermal stability behavior. Those mutants precipitate and become inactivated by heat as will be shown later. The band observed on the 10% SDS-PAGE gel was of the expected size (Supplemental Figure 3).

b) HisTag affinity column and superdex 75 size exclusion chromatography purification of MerA ATII-LCL-NH recombinant protein

The purification procedure described in this section was applied successfully with MerA ATII-LCL-NH and its mutants. Two columns chromatography were used in the purification process, HisTag affinity chromatography and size-exclusion on Superdex 75. The elution profiles obtained from the two columns chromatography were similar for MerA ATII-LCL-NH and its mutants. Therefore, the results of purification of just MerA ATII-LCL-NH are presented. Also, this second step of the purification process on Superdex 75 column resulted in superior yield and purity relative to dialysis.

1. *HisTrap affinity purification*
The cell lysate of the ATII-LCL-NH E. coli clone was uploaded into 1 ml HisTag affinity column equilibrated with 40 mM Imidazole of binding buffer. This relatively high concentration of imidazole was used to enhance the purity of the enzymes and reduce the contamination with cellular proteins. The expressed enzyme has a Poly-Histidine tag and thus binds to the nickel resin of the HisTrap column and the remaining cellular protein pass in the flow-through. Step elution was done using 500 mM Imidazole. The entire purification process was done on the AKTA purifier instrument. The eluted yellow fractions – due to the bound FAD – containing the purified enzyme were pooled as (Supplemental Figure 4).

Aliquots from the HisTrap eluted fractions were analyzed on 10% SDS-PGE to assess the purification process. The MerA recombinant enzyme was found to be the major protein of the cell lysate. The flow-through of the HisTrap column show bacterial cellular proteins that did not bind the HisTrap column. The eluted MerA fractions show a major band of MerA (Supplemental Figure 5).

Different eluted yellow fractions from the HisTrap column were loaded on 10% SDS-PAGE for visualization. The soluble part of the cell lysate shows a remarkable band, while the flow-through show bacterial cellular proteins passing through the column without binding the nickel resin. Eluted yellow fractions show high levels of protein concentration and some light bands that could be lyzed fragments or co-purified cellular proteins. Fractions (yellow) containing the mercuric reductase protein were pooled for the next purification step. Fractions 5 and 6 show the highest purity, while fractions 3 and 4 show the highest yield (Supplemental Figure 5).

2. Superdex size-exclusion chromatography
The MerA ATII-LCL-NH pooled fractions from the HisTrap column (Supplemental Figures 4 and 5) were passed through 24 ml Superdex 75 size-exclusion (gel-filtration) chromatographic column (Supplemental Figure 6). Two major peaks were observed, the first eluted at 6.0 ml containing MerA protein while the second at around 14 ml has the unbound FAD.
Analysis of the first peak on 10% SDS-PAGE (Supplemental Figure 7) shows a MerA single band indicating a highly purified protein. Pooled fractions were concentrated on Amicon Ultra-4 centrifugal filter units with a cut off of 10 KDa to a final concentration of 30 mg/ml in PBS buffer. All enzymatic assays were done immediately after the purification process.

It is important to note that the six mutants generated in this work (Table 2 and Figure 7) were successfully purified in a comparable manner as ATII-LCL-NH. In addition, all the mutants show a single protein band on 10% SDS-PAGE as shown in Supplemental Figure 7 for ATII-LCL-NH.

V. Effect of NaCl on the activity and kinetic parameters of ATII-LCL-NH and its mutants

a) Inhibition by NaCl

As mentioned earlier, previous work with MerA ATII-LCL revealed that the enzyme has distinguished properties, it is activated by NaCl in a salt dependent manner and attains its highest activity in the presence of 4 M sodium chloride [22]. This high molarity of sodium chloride is similar to the salinity concentrations of the ATII LCL environment. This enzyme is suggested to originate from an organism adapted to the salt-in approach.

In contrast to the ATII-LCL enzyme, the ATII-LCL-NH homolog described in this work did not show an increase in acidic and polar uncharged side-chains at the expense of hydrophobic amino acids, a characteristic of non-halophilic proteins [22, 49, 50, 53, 54]. To examine the response of the ATII-LCL-NH and its mutants to NaCl, we assayed the enzyme’s activity in the presence of increasing concentrations of NaCl. Indeed, and as expected from their amino acids composition, the ATII-LCL-NH and its mutants are strongly inhibited by NaCl (Figure 8B). Although all the mutants were found to be inhibited by NaCl, a clear difference in the degree of inhibition was noticed between the different mutants. The effect of the stretch of the 4D residues (414DDDD417) on the response of the mutants to inhibition by NaCl is presented in Figure 8B. The 2D (415DD416) did not change the salinity response of the mutants toward NaCl when
compared with ATII-LCL-NH. However, the 4D enzyme resulted in a mutant which is strongly inhibited by NaCl (Figure 8B).

![Image A](image.png)

![Image B](image.png)

**Figure 8: Inhibition of ATII-LCL-NH and its mutants by NaCl**
To facilitate the comparisons between the ATII-LCL-NH and its mutants the cartoon from Figure 7 is presented in (A). B) The enzymes were assayed in a standard assay mixture in the presence of the indicated concentration of NaCl. Each point is replicated 3 times and mean/standard deviation is plotted using Graphpad Prism version 7.0 for Windows.
Mutants with the two Boxes also show different inhibition with NaCl when compared with the ATII-LCL-NH. The B1B2 mutant was found to be less inhibited by NaCl compared with ATII-LC-NH. However, mutants of the B1B2 with the 2D (2DB1B2) or the 4D (4DB1B2) were strongly inhibited by NaCl similar to the 4D mutant. Although the 2D (415DD416) and Box1 and 2 (B1B2) have no noticeable difference in the degree of inhibition by NaCl compared with the ATII-LCL-NH, mutants containing the 2D together with Boxes B1 and B2 resulted in enzymes that are strongly inhibited by NaCl. The residual activates of ATII-LCL-NH and its mutants at 0.5M NaCl is summarized in Table 4.

**Table 4: Residual activities of ATII-LCL-NH and its mutants at 0.5 M NaCl**

<table>
<thead>
<tr>
<th>MerA</th>
<th>Residual activity at 0.5 M NaCl (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATII-LCL-NH</td>
<td>55.0</td>
</tr>
<tr>
<td>B1B2</td>
<td>74.0</td>
</tr>
<tr>
<td>2D</td>
<td>56.0</td>
</tr>
<tr>
<td>4D</td>
<td>27.0</td>
</tr>
<tr>
<td>2DB1B2</td>
<td>38.0</td>
</tr>
<tr>
<td>4DB1B2</td>
<td>26.0</td>
</tr>
</tbody>
</table>

The results show that both the 4D and 4DB1B2 behave nearly in an identical fashion. So, the reduction in salt-tolerance could result from the addition of the 4D residues rather than B1B2 motifs. In addition, the 2D alone or B1B2 alone are relatively insensitive to inhibition by NaCl when compared with the 2DB1B2. These results suggest that individual mutations could not explain the ability of the enzyme to tolerate salt on their own. Thus, it seems that the molecule behaves as a unit and requires favorable conformation and interactions with the solvent around it to impart the stability in the presence of salt. The results also reveal the complexity of interpretation of the behavior of these mutants toward their response to increasing concentration of NaCl.

b) Kinetic parameters of the ATII-LCL-NH and its mutants

The kinetic parameters of the NaCl inhibited ATII-LCL-NH, 2D and 4D mutants were determined. We selected the 2D and 4D mutants for this comparison because they
show different levels of inhibition by NaCl when compared with ATII-LCL-NH (Figure 8B).

The specific activities, the affinity to the substrate (Hg), and the efficiency – measured as $K_{\text{cat}}/K_m$ – are very similar for the ATII-LCL-NH, 2D and 4D enzymes (Table 5). This result indicates that the different degree of NaCl inhibition (Figure 8B) of the three MerAs is not due to changes of the affinity of the enzyme to the substrate, or the efficiency of the catalytic process.

Table 5 Kinetics data of ATII-LCL-NH, 2D and 4D mutants*

<table>
<thead>
<tr>
<th>MerA Isoform</th>
<th>Specific activity (µmol/min/mg)</th>
<th>$K_m$ (µM)</th>
<th>$V_{\text{max}}$ (µmol/min/mg)</th>
<th>$K_{\text{cat}}$ (s$^{-1}$)</th>
<th>$K_{\text{cat}}/K_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATII-LCL-NH</td>
<td>10.96</td>
<td>11.25</td>
<td>14.21</td>
<td>12.25</td>
<td>1.1</td>
</tr>
<tr>
<td>2D</td>
<td>10.95</td>
<td>10.96</td>
<td>14.05</td>
<td>12</td>
<td>1.1</td>
</tr>
<tr>
<td>4D</td>
<td>11.10</td>
<td>13.90</td>
<td>15.09</td>
<td>12.6</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*K_m and $V_{\text{max}}$ were determined using Lineweaver–Burk (double reciprocal plot) and the Graphpad Prism software.

VI. Thermal stability of ATII-LCL-NH merA and its mutants

In the previous work, it was shown that Box1 and Box2 contribute to the thermostability of MerA ATII-LCL. As shown in Table 1, substitution of Box1 (KPAR), Box2 (KVGKFP), or double mutants of Box1 and Box2, with residues found in the soil homolog, resulted in mutants that retained 60%, 30%, and 20% of their activities respectively [22].

It is worth noting that Box1 and Box2 each have one proline residue. Proline is known to reduce the structural flexibility of polypeptide regions containing it [77, 78, 81]. Additionally, both Boxes have basic residues - Box1 has one lysine and one arginine, and Box2 has a lysine (Table 3). In fact, modeling the three-dimensional structure of the MerA ATII-LCL shows the presence of a salt bridge between lysine 432 in Box1 (a random coiled loop) and aspartic 417 located in the stretch of the 4D residues (aspartic 414 to 417) (Figure 9). The presence of this salt bridge between lysine 432 in Box1 and aspartic 417 within the 4D stretch suggests a possible involvement of this interaction in the thermal stability of ATII-LCL.
Figure 9: Salt bridge between Lys 432 of Box1 and Asp 417 in the 4D. 
A) Shows a MerA monomer in which Box1 (random coiled loop), Box2, (beta sheet) and the 4D are shown in purple. B) Enlargement of the area containing Box1 and the 4D residues. C) The two amino acids residues involved in the potential salt bridge.

Interestingly, although ATII-LCL-NH is missing the two Boxes and all the acidic residues present in its ATII-LCL homolog (Figure 6), it still retains 81% residual activity at 60°C (Table 6). However, the ATII-LCL-NH is significantly less stable at 70°C in comparison with ATII-LCL, as it retains 44% of its activity relative to 70% residual activity in case of ATII-LCL.

Table 6: Thermostability of ATII-LCL-NH

<table>
<thead>
<tr>
<th>MerA isoform</th>
<th>Residual activity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>at 60°C</td>
</tr>
<tr>
<td>ATII-LCL-NH</td>
<td>81</td>
</tr>
</tbody>
</table>

The residual activities of the enzyme were measured after incubation for 10 minutes at the temperature indicated in the table. The values (residual activity) are relative to those obtained after incubation at 25°C for 10 minutes.
To examine if the two Boxes and the 4D residues can increase the thermostability of ATII-LCL-NH, the established set of mutants in which the two Boxes (Box1 and Box2), and the 4D were introduced to substitute their corresponding residues in the ATII-LCL-NH were examined for their thermostabilities (Figure 10A).

ATII-LCL-NH and its mutants were expressed in *E. coli* and their proteins were purified and examined for their thermostabilities. The results of the thermostability of the ATII-LCL-NH and its mutants are presented in Figure 10B.
Figure 10: Thermostability of ATII-LCL-NH and its mutants.
A) A sketch of the ATII-LCL-NH and its mutants indicating the location of the residues mutated in each mutant (indicated in the figures) and their relative position to the cysteines involved in the binding and catalysis, and the NmerA and the catalytic domains. B) The enzymes were incubated at the indicated temperatures for 10 minutes and the residual activities were measured in standard assay conditions (refer to experimental procedures section for details).

The B1B2 mutant was unstable to heat treatment when compared to ATII-LCL-NH, whereas the B2 mutant was much more stable than the B1B2. These results show that
the two Boxes, alone, did not increase the thermostability of ATII-LCL-NH, in fact they cause instability of the molecule towards heat. This may implicate the requirement of the 4D motif together with the Boxes to form the salt bridge between lysine 432 in Box1 and aspartic 417–located in the 4D stretch–to have a mutant that possess high stability to heat treatment. Surprisingly, in comparison with ATII-LCL-NH, the 2D (aspartic 415 and 416) alone increases the thermostability of the molecule whereas the 4D severely caused instability of the mutant toward heat treatment (Figure 10B). The 4DB1B2 mutant, however, was severely unstable to heat as the B1B2 mutant.

Modeling the three-dimensional structure of the 4DB1B2 shows that the salt bridge between lysine 432 in Box1 and aspartic 417 of the 4D was indeed established. This indicates that this salt bridge did not change or alter the negative behavior of the B1B2 mutant molecule toward stability to heat. Introducing the 2D to the B1B2 mutant, however, repels partially the negative effect of the two Boxes toward stability to heat - the 2DB1B2 mutant was found to have intermediate heat stability between the B1B2 and the 2D mutants. The temperature that is required for 50% thermal stability of the ATII-LCL-NH and its mutants is presented in Table 7.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATII-LCL-NH</td>
<td>68.0</td>
</tr>
<tr>
<td>B2</td>
<td>60.0</td>
</tr>
<tr>
<td>B1B2</td>
<td>33.0</td>
</tr>
<tr>
<td>2D</td>
<td>74.0</td>
</tr>
<tr>
<td>4D</td>
<td>33.2</td>
</tr>
<tr>
<td>2DB1B2</td>
<td>53.5</td>
</tr>
<tr>
<td>4DB1B2</td>
<td>33.8</td>
</tr>
</tbody>
</table>

1The residual activities of the enzymes were measured after incubation for 10 minutes
2The data is taken from Figure 10B

The results show that removal of the two Boxes from ATII-LCL–which caused instability of the MerA to heat–did not cause regain of the heat stability when introduced to ATII-LCL-NH which lacks these two motifs. The ATII-LCL, in additional to the two Boxes and the 4D, has other additional differences with ATII-LCL-NH particularly at least 32 acidic residues.
residues distributed along the polypeptide chain (Figure 6). These results therefore indicate that stability to heat does not require particular amino acid residues and motifs, but rather an overall adjustment of the secondary/tertiary/quaternary structure of the molecule to cope with heat.

We therefore analyzed computationally the hydrogen and ionic bonds present in the ATII-LCL-NH dimer and compared it with the 2D, 4D, and 4DB1B2 mutants. The rationale of this is to compare the effect of 2D on increasing the stability, 4D on its destabilization of the molecule to heat, and effect of the two Boxes together with the 4D. The hydrogen bonds between the amino acid side chains, within the MerA dimer molecules, were computed using the Hydrogen Bond calculator version 1.1 web server available at (http://cib.cf.ocha.ac.jp/bitool/HBOND/) [98, 99]. Figure 11A shows the total hydrogen bonds in ATII-LCL-NH and the three mutants analyzed.

Figure 11: Total hydrogen bonds of ATII-LCL-NH and its mutants.
A) Total hydrogen bonds in the indicated MerAs in the figure were calculated using Hydrogen Bond Calculator version 1.1 [99]. B) From the results in Figure 11A, the loss and gain of hydrogen bonds in each MerA indicated in the figure were computationally calculated.
There is no direct relation between the total hydrogen bonds in the molecules and the observed stabilities of the mutants compared to the ATII-LCL-NH. A clear difference that can be correlated with the heat stabilities was observed by comparing the loss and gain of hydrogen bonds between the ATII-LCL-NH and its mutants (Figure 11B). The 2D mutant lost 26 hydrogen bonds present in ATII-LCL-NH, and gained 42 new hydrogen bonds. Whereas the 4D mutant—that is unstable to heat—lost 69 hydrogen bonds and gained 90 new hydrogen bonds. In the case of 4DB1B2 mutant, the loss and gain of hydrogen bonds were very prominent - 174 bonds were lost and the molecule has 150 newly established hydrogen bonds.

Analyzing the Venn diagram of these results (Figure 12), shows that the 45 hydrogen bonds that are present in the ATII-LCL-NH and the 2D mutant, are missing in the 4D heat-labile molecule. In addition, out of the 42 newly hydrogen bonds formed in the 2D mutant, 28 were shared with 4D heat-labile mutant and 14 were found to be specific for the 2D heat stable mutant. Moreover, 62 new hydrogen bonds were specific to the 4D mutant. These results nominate the 45 hydrogen bonds lost from the 4D mutants, and the 62 new formed hydrogen bonds in the 4D as potential candidates for destabilization of the molecule, turning it into heat-labile structure. In support of this potential proposition, the Venn diagram of the ATII-LCL-NH, 2D, 4D, and 4DB1B2 shows that 52 hydrogen bonds out of the 62 hydrogen bonds are still present in the 4DB1B2 thermo-labile mutant. Moreover, the 45 hydrogen bonds that are shared between ATII-LCL-NH and the 2D are absent in the 4DB1B2 mutant as well as the 4D mutant.
Figure 12: Venn diagram of hydrogen bonding in ATII-LCL-NH and its mutants.

Venn diagram of total hydrogen bonds in ATII-LCL-NH, 2D, and 4D mutants, was created and drawn using VENNY software version 2.1 [100], is shown in (A), and ATII-LCL-NH together with 2D, 4D, and 4DB1B2 mutants is shown in (B).

The salt bridges in the ATII-LCL-NH and its mutants were also analyzed using the web server for evaluating salt bridges in proteins (http://bioinformatica.isa.cnr.it/ESBRI/introduction.html) [71]. In comparison with ATII-
LCL-NH, the introduction of the 2D or the 4D has increased the salt bridges by 6 and 7 bridges respectively (Figure 13).

![Diagram](image)

**Figure 13: Total salt bridges of ATII-LCL-NH and its mutants.**
A) Total salt bridges in the indicated MerAs in the figure were calculated using ESBRI software [71]. B) From the results in Figure 13A, the lost and gained potential hydrogen bonds in each MerA were computationally calculated.

Note that the 6 new salt bridges in the 2D mutant are also shared with the 4D mutant. Since the introduction of the 2D residues did not cause loss of the thermostability of the molecule, those 6 salt bridges, most probably, did not contribute to the dramatic decreases of the thermostability in 4D mutant. The only difference that can potentially explain the extreme sensitivity of the 4D to heat is the only one salt bridge that is specific for this mutant. It is known that salt bridge formation or disruption in site-
directed mutagenesis studies may improve or lower thermal stability depending on the location of the bridge and its geometry [70, 71].

In the case of the 4DB1B2 mutant, although did not gain this unique salt bridge, it has gained 16 salt bridges that are not present in the thermostable ATII-LCL-NH molecule and the 2D mutant (Figure 14).

![Venn diagram](image-url)

**Figure 14: Venn diagram of salt bridges in ATII-LCL-NH and its mutants.** Venn diagram of total potential salt bridges in ATII-LCL-NH, 2D, and 4D mutants was created and drawn using VENNY online software [100], is shown in (A). Diagram of ATII-LCL-NH together with 2D, 4D, and 4DB1B2 mutants is shown in (B).
Taken together, the potential hydrogen bonds and salt bridges that may have contributed to the instability of the 4D mutant to heat are: loss of 45 and gain of new 62 hydrogen bonds, and gain of 1 salt bridge in the 4D mutant.
Chapter 4: Conclusion

The well-characterized MerA ATII-LCL from ATII LCL environment is thermostable and is activated by NaCl. These features suggest that this isoform of MerA may have evolved in a microorganism that utilizes the salt-in approach. In this work, an ortholog from the same environment is characterized using reverse-genetics and site directed mutagenesis approaches. This ortholog, ATII-LCL-NH, is thermostable and is inhibited by NaCl suggesting that, unlike ATII-LCL, it may have evolved in a microorganism that use the salt-out mechanism. Although it is thermostable—as expected from proteins that function in microorganism that reside in this environment—the enzyme was found to be less stable at 70°C when compared with its ortholog ATII-LCL. In contrast with ATII-LCL, ATII-LCL-NH devoid of two boxes, that were shown to be involved in the thermostability, and all the acidic residues present in ATII-LCL. The fact that ATII-LCL-NH is also devoid of the acidic residues that were shown to contribute to the halophilicity of ATII-LCL explains the inhibition properties observed in this ortholog. The ATII-LCL boxes did not increase the thermostability of ATII-LCL-NH mutants, on the contrary; they strongly destabilized the molecule towards heat treatment. Moreover, mutants containing the ATII-LCL boxes together with the acidic residues from the 4D did not restore the thermostability of the molecule despite that one mutant has established the salt bridge between Box1 and Aspartic 417 in the 4D.

Interestingly, only the 2D mutant was more tolerant to high temperature relative to both the ATII-LCL and the ATII-LCL-NH. It retains 81% of its activity after 10-minute incubation at 70°C.

Analyzing hydrogen bonding and salt bridge interactions of the ATII-LCL-NH and its mutants suggest that a unique set of newly formed Hydrogen bonds and salt-bridge network could contribute to the increase of the heat stability of the 2D mutant, and the instability of the B1B2 alone or together with different set of residues from the 4D. Most probably, a subtle wide-spread conformational change could have imparted the enhanced thermal stability of the 2D mutant and instability of thermo-labile ATII-LCL-NH mutants. It is worth noting that ATII-LCL-NH and its mutants have their kinetics
properties unaltered indicating that the mutations only affected their stability. Both mercuric reductases isolated from the same environment have shown the complexity by which an enzyme has to adapt to several abiotic stressors simultaneously. Disruption of the geometry and balanced interactions of ATII-LCL-NH and its mutants could explain the remarkable alteration of heat stability. It seems that the molecule behaves as a unit and requires favorable conformation and most probably interactions with the solvent to impart the stability towards heat.
Future prospects

To improve our understanding of the evolutionary distinction between the ATII-LCL and the ATII-LCL-NH mercuric reductase and their mutants, it is suggested to perform thermodynamics analysis to describe the thermal stability behavior. It could also be useful to investigate the effect of combining ectoine and betaine compatible solutes and assess their effect on heat resistance. We could also use additional organic osmolytes and test their effect on their own or in combination.

It is very interesting and useful to create a library of mutants including all the different amino acid substitutions between the ATII-LCL and ATII-LCL-NH mercuric reductases.

To complement our analysis of the molecular structure, it is suggested to include circular dichroism in order to observe the changes in secondary structure (if any) between the wild-type enzyme and its various mutants.

Finally, it is suggested to benefit from advances in nanotechnology and design a carrier for the mercuric reductase enzyme that could enhance its stability, activity, storage, handling and practical use of the enzyme for bioremediation. The thermostable 2D mutant could be nominated as a potential candidate to be used in the bioremediation of mercury burden in gold mines. It is stable and active at the ambient temperatures of such mines.
References

46. Ledwidge, R., et al., NmerA, the metal binding domain of mercuric ion reductase, removes Hg2+ from proteins, delivers it to the catalytic core, and protects cells under glutathione-depleted conditions. Biochemistry, 2005. 44(34): p. 11402-16.


Supplemental Figure 1: PCR amplification of genomic Atlantis-II LCL mercuric reductase.
PCR amplified product, using environmental ATII-LCL DNA and pair of primers, analyzed on 1% agarose gel electrophoresis. The sequences of the primers used in the amplification process are shown in the experimental procedures section. Lane 1, molecular weight marker; Lane 2, amplified DNA fragment.
Supplemental Figure 2: Analysis of randomly selected clones from ATII-LCL merA library.
A) Analysis of recombinant plasmids from ten randomly selected clones on 1% agarose gel electrophoresis. B) Analysis of the size of inserted DNA fragments by PCR using MerA-F and MerA-R primers. MWM refers to molecular weight marker.
Supplemental Figure 3: Induction MerA ATII-LCL-NH at different conditions. SDS-PAGE (10%) of the induced ATII-LCL-NH-merA gene was performed as described in the figure. First lane is the molecular weight marker (MWM).
Supplemental Figure 4: Profile of purification of MerA ATII-LCL-NH recombinant enzyme using HisTrap affinity column.

The ATII-LCL-NH E. coli cell lysate was uploaded into 1 ml HisTrap column equilibrated with 40 mM imidazole using the AKTA purifier machine. On-line AKTA monitor adjusted at 280 nm and 450 nm was used to follow the profile of elution of proteins and FAD respectively. Elution of the protein was achieved by changing the buffer to 500 mM imidazole at 28 ml of the elution volume. 250 µl aliquots were collected for analysis. Pooled fractions (corresponding to elution volume 32-35 ml) containing ATII-LCL-NH MerA are indicated in the figure.
Supplemental Figure 5: Visualization of the HisTrap purification process by SDS-PAGE.

Aliquots from the HisTag purification process were analyzed by SDS-PAGE as described in the experimental procedures. Fractions 1 to 6 are referring to the pooled fractions indicated in Supplemental Figure 4.
Supplemental Figure 6: Profile of purification of MerA ATII-LCL-NH recombinant enzyme on Superdex 75 column.
Pooled aliquots containing the recombinant protein from the HisTag column (Supplemental Figure 5) were uploaded into 24 ml Superdex column equilibrated with PBS buffer using the AKTA purifier machine. On-line AKTA monitor adjusted at 280 nm and 450 nm was used to follow the profile of elution of proteins and FAD respectively. Aliquots, 250 µl, were collected and analyzed. Pooled fractions (corresponding to elution volume 5-8 ml) refer to yellow eluates containing the recombinant protein
Supplemental Figure 7: Assessment of Superdex 75 purification process by SDS-PAGE.
Aliquots from the Superdex 75 purification process were analyzed on 10% SDS-PAGE as indicated in the experimental procedures. Fractions 1 to 4 refer to the pooled fractions shown in Supplemental Figure 6.