ISOLATION, GENOME SEQUENCING, ASSEMBLY AND ANNOTATION OF THERMOPHILIC *Geobacillus thermoleovorans* CCB_US3_UF5 FROM ULU SLIM HOT SPRING, PERAK, MALAYSIA

MUHD KHAIRUL LUQMAN BIN MUHD SAKAFF

UNIVERSITI SAINS MALAYSIA

2011

ISOLATION, GENOME SEQUENCING, ASSEMBLY AND ANNOTATION OF THERMOPHILIC *Geobacillus thermoleovorans* CCB_US3_UF5 FROM ULU SLIM HOT SPRING, PERAK, MALAYSIA

by

MUHD KHAIRUL LUQMAN BIN MUHD SAKAFF

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science

February 2011

ACKNOWLEDGEMENTS

During the whole course of the project I have received a lot of assistance from various parties that enabled me to put together all the information and result. As it is, I sincerely apologize to anyone that I failed to mention due to my oversight and mistake.

First and foremost, my deepest and greatest gratitude and thankfulness to the Allah S.W.T. for without His Grace and Mercy, I would not be able to complete this thesis.

To my main supervisor Professor Maqsudul Alam and my co-supervisor Dr. Rashidah, thank you for everything.

To Dr. Jennifer Saito, thank you for helping me during the writing and submission of the thesis.

My sincerest gratitude to Dr. Shaobin Hou (ASGPB) and Zhemin Zhou (TEDA).

To CCB IT team (Ahmad Rizal, Ahmad Yamin, and Mohd Yusri Ismail), Alexandre Dionne-Laporte and Dr. Bjarne Knudsen, thank you for the assistance in the bio-informatics part of the project.

To Su Yean and Dr. Anton Yuryev, thank you for the assistance in our effort to reconstruct the metabolic pathway of the organism.

To my cherished family, colleague, friend and mentors, everyone, thank you for offering me guidance, helps and advice on not just the matter of research but also on life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AHL	N-acyl homoserine lactones
CDS	Coding sequence
DNA	Deoxyribonucleic acid
ftp	File transfer protocol
GYT	Glucose yeast tryptone
L	liter
M	Molar
meso	mesophilic
mg	miligram
ml	mililiter
OD	Optical density
rRNA	Ribosomal ribonucleic acid
sp.	species
TBE	Tris borate EDTA
µl	microliter
UV	Ultraviolet
UVA	Ultraviolet A
UVB	Ultraviolet B
voc	volatile organic substance

LIST OF SYMBOLS

1X	1 times
10X	10 times
@	At site
°C	Celsius
Ca^{2+}	Calcium ion
H^+	Hydrogen ion
KCl	Potassium chloride
Na+	Sodium ion
NaCl	Sodium chloride
OH^-	Hydroxyl ion
pH	Measure of acidity or basicity of solution
w/v	Weight over volume

PENCILAN, PENJUJUKAN, PENGHIMPUNAN DAN ANNOTASI GENOM BAKTERIA TERMOFILIK *Geobacillus thermoleovorans* CCB_US3_UF5 DARI KOLAM AIR PANAS ULU SLIM, PERAK, MALAYSIA

ABSTRAK

Satu bakteria termofilik *Geobacillus thermoleovorans* CCB_US3_UF5 telah dipencil dari kolam air panas Ulu Slim, Perak, Malaysia. Untuk memahami kemandirian hidup *G. thermoleovorans* CCB_US3_UF5 memerlukan pengetahuan genomnya sebagai pelan perancangan adaptasi terhadap persekitaran panas. Penggunaan teknologi penjujukan generasi terkini telah digunakan untuk penjujukan genom bakteria. Pencilan *G. thermoleovorans* CCB_US3_UF5 menunjukkan perkaitan rapat dengan bakteria termofilik *G. kaustophilus* HTA 426. Draf genom yang diperolehi terdiri dari 7 kontig bersaiz 3506524 bp dengan purata kandungan GC sebanyak 52.3 %. Annotasi meramalkan kewujudan 3955 gen, 73 tRNA, 23 rRNA (7 operon rRNA), 37 gen virus dan 117 gen transposase. Pembinaan semula laluan metabolik berdasarkan genom menunjukkan *G. thermoleovorans* CCB_US3_UF5 mempunyai laluan glikolisis, kitaran asid sitrik, laluan pentos fosfat dan biosintesis asid lemak yang berfungsi.

ISOLATION, GENOME SEQUENCING, ASSEMBLY AND ANNOTATION OF THERMOPHILIC *Geobacillus thermoleovorans* CCB_US3_UF5 FROM ULU SLIM HOT SPRING, PERAK, MALAYSIA

ABSTRACT

A thermophilic bacterium *Geobacillus thermoleovorans* CCB_US3_UF5 was isolated from Ulu Slim hot spring located in Perak, Malaysia. Understanding the survival of *G. thermoleovorans* CCB_US3_UF5 requires knowledge of its genome as the blueprint for high temperature adaptation to the environment. Genome sequencing of the bacterium was performed using next generation sequencing technology. The newly isolated *G. thermoleovorans* CCB_US3_UF5 is closely related to thermophilic bacterium *G. kaustophilus* HTA 426. The draft genome consists of 7 contigs totaling 3,506,524 bp with a mean GC content of 52.3%. The annotation predicts 3955 genes, 73 tRNA, 23 rRNA (7 rRNA operon), 37 phage related genes, and 117 transposase genes. Genome-based metabolic reconstruction indicates that *G. thermoleovorans* CCB_US3_UF5 has a functional glycolysis, citric acid cycle, pentose phosphate and fatty acid biosynthesis pathway.

CHAPTER ONE

INTRODUCTION

During the whole course of the current project, I have come across various publications related to extremophiles. One of them is a book titled ‘Life at the Limits – Organisms in extreme environments’ written by Dr. David Wharton (Wharton, 2002) (Figure 1.1) which helped me in understanding the concept of extremophiles.

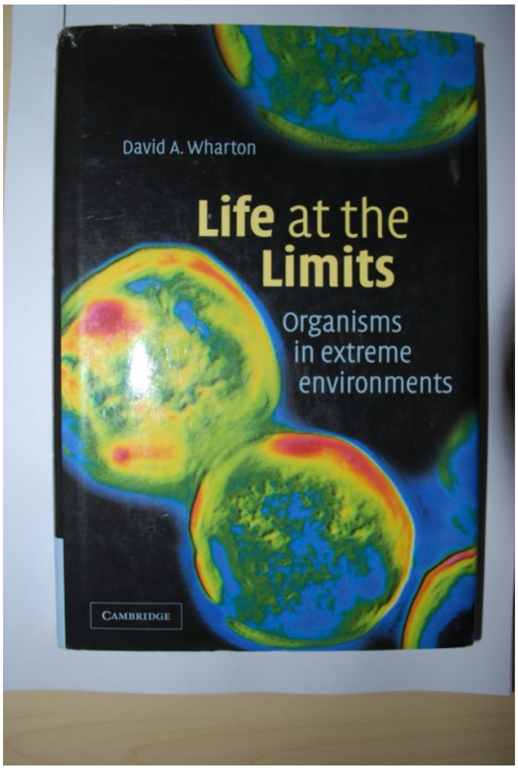

Figure 1.1 ‘Life at the Limits – Organisms in extreme environments’ book

Another informative source is a review article titled ‘Microbial Extremophiles at the Limits of Life’ written by Dr. Elena V. Pikuta (Pikuta *et al.*, 2007) which documented the progress in microbial extremophiles up to year 2007.

Moving beyond these reviews and the book, the field of extremophiles has been enriched with new discoveries on the fundamental and application of extremophiles in life. Also included are the additional findings up to year 2010.

1.1 The Extremophile

Organisms (bacteria, archaea, and eukarya) on Earth live in various kinds of environments. The conditions of these environments are dictated by several variables (stressor) such as physical variables (temperature, pH level, oxygen level, osmolarity, ion balance, access to water, pressure), biological variables (competition, predation, diseases, parasite, food availability), presence of toxic metal, and radiation level. These variables usually work in combination with each other, exposing the organism to various stresses at once (Wharton, 2002).

The range between the lowest and the highest level of these variables that an organism can survive or tolerate can be measured. Using temperature as an example, too low and the organism will have reduced metabolism and freeze, too high and denaturation of protein and irreversible damage to the cell happen (Wharton, 2002).

These variables can be plotted in three dimensional spaces to form a life box (ecological niche) (Figure 1.2). The life box represents the range of conditions where an organism can live or tolerate; going into an environment with variables that lie out of the life box boundary is usually lethal to the organism itself. Having the life box to define the tolerance of an organism makes identification of an extreme organism possible. Majority of the living organisms that we know have their life box clustered together. Compared against the rest, an extreme organism has a different life box pattern. These extremophiles tolerate conditions that are beyond the tolerance of most organisms, thrive and have optimal growth in conditions that are lethal to others. What is extreme to others is normal to these extreme organisms (Wharton, 2002).

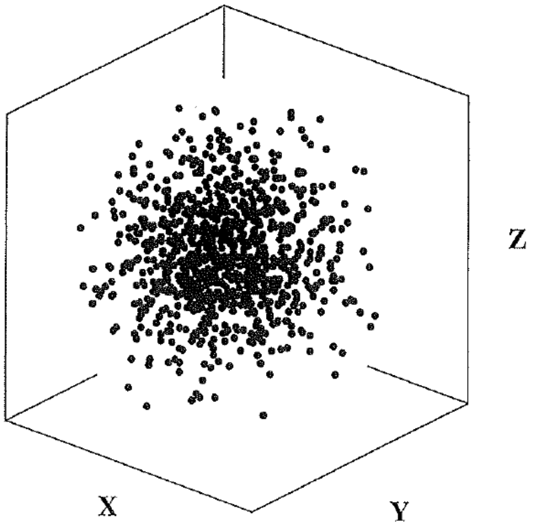

Figure 1.2 A visual representation of a life box. Majority of the organisms are centered in the middle of the box while the extreme organisms are located at the outliers. This figure is taken from the book Life at the Limits (page 5). (Wharton, 2002)

Other types of organisms that could be considered as extreme are the organisms that undergo cryptobiosis (latent state – the metabolic rate of the cryptobiotes drop below 1% of resting state or stop completely) when the environment condition changes beyond their life box boundary, and enter into dormancy instead of dying. These organisms, however, will return to the state of life after the environment becomes favorable again (Wharton, 2002).

For the living organism facing a hostile environment, their survival strategy can be pared down to two main strategies, resistance and capacity adaptation (Wharton, 2002).

Resistance adaptation is based on avoiding or surviving the stress (shown by cryptobiotes by becoming latent) until it returns to normal. On the other hand, capacity adaptation allows the organism to adapt themselves to actually live in harsh conditions (shown by extremophiles as their biological components such as enzymes and membranes are optimized to work in the extreme condition). Most of them can be categorized into two major groups (excluding underground and in organisms), the aquatic extreme environment and also the terrestrial extreme environment (Wharton, 2002).

Aquatic extreme environments tend to be more stable and constant due to water that acts as a buffer to sudden changes in environment variables. This type of environment includes hot spring, deep sea, deep sea hydrothermal vents and cold seep, polar water environment, salt lakes, and soda lakes. Organisms with capacity adaptation thrive in these types of environments (Wharton, 2002).

The terrestrial extreme environments such as desert, temporary desert, temporary water environment, tundra, terrestrial polar region and mountain face temporary extreme condition and the organisms living there usually show resistance adaptation as the main strategy for survival (Wharton, 2002).

1.1.1 Desiccation (absence of water)

The lack of access to liquid water (which results in desiccation) is usually lethal to a living organism due to the requirement of water as the medium of biochemical reactions in the cell. Several strategies are employed to survive desiccations. For larger life forms (plant, animal, insect), they primarily focus on capacity adaptation strategies which includes conservation and prevention of water loss, and also to get and maximize the acquisition of water from sources available in the environment. For smaller organisms (nematode, bacteria, plants) which focus on resistance adaptation, they actually shut down their metabolism, become dried (in turn becoming extremely resistant to majority of the stressors), and remain dormant until water is present. These organisms are called anhydrobiotes. Some of them produce trehalose or other functionally equivalent molecules or specific proteins which help the anhydrobiotes survive during the drying, dried, and hydration phases. (Wharton, 2002)

1.1.2 The thermophile

Moving on from desiccation to the next variable, extremely high temperature is another stressor that influences an organism (Wharton, 2002). Compared to other extreme environments, thermal environments are restricted to hydrothermal areas and environments with sporadic heating like salines and shallow lakes (Cava *et al.*, 2009).

Only microbes are found living in this high thermal environment. Still, the requirement for liquid water limits the highest terrestrial organism existence to 100°C environment while other organisms in deep sea with higher pressure can live in higher temperature due to the water remaining in liquid form at high pressure even beyond 100°C (Wharton, 2002).

The first study into the characterization of thermophiles started in 1888 by Miquel on thermophilic bacteria that grow at 70°C (Adiguzel *et al.*, 2009; Miquel, 1888). Modern thermophile research was triggered by discovery of *Thermus aquaticus* in 1969 (Brock and Freeze, 1969; Pikuta *et al.*, 2007). Currently, *Pyrolobus fumarii* holds the record for maximum temperature growth at 113°C (Blochl *et al.*, 1997; Pikuta *et al.*, 2007) while another hyperthermophilic microorganism known as strain 121 (Kashefi and Lovley, 2003; Pikuta *et al.*, 2007) can survive for a short time at 130°C (Cowen, 2004; Pikuta *et al.*, 2007).

Thermophilic bacteria and archaea are divided into three main groups. The first group comprises of moderately thermophilic microorganisms with optimum growth at temperatures between 50°C to 60°C. The second group, classified as thermophilic has their optimum growth temperature higher than 70°C while the third group, which is the hyperthermophilic, has optimum growth temperature higher than 80°C (Pikuta *et al.*, 2007).

Members of archaea found in thermophilic environment come from phyla Crenarchaeota (Sulfolobales-Thermoproteales branch), Euryarchaeota, and Nanoarchaeota. Nanoarchaeota, represented by *Nanoarchaeum equitans,* draws the interest of researchers due to its nature as a symbiont of another archaea (Huber *et al.*, 2003; Pikuta *et al.*, 2007). It has the smallest genome among all the archaea (Pikuta *et al.*, 2007).

Thermophilic bacteria come from phyla that either only contain thermophilic bacteria as members or have taxa of thermophilic members that are mixed with members from mesophilic bacteria. Aquificae, Thermotogae, Thermodesulfobacteria, Thermomicrobia, and Dictyoglomi only contain thermophilic members, while phyla Deinococcus-Thermus, Nitrospirae, Deferribacteres, Chloroflexi, Cyanobacteria, Proteobacteria, Fermicutes, Actinobacteria, and Spirochaetes are known to have thermophilic bacteria (Lebedinsky, 2007).

Out of these phyla, Aquificae, Thermotogae, and Thermodesulfobacteria have bacteria members that are classified as hyperthermophilic (which are dominated by archaea) while the rest of the phyla occupy the moderately thermophilic environment (Lebedinsky, 2007).

Going down from the hyperthermophiles to the moderately thermophilic microorganisms, the number of species increases as the temperature for the optimal growth is reduced. The reason for the huge diversity of moderately thermophilic bacteria is probably due to the horizontal gene transfer (Pikuta *et al.*, 2007).

At high temperature, the kinetic energy of the environment causes the bonds involved in protein folding to break, denaturing the protein and also the DNA of the organism. Capacity adaptation strategies implemented in a high temperature environment includes alteration to membrane, protein, DNA, and repair and protection mechanism of the cell. The cell membrane of thermophilic organisms generally have a higher saturated fatty acid ratio. In the case of the archaea, their cell membrane forms a monolayer instead of the lipid bilayer, conferring extra stability at high temperature. The amino acid changes increase heat resistant protein folding, increased bonds and bridges, and having a tightly packed interior help the protein to become more resistant to denaturation by heat. DNA bases ratio adjustment and super coiling help to stabilize the microbes DNA at high temperature. Additional measures to survive at high temperature also includes efficient repair and extra protection mechanism like the production of trehalose, 2,3 diphosphoglycerate, chaperones, and histones to stabilize the cells (Wharton, 2002).

Proteins of thermophiles are well adapted to function in high temperature due to several modifications. They have a higher amount of α-helices and β-sheets compared to their mesophilic counterparts (Nagi and Regan, 1997; Pikuta *et al.*, 2007). The unfolding rate of the thermostable protein is very slow, allowing the retention of native structure in denatured condition (Pikuta *et al.*, 2007; Vieille and Zeikus, 2001). They also incorporate more charged and hydrophobic amino acid residues (Haney *et al.*, 1999; Pikuta *et al.*, 2007). Further findings based on the ‘Three principle’ methods shows there are additional factors that contribute to the enzyme thermostability. This includes a decreased loop length, increased secondary structure, decreased labile amino acid residues, increased aromatic stacking, increased hydrophobic interactions, increased metal binding capacity and increased oligomerization (Pikuta *et al.*, 2007; Yano and Poulos, 2003).

Thermophiles have served as sources of many thermostable enzymes we use today (Adams and Kelly, 1998; Hawumba *et al.*, 2002). Due to their thermostable nature, enzymes from high temperature organisms are of interest to the industry. One famous example is the *Taq* DNA polymerase that is extensively used in the biotechnology industry (Wharton, 2002).

Moving beyond the enzyme application of thermophiles, these thermophiles hold a huge interest to the structural biology field as enzymes and macromolecular complexes originating from thermophilic organism are easier to crystallize compared to the mesophiles (Cava *et al.*, 2009).

In the study of the evolution of life, the extreme thermophiles and hyperthermophiles are believed to be the closest representative to the last common ancestor of life (Cava *et al.*, 2009). Knowledge from these thermophiles will provide a better knowledge of the ancient life's biological trait (Cava *et al.*, 2009).

1.1.3 The psychrophile

The temperature of a cold environment presents a different challenge to the organisms than at high temperature. At low temperature the challenges come from lowered kinetic activity (reduced metabolic rate) and also damage from freezing. When facing cold conditions, there are two outcomes depending on the adaptation strategy utilized by an organism. Resistance adaptation strategy (by becoming dormant) is favored by most organisms in an environment with temporary cold conditions (Wharton, 2002).

While psycrophiles are found in the coldness of the extremely vast deep sea, other cold locations like polar and near polar regions also harbor psycrophiles (Pikuta *et al.*, 2007).

Many eukaryotes (dominated by algae) and a diverse physiological group of bacteria and archaea are found in these environments (Hoover, 2009). These also include anaerobes from fermentative bacteria, methanogens, acetogens, sulfate-reducers, iron-reducers, and nitrate reducers. The main producer in these cold ecosystems not only comes from the photosynthetic microorganisms but also chemolithotrophic bacteria (Pikuta *et al.*, 2007).

Survival strategies to a cold stressor includes seasonal dormancy, pigment production, exopolysaccharide production and sheath production (protection against UV) (Hoover, 2009). Other adaptation strategies involve changes of protein functionality (optimal at low temperature) and increased fluidity of membranes by increased unsaturated fatty acid ratio, modifying ante-iso-/iso- branching pattern and short fatty acid length (Pikuta *et al.*, 2007). The cold adapted enzymes traded their thermo stability for more plasticity at lower temperature (Hochachka, 1984; Hoyoux *et al.*, 2004). They also have active sites that are more accessible to ligand (Aghajari *et al.*, 2003; Hoyoux *et al.*, 2004; Russell *et al.*, 1998; Smalas *et al.*, 2000).

Cold shock responses are induced by a sudden reduction of temperature. These lead to production of cold shock protein (Csps) and cold acclimatization protein (Caps). Csps level depend on the degree of the lowered temperature severity while Caps are continuously synthesized in long duration of growth in cold (Pikuta *et al.*, 2007).

The proteins and enzymes from psycrophiles have been utilized in various applications such as cold water detergents, food (flavor) additives, biosensors and bioremediation (Gounot, 1991; Hoover, 2009). The cryoprotectants from these organisms have been extensively used in cosmetics, medicine, and agriculture (Hoover, 2009).

In the study of astrobiology, the knowledge of psychrophiles are relevant as most celestial bodies that exist in the solar system have water (which are required for life) in liquid or frozen form. In addition, microorganisms have the ability to remain alive after being trapped for a long period in polar ice. These ices may provide the additional viable habitat and protection against stress like vacuum and radiation that are prevalent in space (Hoover, 2009).

1.1.4 The barophile

High pressure environments are vast and diverse. This includes the deep sea, deep lakes and deep subsurface region (Pikuta *et al.*, 2007). The deep sea comprises 75% of the total ocean volume (Fang, 2008; Fang *et al.*, 2010). Organisms that live here are called the piezophile or piezotolerant. (Wharton, 2002)

Most of the piezophiles are Gram-negative, facultative aerobic bacteria and many of them are also psychrophilic (Fang *et al.*, 2010). Isolation by pure culture indicates that genus *Shewanella* are widely spread in this environment (Fang *et al.*, 2010; Kato and Nogi, 2001).

Thermophilic piezophile from both Bacteria and Archaea Domain has been isolated mostly from the deep sea thermal vent and the most thermophilic of these are the archaea with ability for chemolithoautotrophs or chemoorganoheterotrophs (Fang *et al.*, 2010).

A high pressure environment causes changes in biochemical reactions, suppressing reactions that cause the volume to increase while enhancing reactions that decrease the volume. Membrane fluidity is reduced and becomes rigid while protein folding is altered. The organisms (microbes and deep sea inhabitant) living in this environment counteract the effect of high pressure using several strategies. Increased amount of unsaturated fatty acid to maintain membrane fluidity, pressure insensitive enzyme, and production of cell membrane protein dictated by the pressure level is part of the capacity adaptation to this environment (Wharton, 2002).

These barophiles are involved in weathering of the volcanic glass that form the new sea floor (Fisk *et al.*, 1998; Pikuta *et al.*, 2007), together with salt water circulation they end up altering the composition of the saltwater itself (Pikuta *et al.*, 2007; Thorseth, 1995).

Some of the discovered piezophiles have shown potential in medical-related application; one species is investigated for potential heparin-like compound (Pikuta *et al.*, 2007; Raguenes *et al.*, 1997) while another piezophile having the ability to tolerate and remove cadmium is being considered for environmental remediation and heavy-metal recovery (Pikuta *et al.*, 2007; Wang *et al.*, 1997).

Interestingly, the high pressure-regulated system for genes found in the piezophile is also in bacteria growing in atmospheric pressure (Pikuta *et al.*, 2007; Sato, 1995; Welch and Bartlett, 1996).

1.1.5 Extreme pH

While most environments, including the deep sea have neutral or near neutral pH, there are environments where the pH level is either extremely low (too acidic with high H^+ concentration) or extremely high (too alkaline with low or absent H^+ concentration with higher OH^- concentration). These environments are not only corrosive to the organisms, they also dissolve and destroy most biological materials.

Organisms that optimally live in acidic pH are called acidophiles and acidotolerant, while the ones that live in high pH are called alkaliphiles and alkalitolerant (Wharton, 2002).

The organisms existing in these extreme pH levels adapt their outer cell component and protein to be resistant to the corrosive environment. They also maintain the internal cellular pH closer to neutral pH by regulation of hydrogen ion across the cell membrane (which is important as most biochemical reaction takes place at near neutral pH). Lastly, these organisms also adapt their enzymes to work in their native harsh condition (Wharton, 2002).

1.1.5(a) The acidophile

The most extreme acidic environments are only accessible to microorganisms. These microorganisms obtain energy by being chemolitothrophic, chemolothomixotrophic, or chemoorganoheterothrophic (Pikuta *et al.*, 2007).

The known acidophiles (hyperthermophiles, moderately thermophiles, thermophiles, and mesophiles) have diverse characteristics. This includes spore-forming, non-spore forming, aerobic, microaerophilic, or obligate anaerobic with positive or negative reaction to gram stain. Some of them may only have a single membrane that made up the cell wall (Pikuta *et al.*, 2007).

Some of the acidophiles are able to maintain their internal pH close to neutral, thus their internal components requiring no adaptation to the acidic environment (Norris, 1998; Pikuta *et al.*, 2007). However, their extracellular proteins are still adapted for optimum function at low pH environment (Pikuta *et al.*, 2007).

Not only facing the extremely acidic environment, these acidophiles also encounter the presence of high soluble metal concentration which blocks the functional groups of enzymes, inhibits transport systems, displaces the native metal from the binding site, and disrupts the cell membrane. These effects are countered by mechanisms like efflux of toxic metal out of the cell; enzymatic conversion, sequestration, exclusion by permeability barrier, and reduced sensitivity of cell (Dopson *et al.*, 2003; Pikuta *et al.*, 2007).

1.1.5(b) The alkaliphile

Alkaliphiles are composed of two major physiological groups, alkaliphiles and haloalkaliphiles (found in soda lakes) (Horikoshi, 1999; Pikuta *et al.*, 2007). Haloalkaliphiles have the additional requirement for high NaCl (Pikuta *et al.*, 2007).

The first alkaliphiles were isolated in 1928 (Downie, 1928; Pikuta *et al.*, 2007). Alkaliphiles have also been isolated from non-alkaline environment instead of traditional alkaline environment. Due to high Na^+ concentration and high pH, the alkaliphiles develop a similar strategy of using osmoprotective compounds to balance the osmotic pressure while maintaining internal pH and developing alternative ion transport pathways (Pikuta *et al.*, 2007). These alkaliphiles have been used extensively in detergents industry (Pikuta *et al.*, 2007).

1.1.6 The halophile

A high salt environment causes several problems to the organisms; they distort protein structure causing these proteins to stick together and no longer function. A high salt environment also causes osmotic shock to the organism, causing loss of water to the environment (Wharton, 2002).

However, we do find organisms that thrive optimally in these conditions. These microorganisms are called the halophiles (which survive in high salt concentration) and the halotolerant (which survive optimally in low salt environment but can tolerate high salt environment) (Wharton, 2002).

Halophiles hold a wide range of metabolic diversity. As salinity increases, the diversity decreases due to the cost of osmotic adaptation to the environment. Except for oxygenic and anoxygenic photosynthesis, aerobic respiration, fermentation and denitrification, other metabolic types do not exist here (Oren, 2002; Pikuta *et al.*, 2007).

The strategies to survive in high salinity are either the "salt in" strategy or the "salt out" strategy (Pikuta *et al.*, 2007). The "salt out" strategy maintains a low salt internal environment through compatible solutes and do not require internal component adaptation (Oren, 1999; Pikuta *et al.*, 2007). However, this requires them to expend energy to pump Na^+ out (Pikuta *et al.*, 2007). Beyond a certain concentration this becomes too energy intensive and nonviable (Pikuta *et al.*, 2007). They also have the ability to take up these compatible solutes from the environments (Pikuta *et al.*, 2007). Archaeal solutes, while similar to bacterial ones, are different only in the charge they carry; Archaeal ones carry a negative charge (Averhoff and Muller, 2010; Martin *et al.*, 1999; Roesser and Muller, 2001). The pathway to produce these solutes have been identified, however the signaling mechanism that triggers them are not understood (Averhoff and Muller, 2010; Wood *et al.*, 2001).

Some organisms in high salt environment do not use osmolytes due to its effects on the cell at high concentration. These organisms instead accumulate potassium chloride (KCl) inside the cell while expelling Na^+ through an ion pumping mechanism. These strategies require the internal cell component to be adapted to work in high K^+ concentration (Wharton, 2002).

The "salt in" solution involves adaptation of internal cellular component to work in high salt. This requires a long evolutionary process and is used only by a few groups of prokaryotes (Pikuta *et al.*, 2007).

Proteins from halophiles are stable and soluble in high KCl. This is due to the presence of high negative charge of the proteins that counteract the salt hydrophobic effect. They have low affinity binding of salt ion which explains their need for high salt concentration to function properly (Pikuta *et al.*, 2007).

Some of these halophiles produce halocins that inhibit many bacteria. Their enzymes are useful due to the ability to function in a high salt environment. Their compatible solutes act as stabilizers in various biological materials, while their exopolysaccharides are useful in remediation of oily environments (Pikuta *et al.*, 2007). These organisms are found in ancient salt crystal from ages ago (McGenity *et al.*, 2000; Pikuta *et al.*, 2007). For the nuclear waste disposal, understanding the role of halophiles in waste container biodeterioration is important as deep salt caves are used for storage of these wastes (Pikuta *et al.*, 2007).

1.1.7 Anaerobiosis

Another important variable is the availability of oxygen in the environment. An environment without oxygen is known as an anaerobic environment. This type of environment was prevalent during the early period of earth. Most of the anaerobic microorganisms (the obligate anaerobic microbes) are vulnerable to oxygen due to the toxicity effect of oxygen radicals. While anaerobic energy production is only 1/18 of the aerobic energy production, they excrete many important compounds due to incomplete metabolism like production of ethanol. For higher organisms, some of the parasites, animals, and plants may face anaerobic condition in certain condition for short time. However, these environments are mainly dominated by microorganisms (Wharton, 2002).

1.1.8 Radio tolerant microorganism

Radiations (UV, X-ray, gamma ray) mostly from the sun represent another variable that affects organism survival. The danger from radiation is due to the high energy level in the radiation itself. These radiations interact with the molecules in the living organism, causing damage, forming ion and free radicals, and causing mutations in the genes due to DNA damage. While the earth's atmosphere absorbs most of the UV (UVA and UVB) radiation from the sun, UV radiation that manages to pass through is prevented from causing damage to the organism by production of pigments that absorb these radiations. They also have efficient DNA repair mechanisms to repair damage from radiation that manages to pass through the entire barrier. Some organisms have the capability to tolerate the high amounts of radiation that are found in environment of nuclear reactor and environment exposed to gamma ray sterilization (Wharton, 2002).

The studies into radio resistant organism began in 1956 with the discovery of *Deinococcus radiodurans*. Since then, other types of radio resistant microorganisms have been found from various environments (Hastings *et al.*, 1986; Lacroix, 2004; Pikuta *et al.*, 2007). There is also effort to isolate these types of bacteria without using the method of exposure to radiation; several microorganisms have been identified through this way (Albuquerque *et al.*, 2005; Pikuta *et al.*, 2007; Sanders and Maxcy, 1979).

The mechanism to resist radiation probably rise due to chronic exposure to a non-radioactive source of DNA damage. Studies on the genome of *D. radiodurans* conclude that the presence of Nudix hydrolase, a homolog of plant desiccation resistance protein, may play a role in the resistance mechanism (Makarova *et al.*, 2001; Pikuta *et al.*, 2007).

Radioresistant bacteria accumulate high intracellular manganese while having low iron level (Ghosal *et al.*, 2005; Pikuta *et al.*, 2007). This might be due to the presence of iron related compounds which may cause oxidative stress and cellular damage after exposure to radiation and lead to cell death. The manganese probably acts as an antioxidant to counteract the oxidative stress on the cell after the radiation exposure (Pikuta *et al.*, 2007).

The potential targets controlling recovery after irradiation includes respiration, tricarboxylic acid cycle activity, peptide transport, and metal reduction (Ghosal *et al.*, 2005; Pikuta *et al.*, 2007). Exposure to vacuum, however, is shown to reduce the survival rate of these microorganisms (Pikuta *et al.*, 2007; Saffary *et al.*, 2002).

1.1.9 Summary of extremophiles

Table 1.1 shows the known types of microorganisms in various environments on Earth and the range of pH, salinity, and temperature where they could be found excluding the deep sea and the subsurface environments.

The extremophile microbes are shown to have huge diversity in the soil. Their diversity is not determined by the temperature, latitude, and variables that affect plants and animals (Fierer and Jackson, 2006; Pikuta *et al.*, 2007). Instead, they depend on the pH and particle size of the soil (Pikuta *et al.*, 2007; Ranjard *et al.*, 2000). Many extremophiles have been found outside of their usual environment (Pikuta *et al.*, 2007).

The archaea, which was first found in extreme environments, exist in many places worldwide in mesophilic environments. They are shown to be interacting with the root of the plant, suggesting that they may have ecological role in the native environment (Bintrim *et al.*, 1997; Pikuta *et al.*, 2007). Methanogens are found in flooded areas and their ability to use H_2-CO_2 or acetate comes in handy due to anoxic conditions imposed by the flood (Liesack *et al.*, 2000; Pikuta *et al.*, 2007).

Table 1.1 Summary of known types of microbial communities and their environment. Not included in this table is the information on deep sea and subsurface community. (Taken from Table 1 in page 184, Pikuta *et al.*, 2007).

Number	Types of microorganism community	pH	NaCl % (w/v)	Temperature (°C)
1	Freshwater (inland) • psychrophilic • meso-thermal • moderately thermophilic • thermophilic	5 - 7	0 - 1	Less than 10 15 – 40 50 – 60 70 – 100
2	Marine (sea) • psychrophilic • meso-thermal • moderately thermophilic • thermophilic	8	3 - 4	Less than 10 15 – 40 50 – 60 70 – 120
3	Alkaliphilic • psychrophilic • mesophilic • moderately thermophilic • thermophilic	9 - 11	0 - 1	Less than 10 15 – 40 50 -60 70 – 110
4	Haloalkaliphilic • psychrophilic • mesophilic • moderately thermophilic	9 - 10	3 - 25	Less than 10 15 – 40 50 – 60
5	Halophilic • psychrophilic • mesophilic • moderately thermophilic	8	3 - 30	Less than 10 15 - 40 50 – 60
6	Acidophilic • mesophilic • moderately thermophilic • thermophilic	0 - 4	0 - 2	15 – 40 50 – 60 70 – 120

Figures 1.3, 1.4, 1.5, and 1.6 sums up the known extremophile limits based on pH, temperature, salinity, and radiation.

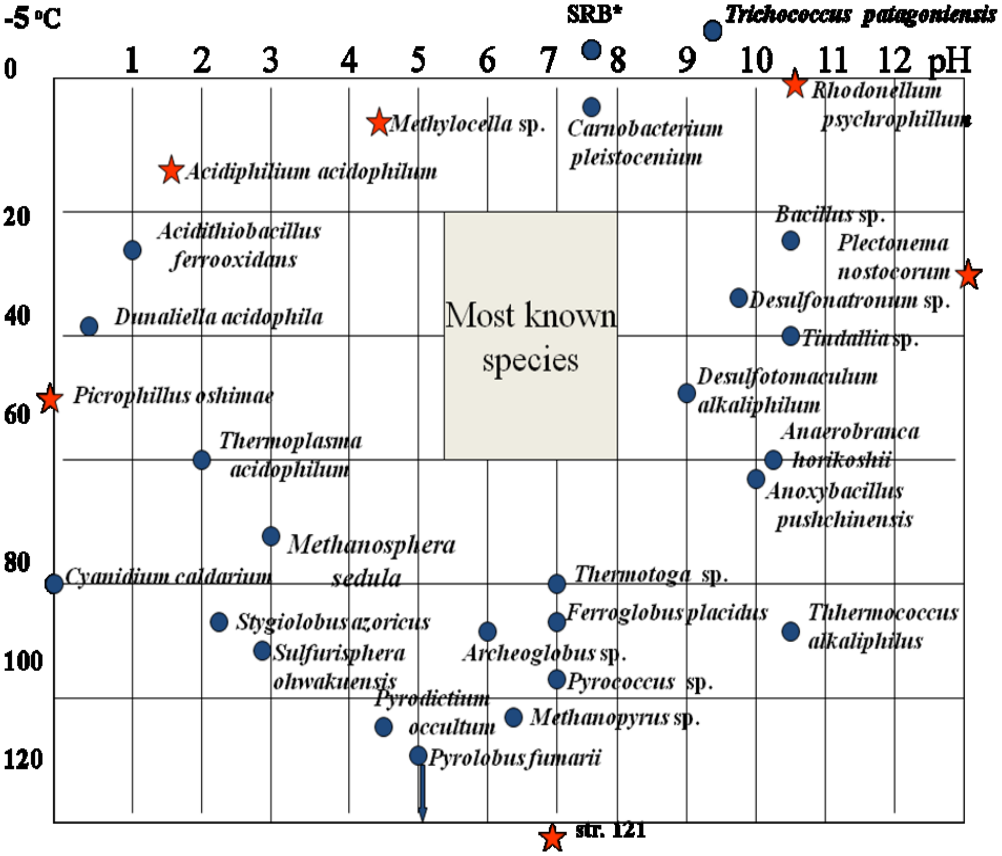

Figure 1.3 Diagram of all known validated microorganisms distributed within matrix pH/temperature. (*- species of five genera of psychrophilic sulfate-reducing bacteria) (★, ● - organism) (↓ - can grow up to the following temperature). This figure was taken from Figure 5.2 in page 129, Hoover, 2009.

Referring to Figure 1.3 which plotted temperature against pH, majority of the organism live at 20 - 60°C and neutral pH (shown as grey box). No true acidic psychrophiles are found until now. Mesophilic cyanobacteria are the most alkalitolerant species known, while one alkaliphilic bacterial species demonstrated psychrophilic nature (Pikuta *et al.*, 2007).

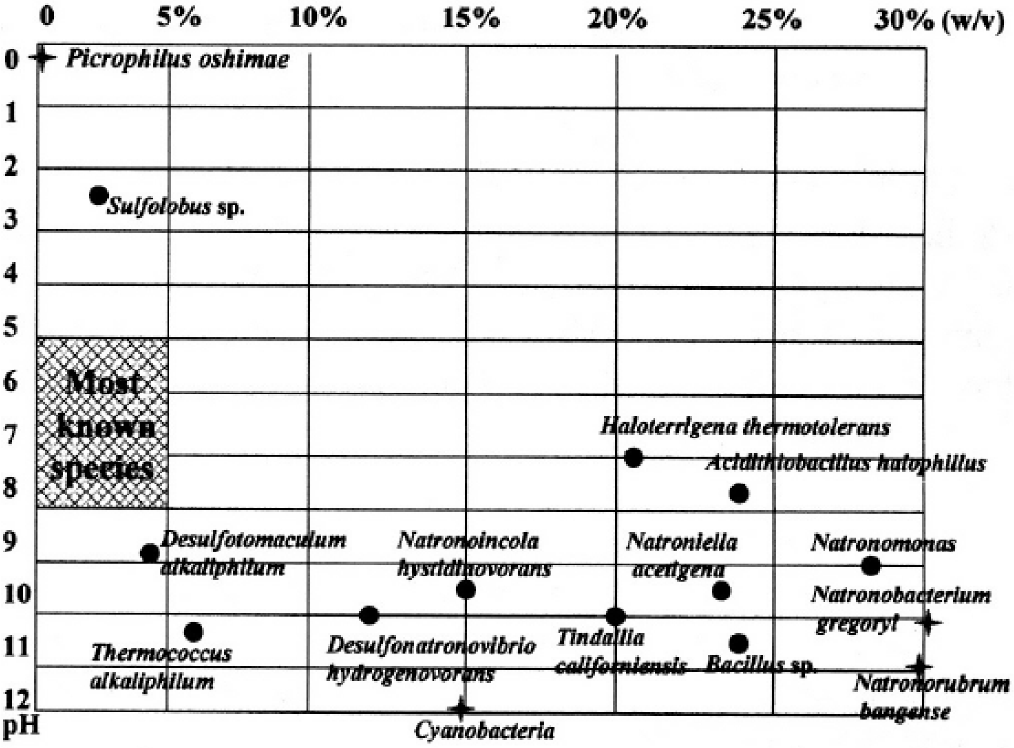

Figure 1.4 Diagram of all known validated microorganisms distributed within % (w/v) salinity vs pH. (●,✦- organism). This figure was taken from Figure 1 in page 202, Pikuta *et al.*, 2007).

In Figure 1.4, majority of the microorganisms tolerate up to 5% salinity (crisscross square box). The most extreme acidophiles tolerate low salinity. While some species can tolerate up to 30% salinity, Cyanobacteria are limited up to 15% salinity. This chart also shows the possibility of halophiles coming into existence later during evolution due to accumulation of Na^+ and Cl^- ion in water in the later stage of Earth development (Pikuta *et al.*, 2007).

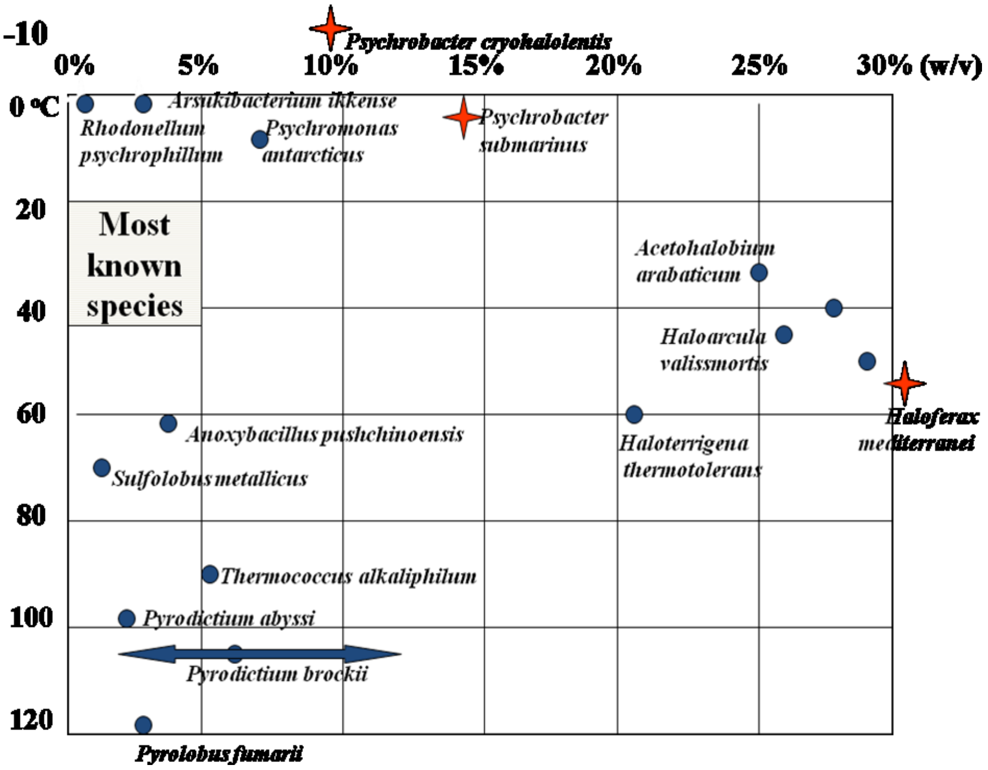

Figure 1.5 Diagram of all known validated microorganisms distributed within NaCl -salinity/temperature. (✦, ● - organism) (⟷ - found in the following range of NaCl concentration). This Figure is taken Figure 5.2 in page 129, Hoover, 2009.

No extreme thermophiles or psychrophiles exist in high salinity environment (Figure 1.5) (Hoover, 2009). The mesophilic and moderately thermophilic species dominate the highly saline environment. Possibly the early ancestors of life might not be an alkalophile or the ancient ocean is not salty (Pikuta *et al.*, 2007).

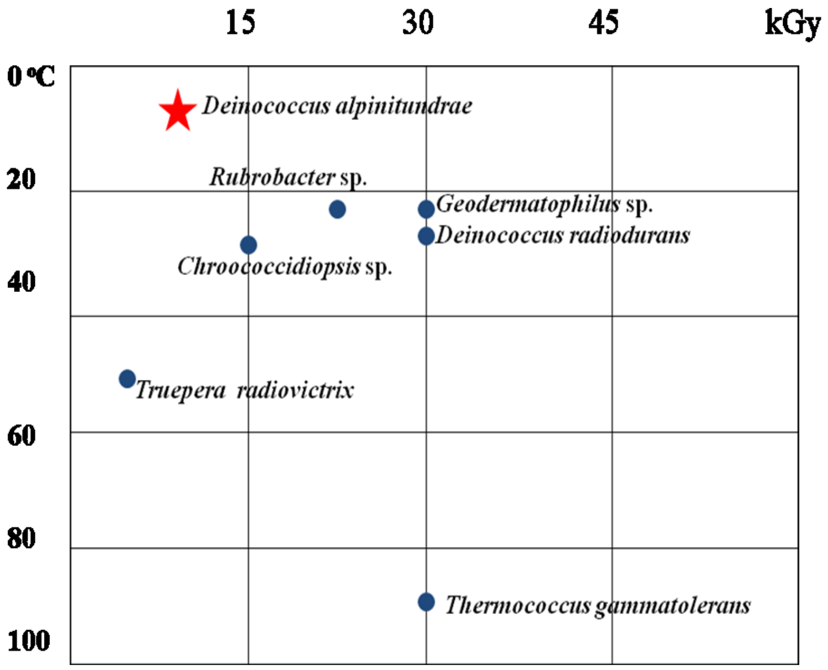

Figure 1.6 Diagram of all known validated microorganisms distributed within matrix temperature/gamma-radiation. (★, ● - organism). This figure was taken from Figure 5.2 in page 130, Hoover, 2009.

In many theories regarding the origin of life on Earth, the extremophiles are theorized to play a huge role in the development of life. As the early era Earth was hot, thermophilic microorganisms were likely the first ancestors of life. However, some researchers do not agree with this theory and instead concluded that thermophilic existence originates from mesophiles which gradually adapted to thermal conditions (Boussau *et al.*, 2008; Pikuta *et al.*, 2007).

With Earth cooling down with passing time, the psycrophiles came into existence and started occupying the low temperature niches where the thermophiles were unable to survive (Pikuta *et al.*, 2007).

Returning back to the concept of the life box, the changes of the earth's environment throughout time are also reflected by the organisms as they keep up with the changes which create new environments while extinguishing the previous ones. The life box of the majority of them also changes following these adaptations.