Structural and Functional Insights of Virulence Proteins and Associated Factors from *Pseudomonas aeruginosa*

A thesis submitted for the award of degree of

Doctor of Philosophy (Science)

In Life Science and Biotechnology

by

Atanu Pramanik

Structural Biology and Bioinformatics Division

CSIR-Indian institute of Chemical biology, Kolkata, India

Department of Life Science & Biotechnology

Faculty Council of Science, Jadavpur University

Kolkata – 700032, INDIA

2024

सी.एस.आई.आर–भारतीय रासायनिक जीवविज्ञान संस्थान
वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद की एक इकाई
विज्ञान एवं प्रौद्योगिकी मंत्रालय के अधीन, एक स्वायत्त निकाय, भारत सरकार
4, राजा एस. सी. मल्लिक रोड, यादवपुर, कोलकाता - 700 032
CSIR - INDIAN INSTITUTE OF CHEMICAL BIOLOGY
A Unit of Council of Scientific & Industrial Research
An Autonomous Body, under Ministry of Science & Technology, Government of India
4, Raja S. C. Mullick Road, Jadavpur, Kolkata-700 032

CERTIFICATE FROM THE SUPERVISOR(S)

This is to certify that the thesis entitled " **Structural and Functional Insights of Virulence Proteins and Associated Factors from *Pseudomonas aeruginosa*** " Submitted by Sri / Smt. **Atanu Pramanik** who got his / her name registered on **20.09.2018** for the award of Ph. D. (Science) Degree of Jadavpur University, is absolutely based upon his own work under the supervision of **Dr. Saumen Datta** and that neither this thesis nor any part of it has been submitted for either any degree / diploma or any other academic award anywhere before.

डॉ. सोमेन दत्ता / Dr. Saumen Datta
मुख्य वैज्ञानिक / Chief Scientist
सीएसआईआर-भारतीय रासायनिक जीवविज्ञान संस्थान
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
CSIR-Indian Institute of Chemical Biology
(Council of Scientific & Industrial Research)
4, राजा एस. सी. मल्लिक रोड, यादवपुर, कोलकाता - 700 032
4, Raja S. C. Mullick Road, Jadavpur, Kolkata - 700 032

SDatta

25-06-2024

Dr. Saumen Datta
Chief Scientist
Structural Biology and Bioinformatics Division
CSIR-Indian Institute of Chemical Biology
Jadavpur, Kolkata-700032

Dedicated

to

My Family

ACKNOWLEDGEMENTS

Since my graduation, I have dreamed of nurturing my ideas and finding new things, ultimately resulting in the work done for this thesis. With a great deal of patience and perseverance, I was able to maintain my flow during the entire arduous process. It would not have been feasible without the faith and encouragement of many people during this whole journey at IICB. Now is the perfect moment to express my gratitude to everyone with whom I have been connected along this arduous and lengthy journey.

First of all, I would like to express my gratitude to my supervisor, ***Dr. Saumen Datta*** *(Chief Scientist), without whom this would not have been possible. I am fortunate and appreciative of him for giving me the chance to work with him. I am also very fortunate to have learned Macromolecular Crystallography under his guidance. I want to sincerely thank him for all of his assistance and guidance throughout this course. I was able to finish this work because of his continuous supervision and mentoring.*

Next, I want to thank ***Prof. Arun Bandyopadhyay*** *(Ex-Director, CSIR-IICB) and* ***Prof. Vibha Tandon*** *(Director, CSIR-IICB) for giving me the opportunity to work here at IICB.*

In order to proudly complete my work here, I also want to thank the ***University Grant Commission*** *of the Government of India for providing me with a fellowship during my doctoral studies. I thank all of the faculty members of the Department of Life Science and Biotechnology at Jadavpur University for allowing me to register for Ph.D.*

I would also like to express my heartfelt gratitude to a few scientists at Department of Structural Biology and Bioinformatics of IICB, starting with ***Dr. Jayati Sengupta*** *(Senior Principal Scientist),* ***Dr. Subrata Adak*** *(Chief Scientist), and* ***Dr. Krishnananada Chattopadhyay*** *(Chief Scientist) who have supported me during my research work.*

I am constantly appreciative of my lab colleagues starting with my seniors ***Chittran Roy, Abhisek Mondal, Basavraj khannpnavar, Arkaprabha Chowdhury,*** *and* ***Gourab Basu Chowdhury****. I want to thank* ***Rajeev Kumar*** *from the bottom of my heart for all of his assistance and guidance along this entire long journey. His guidance regarding cloning and other experimental techniques helped me a lot in fulfilling my work during this entire Ph.D. journey. I also want to thank my juniors* ***Bidisha, Angira, Maruf, and Debayan****. A special thanks to* ***Sushanta Da*** *for helping with my research work during the whole Ph.D. journey.*

Besides my lab members, I want to express my heartfelt gratitude to my colleagues from other labs at IICB starting from ***Shashi Kant, Sumit Das and Vivek Kumar Gupta*** *for their immense help in my entire research work. Without their help in my research work and staying by my side during this entire journey, it would have been impossible to complete my work. Also, thanks to other lab members starting from* ***Bhakta Da, Bani Da, Manideep Da, Biprasekhar Da, Sirin Di, Saroj Da, Ayan Da, Chinmoy Da, Shiladitya Da, Priya, Krishna, Aneek, Dipanjan, Sumangal, Yuthika, Puja, Gourab, Saikat, Debanjan, Sinjini, Subhasish, Arnab****. I will be always thankful to them for their generous support. A special thanks to* ***Aditya*** *from Microbiology Department of Calcutta University, for giving support in my work.*

In addition, I would like to express my gratitude to the personnel at IICB, particularly the technical staff at the Central Instrument Facility, who have consistently offered assistance and recommendations while conducting the experiments there. Special thanks in this regard to ***Ramdhan, Jishu, Vignesh*** *and* ***Beni****.*

I also want to extend my sincere gratitude to my childhood friends ***Sourav, Daipayan****, and* ***Rony*** *who always encouraged and believed in me and my dedication to my research work. I will always treasure the sweetest moments spent with them since my childhood, which will be one of my favorite recollections in my life.*

Finally, I would want to express my gratitude to my ***brothers*** *and* ***sisters*** *as well as my entire family for their unwavering support over the years. Not enough words can adequately convey how grateful I am for blessing, cooperation, and kind patience of my parents throughout my whole Ph.D. journey. Thanks to Lord Shiva for his blessings in my life. Thank you all.*

Atanu Pramanik
CSIR-IICB, Kolkata
2024

Table of contents

ABBREVIATIONS

ACOD 1	Aconitate decarboxylase 1
CLYBL	Citrate lyase subunit beta-like
CF	Cystic Fibrosis
K_d	Dissociation constant
EDTA	Ethylenediaminetetraacetic acid
GAP	GTPase activating protein
IRG	Immunoresponsive gene
ICL	Isocitrate lyase
IPTG	Isopropyl-β-d-thiogalactoside
ITC	Isothermal titration calorimetry
kDa	Kilodalton
LPS	Lipopolysaccharides
LB	Luria Broth
MCM	Methylmalonyl-CoA mutase
µg	Microgram
µl	Microlitre
µM	Micromolar
PAMP	Pathogen associated molecular pattern
PMSF	Phenylmethylsulfonyl fluoride
PLC	Phospholipase C
PEG	Polyethylene glycol
RIP	Required for intracellular proliferation
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SDH	Succinate dehydrogenase
SDH	Succinate dehydrogenase
TCA	Tricarboxylic acid

List of Figures

List of Tables

ABSTRACT

Pseudomonas aeruginosa is an opportunistic multi-drug resistance pathogen which secretes some virulence factors or proteins either extracellularly or via some specialized secreting nanomachines called injectisomes into the host cells. These virulence proteins also referred to as effectors molecules subvert the host immune response by targeting some essential proteins or some immunomodulatory substances in the host cells. Therefore, it is essential to comprehend these pathogenic variables which are necessary for the development and maintenance of bacterial pathogenicity. Atomistic level insights into these virulence factors or proteins with or without their targeting partners decipher the molecular level information in pathogenesis.

To overcome the infection caused by these virulent organisms, natural killer cells like macrophages secrete itaconate, an immunometabolite substance which specifically inhibits isocitrate lyase of bacterial glyoxylate cycle. To overcome the inhibitory effect of itaconate, *P. aeruginosa* release three enzymes: succinyl-CoA:itaconate CoA transferase (Ict), (*R*)-specific itaconyl-CoA hydratase (Ich) and (*S*)-citramalyl-CoA lyase (Ccl) which alter itaconate to pyruvate and acetyl-CoA. Here, we report the first crystal structure of itaconyl-CoA hydratase from *P. aeruginosai* (*Pa*Ich) at 1.98 Å resolution. The overall structure of *Pa*Ich resembles the structure of MaoC family (*R*)-specific enoyl-CoA hydratase consisting of two domains, N-terminal domain and C-terminal domain connected by a long intervening loop. Each domain is comprised of a 'hotdog fold' where a central α-helix is represented as sausage wrapped by a β-sheet scaffold represented as a bun. Crystal structure analysis of *Pa*Ich showed that a unique N-terminal hotdog fold containing a 4-residue short helical segment 'α3', named as an 'eaten sausage', followed by a flexible loop region slipped away from the conserved β-sheet scaffold, whereas the C-terminal hotdog fold is similar to all MaoC family hydratases.

PvrA was discovered to be an upregulated regulatory gene during infection, which increases the virulence of the bacteria in the host. Here, we report the crystal structure of PvrA from *P. aeruginosa* at 2.30 Å resolution. Structural analysis of PvrA showed that it present as a dimer which is very similar to other TetR family transcriptional regulators such as AmtR and AcrR. It consists of a common N-terminal H-T-H DNA binding domain similar to other TetR family regulators. We have found that region consists of ~5 residues (Val_8-Arg_{12}) of H-T-H motif of chain B is slightly bent downwards compared to chain A.

Chapter 1

Introduction

From premodern times to the modern era, infectious diseases are still the main cause of overall mortality rate of more than 13.5 million deaths annually. Infectious diseases are of many different types and pose a quite significant threat to public health at the time of emergence on various time scales. Due to increasing number of populations worldwide, the rate of transmission of diseases has exponentially proliferated with the emergence of more virulent strains. With the advent of novel therapeutics, the prevalence of morbidity and mortality associated with infectious diseases is reduced nowadays. However, it still presents adverse effects in our society. Of them, lower respiratory tract infections and diarrhoeal diseases are major health concerns at present and new-generation drugs cannot exterminate them properly.

1.1 The genus *Pseudomonas*

Being an opportunistic pathogen, *Pseudomonas aeruginosa* is a gram-negative, rod-shaped, motile, non-fermentative bacteria which is ubiquitously found in soil and water, as well as in plants and animals including humans [1, 2]. It resembles a rod-shaped cell measuring 1-5 μm in length and 0.5-1.2 μm wide. *P. aeruginosa* is a facultative anaerobe which can proliferate in partially oxygen-deprived conditions where nitrate or nitrite is used as a terminal electron acceptor. In aerobic conditions, it utilizes glucose as a sole carbon source, catabolized by a glycolytic pathway including entner-doudoroff pathway where oxygen is used as a final electron acceptor [3]. Some organic compounds such as phenol sulfates, phenylalanine, benzoate and 2,4-butanediol are sometimes used as sole carbon sources [4]. *P. aeruginosa* is adaptable to different environmental conditions and also metabolically versatile, so they are the most predominant organism on earth.

1.2 Genome of *Pseudomonas aeruginosa*

As of 2023, a total of 30193 strains of *P. aeruginosa* from both clinical and environmental sources have been fully or partially sequenced, according to NCBI database. With a genome size of over 7 million base pairs and a G + C content of approximately 66%, *P. aeruginosa* has one of the largest genomes among the numerous pathogenic or non-pathogenic bacterial genomes that have been sequenced to date [5]. Nearly 90% of the entire genome is found in the highly conserved core genome of all *P. aeruginosa* strains, together with accessory genome components. The accessory genome elements range from a few hundred base pairs to more than 200 kbp. The 5109 protein-coding genes comprise the core of *P. aeruginosa* strains [6]. In contrast, the accessory genome elements consist of genes which are not thoroughly

distributed in all strains. Moreover, these segments termed ‘regions of genome plasticity’ are not distributed randomly in core genome elements; rather, they are specified in certain loci. The genetic sequences of many RGPs (regions of genomic plasticity) are often termed genomic islands (>10 kb) or islets (<10 kb), evolved by horizontal gene transfer between various strains for adaptability in adverse environmental conditions [7]. The accessory genome elements of *P. aeruginosa* possess significant therapeutic interest due to their ability to improve virulence, antimicrobial resistance and overall robustness through horizontal gene transfer [8].

Among many strains, *P. aeruginosa* PAO1 and PA14 have been extensively studied for therapeutic interest and to understand the genetics, physiology and metabolism of the spectrum. The genomes of both strains are fully sequenced and also available in NCBI database. PAO1 strain is a clinical isolate, the derivative of initial PAO strain, obtained from a wound patient in Melbourne, Australia, in 1941 [9]. PA14 or UCBPP-PA14 is also a clinical isolate obtained from a burn wound patient and a much more virulent strain for plants and animals [10]. This strain was initially part of a collection from University of California Berkeley Plant Pathology Laboratory, hence the name “UCBPP-PA14”. Initially, the research was started to find the link between plant pathogens and human infections. Nowadays, it is the most preferred model for studying virulence and developing novel therapeutics to prevent dissemination and pathogenesis.

1.3 Pathophysiology of *Pseudomonas aeruginosa*

P. aeruginosa can infect a broad spectrum of host targets including nematodes, insects, plants and mammals. It is generally found in the gut microflora of a healthy individual but causes infection in immunocompromised patients; hence, they are called opportunistic pathogens [11]. Being resistant to multiple drugs, *P. aeruginosa* causes acute and also chronic infections in immunocompromised individuals with cystic fibrosis, coronary obstructive pulmonary disorder (COPD), burns, cancer, trauma, sepsis and ventilator-associated pneumonia (VAP), including those caused by COVID-19 [12, 13]. It is a member of the MDR ESCAPE pathogens where ESCAPE stands for *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanii*, *P. aeruginosa*, and *Acinetobacter baumanii*. The carbapenem-resistant strain of *P. aeruginosa* belongs to ‘critical’ group of pathogens recommended by WHO [14]. *P. aeruginosa* has a frequency of 7.1–7.3% among people who are connected to the healthcare industry and is responsible for hospital-acquired infections

such as pneumonia, wound infections, urinary tract infections, and bacteremia **(Figure 1.1)**. It is mostly found in nosocomial pneumonia and the prevalence is rising over time in intensive care unit (ICU) patients. It has been estimated that *P. aeruginosa* was the cause of 23% of all ICU-related infections. Healthcare-associated pneumonia (HAP) and ventilator-associated pneumonia (VAP) are a significant threat to healthcare personnel, reporting 22% of all hospital-acquired infections [15]. Mortality from VAP was estimated to be 32-43% [16] [17].

P. aeruginosa causes nosocomial urinary tract infections, mostly catheter-associated UTI (CAUTI). They are the reason for ~10% of all CAUTIs and account for approximately 16% of all UTIs in ICU patients [18]. Studies showed that ICU patients had greater resistance rates—more than 40%—to antibiotics such as meropenem, fluoroquinolones, and piperacillin-tazobactam.

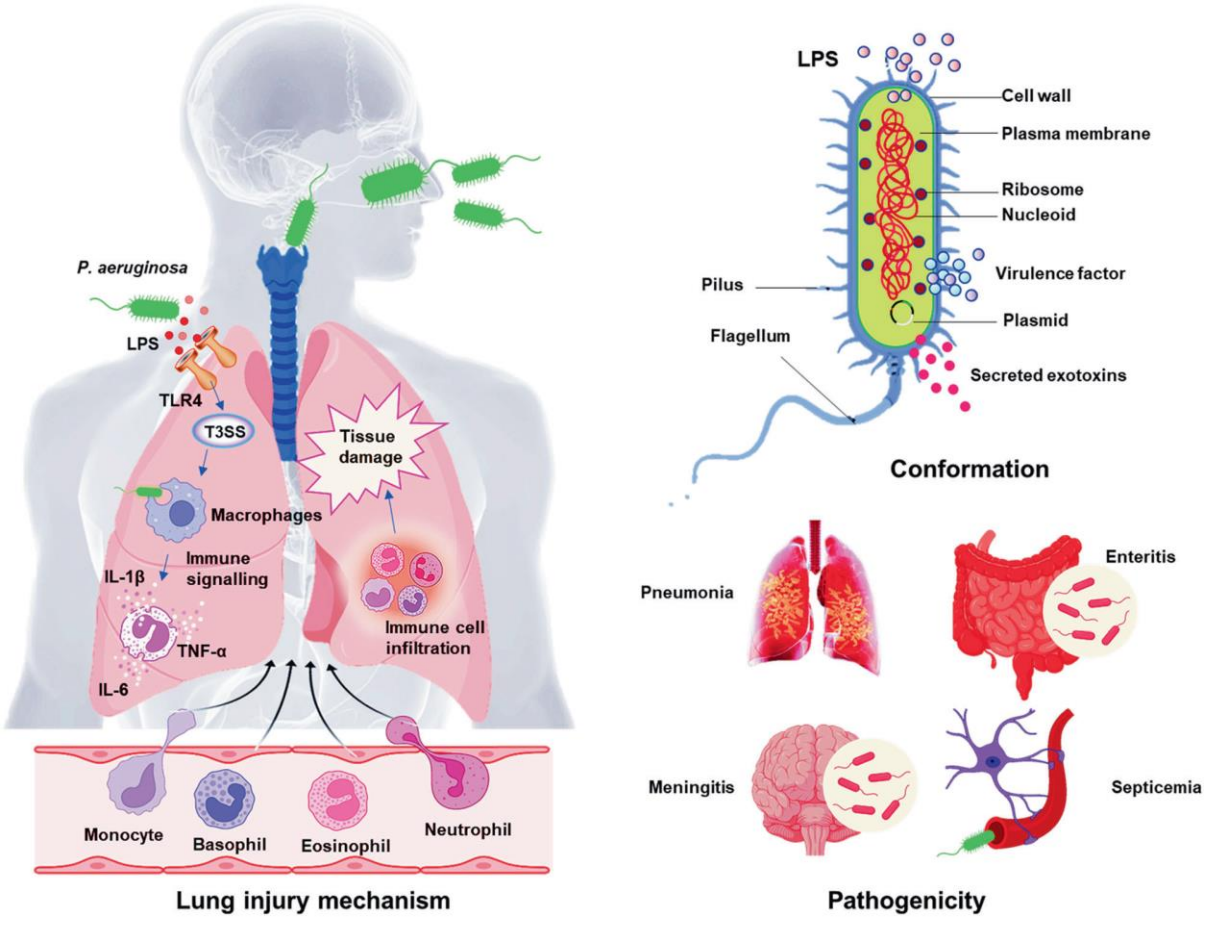


Figure 1.1: Overview of pathogenesis of *P. aeruginosa*. It is found everywhere and causes infections to any organs. LPS stimulates inflammatory responses in lung after infection. Activation of immune cells and recruitment of neutrophils is a promising sign of inflammatory response in the body. Excessive immune cells infiltration cause severe tissue damage and aggravate bacterial infections.

Adapted from Qin, S., Xiao, W., Zhou, C., Pu, Q., Deng, X., Lan, L., Liang, H., Song, X., & Wu, M. (2022). Pseudomonas aeruginosa: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics. *Signal transduction and targeted therapy, 7(1)*.

In cystic fibrosis (CF) patients, *P. aeruginosa* is a predominant cause of morbidity and mortality [19]. CF patients possess complications including chronic respiratory tract infection, structural lung disorder, bronchiectasis, airflow obstruction, and ultimately death. *P. aeruginosa* may thrive in CF-associated lungs due to its ability to survive in hypoxic conditions [20]. The inherited genetic abnormalities in the CF transmembrane conductance regulator (CFTR) gene caused these issues in CF patients. The reason behind the morbidity and mortality of CF patients is recurring bacterial infections in abnormal mucus layer [21, 22]. The CFTR regulator controls the movement of chloride ions and electrolytes across epithelial cell membranes to preserve homeostasis and typical mucus characteristics. As a result, the CF lungs have an excessively thick, sticky, and dried layer due to the loss of activity of the CFTR regulator [23]. Therefore, *P. aeruginosa* respiratory tract infections are a major risk factor for CF patients starting in infancy. *P. aeruginosa* strains that cause acute infections can be obtained by nearly 30% of CF newborns from their surroundings. *P. aeruginosa* supports a variety of phenotypes and molecular pathways necessary for survival throughout pathogenesis and antibiotic treatment utilizing complex genotypic events. Thus, several virulence and intrinsic antibiotic resistance mechanisms mediate survival during the early phases of CF lung colonization. Following infection, bacteria are subjected to oxidative stress and other inflammatory reactions before receiving antibiotic therapy [24, 25]. *P. aeruginosa* can adapt and change to resistant and persistent phenotypes by expressing distinct sets of genes in response to certain environmental stress factors [26, 27].

The symbol of bronchiectasis is the thickening and dilatation of the airways, which causes airway inflammation and persistent infection. Compared to CF, non-CF bronchiectasis is more prevalent and exhibited by some conditions, including immunodeficiency, primary ciliary dyskinesia, post-infectious, and idiopathic bronchiectasis [28]. One of the most often isolated species in non-CF bronchiectasis patients is *P. aeruginosa*, which is linked to decreased lung function and higher mortality rates [28, 29].

In addition, it is also more commonly found in immunocompromised patients suffering from neutropenia. It is also a clinically significant pathogen found in patients with haematological

malignancies [30, 31]. The number of therapeutic choices for treating infections is steadily declining due to the development of a molecular arsenal that confers resistance to several types of antibiotics. Simultaneously, the prevalence of infection and the emergence of multidrug-resistant strains are still increasing.

1.4. Virulence factors

P. aeruginosa secretes a range of virulence factors to adapt to the unfavourable conditions of its hosts. These factors aid in the infection process and the spread of disease. It possesses a wide range of virulence factors such as secretion systems that transport toxins and effectors into the host, flagella, pilli, and LPS that aid in bacterial adhesion and colonization, quorum sensing and biofilm that provide drug resistance and proteases that cause tissue damage [32, 33]. The virulence factors of *P. aeruginosa* were extensively studied and some of them are summarized below.

Gram-negative bacteria have an outer membrane comprised of mainly lipopolysaccharide (LPS), which is present in all strains of *P. aeruginosa*. Toll-like receptors (TLRs) and other pattern recognition receptors (PRRs) can detect pathogen-associated molecular patterns (PAMPs), which are tiny molecular motifs conserved in a class of microorganisms. This allows PAMPs to trigger innate immune responses, thereby shielding the host against infection [34]. Although LPS shields bacteria from phagocytosis, it can cause neutrophils to release neutrophil extracellular traps (NETs) to catch invasive pathogens. Interestingly, the majority of LPS-targeted vaccinations developed against *P. aeruginosa* have not demonstrated adequate efficacy, despite promising outcomes in their development [35].

P. aeruginosa secretes several proteases which have crucial roles in pathogenesis. The type I secretion system secretes alkaline protease (AprA), a virulence factor regulated by the quorum-sensing circuit. AprA can break down TNF-a, IFN-g, and complement components, which can exacerbate infections in the body and weaken host immune function [36, 37]. *P. aeruginosa* produces the enzymes LasA and LasB, which break down elastin, a crucial part of blood vessels and lung tissue, affecting lung function and leading to pulmonary bleeding [38]. The corneal pathogenicity of *P. aeruginosa* is strongly correlated with protease IV [39]. Protease IV can damage host tissues and promote bacterial infection by breaking down fibrinogen, lactoferrin, transferrin, and elastin, among other physiologically significant molecules. It can also compromise host immunity by breaking down immunoglobulins, complement components, and surfactant proteins [40]. Since protease IV is a serine protease,

therapies to stop *P. aeruginosa* infection-related tissue damage may be developed with known serine protease inhibitors which restrict its enzymatic activity.

ExoS, ExoT, ExoU and ExoY are effectors, also known as toxins of *P. aeruginosa*, secreted by the type III secretion system. They both are bifunctional cytotoxins that have both adenosine diphosphate ribosyl transferase (ADPRT) and GTPase activating protein (GAP) activity. They perturb cell-to-cell adhesion by disrupting host actin cytoskeleton which ultimately leads to apoptosis of host cells. ExoU is the most virulent T3SS effector and it is also a more potent phospholipase which causes severe necrosis of host cells, ultimately leading to cell death. ExoY is categorized as an adenylate or nucleotidyl cyclase that substantially raises cAMP, cGMP, cUMP, and, to a lesser degree, cCMP levels.

ExlA of *P. aeruginosa* exhibits cytolysin activity and is encoded by the unique two-gene genetic element exlA-exlB. ExlA causes permeabilization of the plasma membrane, which leads to necrotic cell death. This mechanism depends on the type IV pili for bacterial attachment [41]. Among the most toxic exocellular factors secreted by the T2SS, exotoxin A of *P. aeruginosa* (PEA) is controlled by the metabolisms of iron and glucose [42, 43]. It causes programmed cell death and limits the synthesis of host proteins by catalyzing the ADP ribosylation of cell elongation factor 2 (EF-2) [44].

The primary extracellular lipase named Lipase A (LipA) is released by type II secretion system of *P. aeruginosa*. It can cause severe damage to tissues by breaking down host cell membranes and dipalmitoylphosphatidylcholine, the primary lung surfactant lipid [45, 46]. *P. aeruginosa* produces hemolytic and non-hemolytic forms of phospholipase C (PLC). Numerous investigations have demonstrated that hemolytic PLC can cause organ damage, host vascular permeability, and cell death [47]. Consequently, it helps bacteria survive in neutrophil-rich environments and induces chronic bacterial infection [48, 49].

P. aeruginosa secretes a lipoxygenase called LoxA, which can regulate the bacterial invasion process and impair host lipid signaling [50]. When it infects the lung, LoxA oxidizes a range of host polyunsaturated fatty acids and generates lipoxin A4, a lipid peroxidative mediator that kills host cells [51]. LoxA also prevents the recruitment of immune cells by blocking the production of key chemokines including keratinocyte chemoattractant (KC) and macrophage inflammatory protein (MIP) [52]. The severity of the *P. aeruginosa* infection may be lessened by inhibiting the activity of LoxA.

In addition, *P. aeruginosa* has some catabolic pathways which are used for degradation of anti-microbial products secreted by macrophages and neutrophils, the secondary line of defense systems in humans. These immune cells release various substances upon bacterial infection. Among these compounds is itaconate, a C5-dicarboxylic acid that functions as a strong inhibitor of isocitrate lyase in the glyoxylate cycle, a process that is necessary for the survival of bacteria during infection. Three enzymes aid in the degradation of this immunomodulatory compound, itaconate [53]. succinyl-CoA:itaconate CoA transferase (Ict) first converts itaconate to itaconyl-CoA, then itaconyl-CoA is hydrated to (*S*)-citramalyl-CoA by (*R*)-specific enoyl-CoA hydratase (itaconyl-CoA hydratase) and later, itaconyl-CoA is cleaved to end products acetyl-CoA and pyruvate. Itaconyl-CoA hydratase (Ich) is part of this catabolic pathway and is required for the conversion of itaconyl-CoA to (*S*)-citramalyl-CoA which is further catabolized to acetate and pyruvate. Structural and functional analysis of two enzymes, Ict and Ccl, have been studied so far, but any idea of the Ich, nexus of those two enzymes has remained obscure. In this way, *P. aeruginosa* circumvents the host defense mechanisms which ultimately leads to uncontrolled proliferation and sepsis in the human body. The role of itaconate as a host defense system and how this host defense system is circumvented by specific itaconate degradation pathways of some pathogenic bacteria, such as *Pseudomonas,* is the major primary focus of this thesis. Please refer to the review of literature section for a more thorough analysis of the itaconate breakdown pathway.

Pseudomonas aeruginosa coordinates the expression of global genes to adapt to the host environment and fend off immunological responses. Numerous regulatory genes are present in *P. aeruginosa* and are crucial for regulating the expression of certain virulent genes [54]. Previous work has shown that PC is one of main nutritional sources during lung infection of *P. aeruginosa*, enabling the high-cell-density (HCD) development of bacteria in humans [55]. PC is broken down by *P. aeruginosa*-secreted lipases and phospholipase C (PlcH), which produces phosphorylcholine, glycerol, and long-chain fatty acids (mostly palmitic and oleic acids). So, many transcriptional regulatory factors stimulate the expression of certain types of virulent genes such as *plcH, fadD1, fadD6, glcB,* and *maeB* which actively participate during infection in host, especially in CF patients.

1.5. Overview of Thesis

The preceding introduction section gives a background on several types of virulence factors secreted by *Pseudomonas aeruginosa* and their role in virulence. Among them, the role of itaconate catabolic pathway in the degradation of itaconate secreted by macrophages in subverting host immune responses is the primary focus of this thesis. In addition, this thesis also includes the role of virulence-associated transcriptional factors controlling the expression of certain virulent genes in P. aeruginosa.

Chapter 2 describes the contribution of itaconate secreted by macrophages as a host immune response during *Pseudomonas* infection, and how *Pseudomonas* subverts this host immune response through catabolism of itaconate is thoroughly explained in this Review of Literature section. Moreover, structural and functional analysis of two enzymes, Ict and Ccl, of this pathway and especially the 'hotdog fold' characteristic of Ich, have been described in this section. In addition, a brief description of some virulence-associated TetR family transcriptional regulators controlling the expression of virulent genes is also given.

Chapter 3 (General Methodology section) represents the general methodology used in this study.

Chapter 4 represents the first crystal structure of Itaconyl-CoA hydratase (Ich) from *P. aeruginosa* of the itaconate catabolic pathway. In-depth structural and functional insights from Ich are explained in this section. Various biophysical methods like mass spectrometry, size-exclusion chromatography and also X-ray crystallography confirm the oligomeric nature of Ich as a dimer which consists of a novel N-terminal 'hotdog fold', not shown previously in any other (*R*)-specific enoyl-CoA hydratases.

Chapter 5 represents the structural analysis of PvrA, a TetR family transcriptional regulator involved in the regulation of the expression of certain virulent genes in *P. aeruginosa* during infection in humans, especially in the lungs of CF patients. Structural and size-exclusion chromatography analysis of PvrA showed that it presents as a dimer where a conserved long N-terminal H-T-H DNA binding region is present.

Chapter 2

Review of Literature

2.1. The biosynthesis and metabolism of itaconate

Mitochondria are referred to as "powerhouses of cells" due to the significant role they play in sustaining calcium homeostasis and supplying metabolic energy through the body. The mitochondrial matrix produces adenosine triphosphate (ATP) through an intricate biological process called the TCA cycle, which is mediated by a sequence of enzyme-based processes [56]. The decarboxylation of cis-aconitate generated by the dehydration of citrate is a unique and significant step in the biosynthesis pathway of itaconate. Aconitate decarboxylase 1 (ACOD1), also known as Irg1, is the enzyme responsible for the decarboxylation of cis-aconitate to itaconate [57]. Endogenous itaconate production is enhanced by Irg1 overexpression [58]. Accumulation of citrate (itaconate precursor) is further caused by the catalytic conversion of isocitrate to a-ketoglutarate dehydrogenase [59]. Furthermore, the process of converting pyruvate into the citrate precursor acetyl-CoA is irreversibly facilitated and regulated by the pyruvate dehydrogenase complex (PDC) [57]. Pyruvate dehydrogenase kinase 1 (PDK1) phosphorylates PDC, inactivating the enzyme and reversing this conversion [60, 61]. Itaconate can be catabolized into itaconyl-coenzyme A (CoA), which deactivates mitochondrial CoA B_{12}. This lowers the activity of methylmalonyl-CoA mutase (MUT) and produces branched-chain amino acids (BCAAs) that are dependent on MUT [62, 63]. Beyond the TCA cycle, itaconate and its metabolites can influence other pathways involved in energy metabolism. The biosynthesis and metabolism of itaconate are shown in **Figure 2.1**.

2.2. General fetaures of itaconate

Itaconate is a dicarboxylic acid (methylenesuccinate) consisting of two carboxyl groups and a double bond. It is an electrophile that possesses the active chemical capacity to bind to a nucleophile by taking up two electrons, exhibiting properties similar to the fundamental processes of enzymes. The conjugated unsaturated double bond structure of itaconate allows it to covalently modify the cysteine residues in proteins through Michael's addition reaction. Subsequently, it can alter the activity and function of substrate proteins and have a significant inhibitory effect on the inflammatory signaling pathway [64, 65].

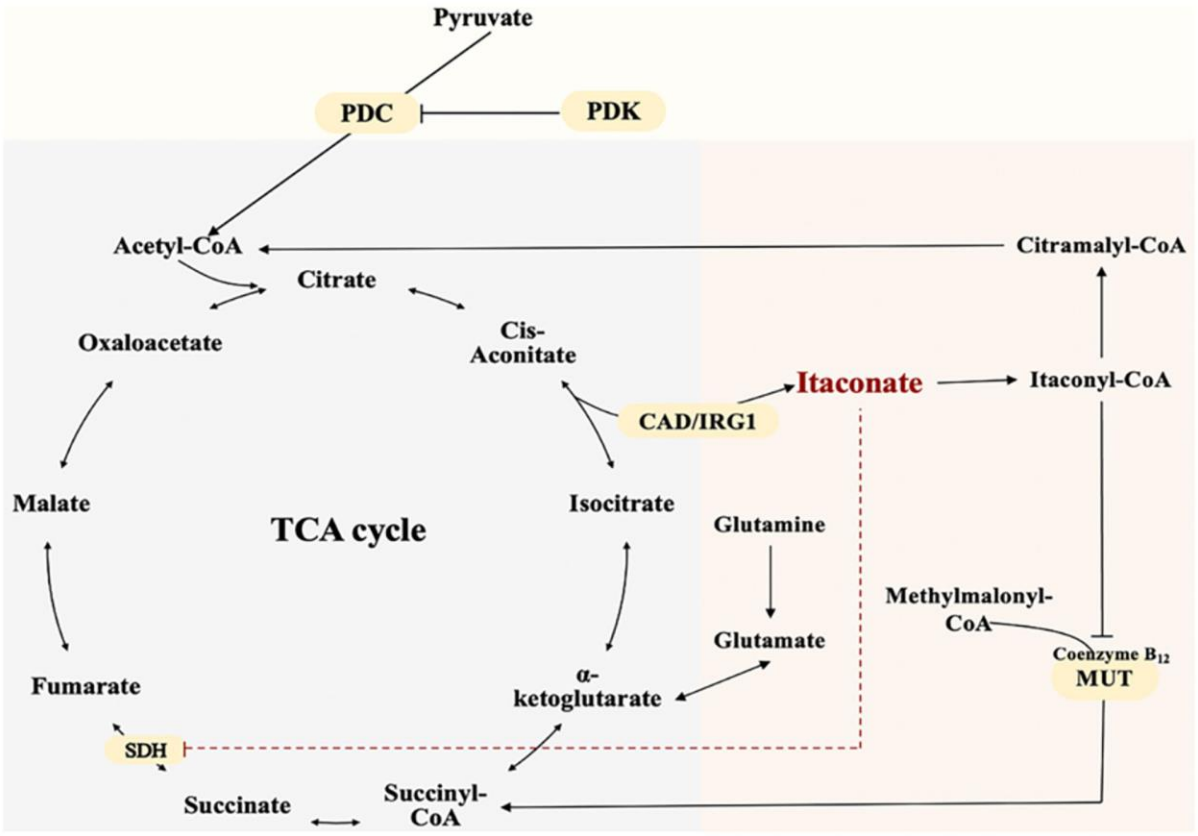


Figure 2.1: Itaconate biosynthesis and its metabolism. Aconitate decarboxylase 1 encodes cis-aconitate, which is then decarboxylated to create itaconate. SDH is inhibited by itaconate and succinate builds up. Acetyl-CoA is produced from pyruvate by the catalytic conversion of pyruvate dehydrogenase complex (PDC). The conversion of itaconate to itaconyl-CoA inhibits methylmalonyl-CoA mutase and methylmalonyl-CoA conversion by deactivating mitochondrial CoA B_{12}. Adapted from Lin, J., Ren, J., Gao, D. S., Dai, Y., & Yu, L. (2021). The emerging application of itaconate: promising molecular targets and therapeutic opportunities. *Frontiers in Chemistry, 9, 669308.*

To support further investigation, Wang et al. designed the biorthogonal probe Itaconate-alkyne (ITalk) in 2020 to gain a better understanding of the true characteristics of itaconate. ITalk retains the long carbon chain and the α, β-unsaturated carboxylic acid group [66]. It was created to directly capture altered proteins in living cells to locate legitimate targets for itaconate on a broad scale. Many important proteins including those linked to cell death, signal transduction, transcription, and the inflammasome, have undergone significant modifications. These proteins are crucial in the regulation of the inflammatory immune

response and host defense. As a result, itaconate may influence macrophage function by controlling multiple pathways **(Fig. 2.2)**.

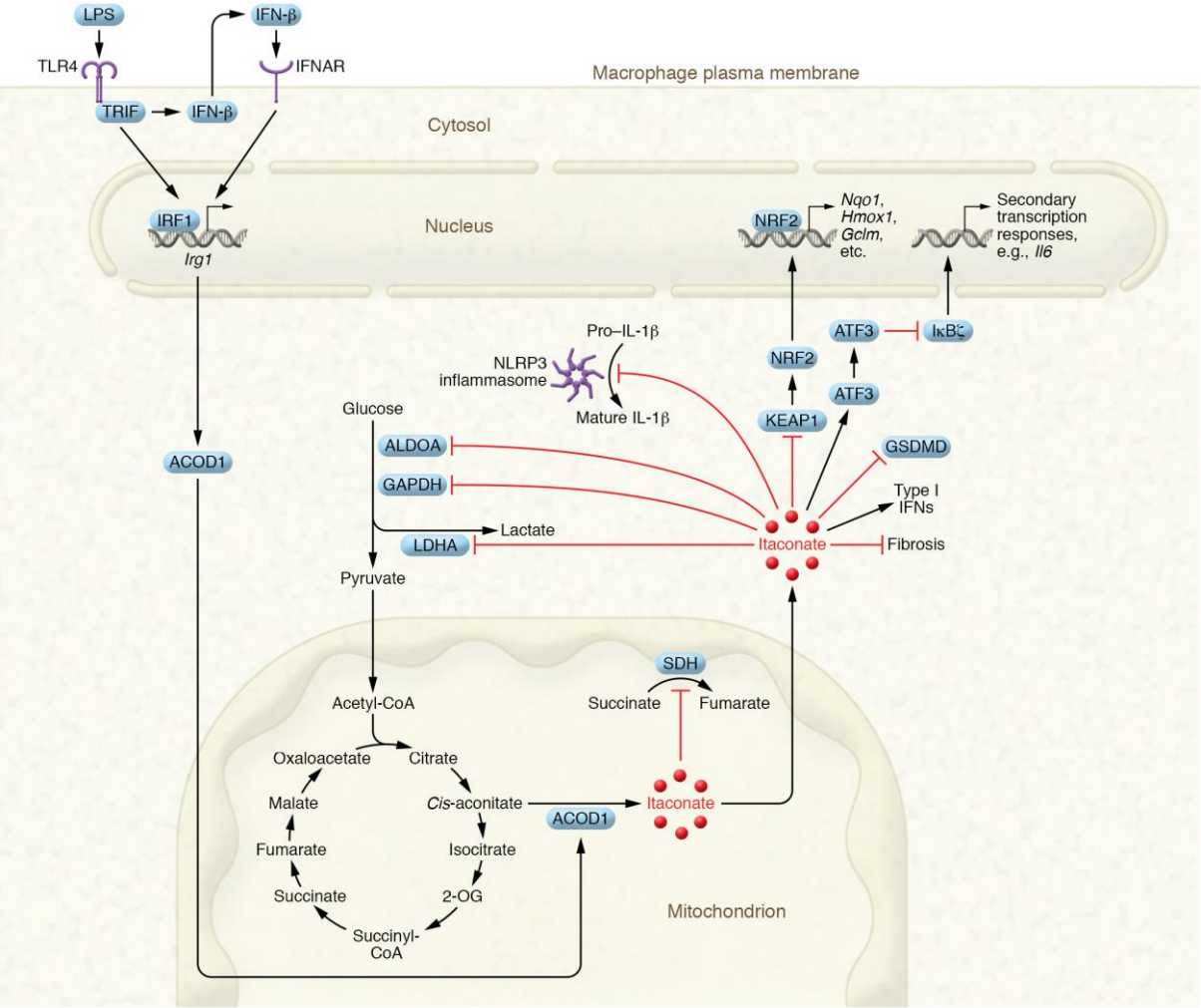


Fig. 2.2: Immunomodulatory properties of itaconate. Irg1 is induced by LPS, leading to production of itaconate. Itaconate targets succinate dehydrogenase (SDH), lactate dehydrogenase A (LDHA), glycolytic enzymes aldolase A (ALDOA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and the NLRP3 inflammasome, which will prevent generation of IL-1β, IL-18, and gasdermin D (GSDMD). Kelch-like ECH-associated protein 1 (KEAP 1) and activating transcription factor 3 (ATF3) were also possible targets. Adapted from Peace, C. G., & O'Neill, L. A. (2022). The role of itaconate in host defense and inflammation. *The Journal of clinical investigation, 132(2).*

2.3. Antibacterial function of itaconate

The antibacterial properties of itaconate were identified about fifty years ago. *Pseudomonas indigofera* depends on isocitrate lyase (ICL) activity and its growth was inhibited by itaconate. ICL is an essential enzyme for glyoxylate shunting during bacterial infections [67]

[68, 69]. Subsequently, numerous investigations demonstrated that itaconate was capable of efficiently preventing the growth of numerous bacteria, including *Legionella pneumophila*, *Mycobacterium tuberculosis*, and *Salmonella* [57, 70-72]. Significant studies have recently shown that itaconate regulates macrophage antibacterial immunity through new pathways [73-75]. Rab32 is a GTPase that coordinates host defensive mechanisms within cells and inhibits the growth of intracellular pathogens like *Salmonella*. BLOC3 is an exchange factor that is necessary for Rab32 to function [76]. Chen et al. found that the production of itaconate by IRG1 was necessary for this antibacterial mechanism to function [73]. During *Salmonella* infection, organelle interactions mediated by the IRG1-Rab32-BLOC3 pathway may transfer itaconate from mitochondria into *Salmonella*-containing vacuoles, exposing the pathogen to elevated itaconate levels and suppressing *Salmonella* development. The researchers also employed a mutant strain of Salmonella with virulence defects that showed decreased capacity to proliferate in wild-type macrophages and was unable to resist the Rab32 defensive mechanism. $IRG1^{-/-}$ macrophages were able to recover the replication-deficient phenotype, which allowed the *Salmonella* mutant to replicate to nearly wild-type bacterial levels [73]. These results demonstrate the role that itaconate plays in the management of bacterial infections and provide evidence of a connection between itaconate and the Rab32 pathway, a cell-autonomous defensive mechanism.

Additionally, Schuster et al. discovered that TFEB (Transcription Factor EB), a transcription factor, regulates the interaction between phagolysosomes and mitochondria in macrophages [74]. Itaconate synthesis and IRG1 transcription were enhanced by TFEB activation in response to bacterial stimulation. Subsequently, the IRG1-Rab32-BLOC3 system was used to transport TFEB-driven itaconate into *Salmonella*-containing vacuoles, thereby limiting *Salmonella* growth. Remarkably, it was discovered that itaconate stimulated TFEB to increase lysosome biogenesis [75]. According to Zhang et al., itaconate alkylated human TFEB at Cys_{212} (Cys_{270} in mice), which inhibited the phosphorylation of Ser_{211} by protein kinase mTOR and interfered with the interaction of phosphorylated TFEB and 14-3-3 regulatory proteins. This led to the translocation of TFEB from the cytosol to the nucleus, which triggered lysosomal biogenesis and enhanced macrophage defense against bacterial invasion. These findings point to a potential interaction between itaconate and TFEB that needs to be investigated further. Itaconate may have antibacterial therapeutic benefits because of its combined role in facilitating the cooperation of mitochondria and lysosomes to resist infections in macrophages.

2.3.1. Itaconate as an inhibitor of isocitrate lyase (Icl)

An important second line of defense against microbial infections is composed of phagocytic cells such as neutrophils and macrophages. At the site of infection, these cells release their secretory products and engulf the invasive cells to create the phagosome and subsequently the phagolysosome, an organelle that typically kills and digests bacteria. Macrophages release more than 100 different chemical substances that have a range of biological actions, from stimulating cell proliferation to causing cell death [77]. Reactive oxygen and nitrogen species causing non-specific damage to DNA and other biological components are the primary antimicrobial agents produced by macrophages, but they also produce some products that block specific processes in pathogens. As mentioned earlier, macrophages create C5-dicarboxylic acid itaconate, also known as methylenesuccinate and in response to lipopolysaccharide and interferon γ, the amount of this compound in cells and culture supernatants significantly rises [57, 78].

Itaconate is a strong inhibitor of isocitrate lyase, an essential glyoxylate cycle enzyme that many pathogenic or non-pathogenic bacteria employ to assimilate acetyl-CoA generated during fatty acid breakdown **(Figure 2.3)** [67, 79, 80]. This is why itaconate secretion is thought to be a component of antibacterial response of macrophages. Plants, bacteria, protists, and fungi all contain the glyoxylate cycle, an anabolic process that is a variant of the tricarboxylic acid cycle. The synthesis of carbohydrates requires the conversion of acetyl-CoA to succinate, which is the main function of the glyoxylate cycle. In the absence of simple carbohydrates like glucose or fructose, the glyoxylate cycle of bacteria replenishes the carbon requirement by using two carbon molecules such as acetate.

Studies showed that itaconate inhibits various Icl isoforms **(Table 2.1)**. Specifically, itaconate inhibits multiple bacterial Icl isoforms: *Mycobacterium tuberculosis* isoforms (Icl 1, K_i = 120 μM and Icl 2 (also known as AceA), K_i = 220 μM), *Pseudomonas indigofera* (K_i = 0.9 μM), and *Corynebacterium glutamicum* (K_i = 5.05 μM) [69] [80]. Icl in association with inhibitors such as bromopyruvate (PDB ID: 1F8M) and 3-nitropropionate (PDB ID: 1F8I, 6C4A, and 6C4C) were described in several crystal structures. However, Icl in complex with itaconate is the only structure reported till now (PDB ID: 6XPP).

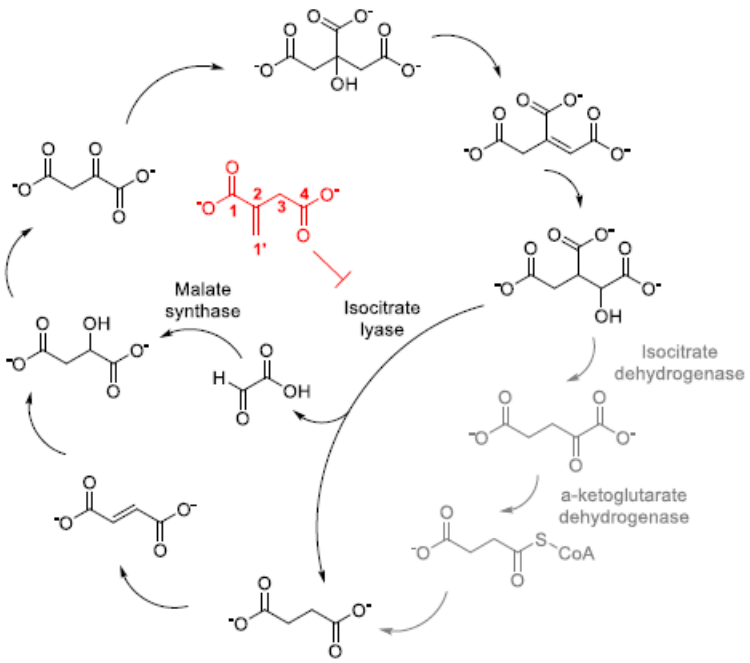


Figure 2.3: Intersection of TCA and glyoxylate cycles. Isocitrate dehydrogenase and α-ketoglutarate dehydrogenase of TCA cycle (full circle) are both bypassed by glyoxylate cycle (in black). Isocitrate lyase is inhibited by itaconate (in red). Adapted from Duncan, D., & Auclair, K. (2022). Itaconate: an antimicrobial metabolite of macrophages. *Canadian Journal of Chemistry, 100(2), 104-113.*

A Michael-addition between a conserved Cys_{191} and itaconate results in the development of a covalent adduct between the enzyme and the inhibitor in this complex of *M. tuberculosis* Icl 1 [81-83]. Despite being a stronger Michael acceptor, the dimethyl ester of itaconate demonstrates 20–25 times less inhibition for *M. tuberculosis* Icl 1 [83]. This implies that key residues involved in interactions between carboxylate group of itaconate and the active site of Icl are Arg_{228}, Asn_{313}, Ser_{315}, Ser_{317}, and Thr_{347}. Moreover, the inhibitory effect is significantly diminished by substituents added at C-1 of itaconate, most likely due to steric hindrance **(Figure 2.4)**. It was found that the formation of itaconate-bound Icl 1 complex needed Mg^{2+} and the formation of all covalent bonds took about five hours [83]. It is perhaps predictable that the duration of complex formation is slow since α,β-unsaturated carboxylates are very weak Michael acceptors.

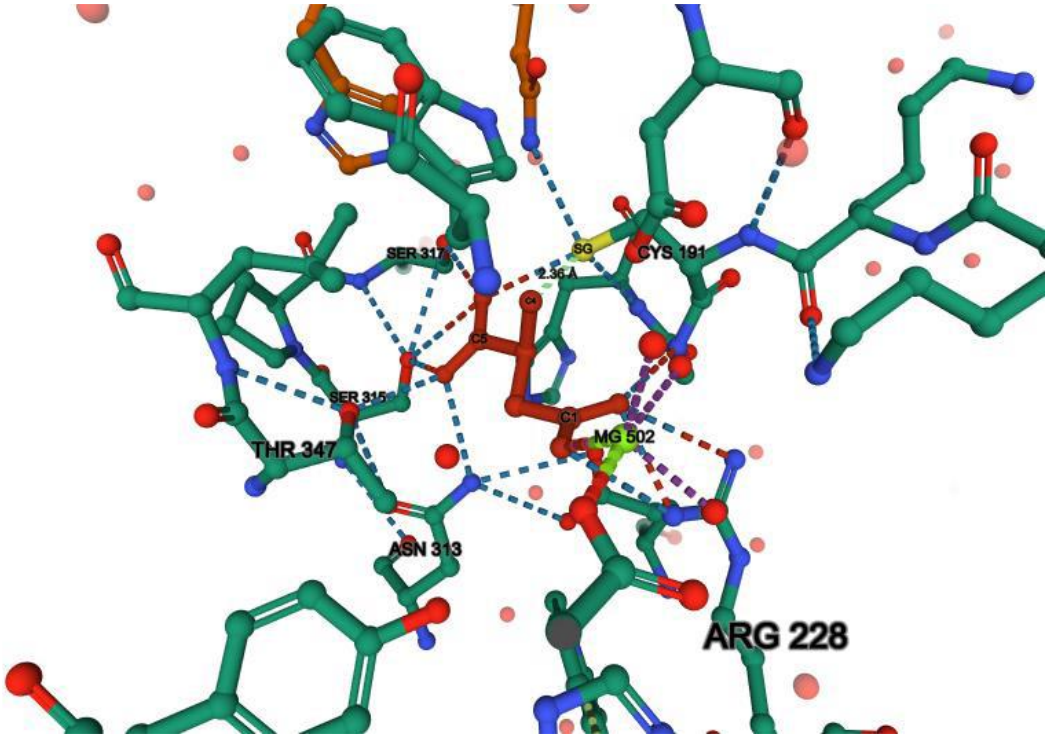


Figure 2.4: *Mt*Icl 1 bound with itaconate (red) (PDB ID: 6XPP). An adduct between the inhibitor and the enzyme is formed when C4 of itaconate covalently links to Cys_{191} of the enzyme (shown in red). C1 of itaconate interacts with a Mg^{2+} ion in addition to forming hydrogen bonds with Arg_{228} and the backbone amide of Cys_{191}. Adapted from Duncan, D., & Auclair, K. (2022). Itaconate: an antimicrobial metabolite of macrophages. *Canadian Journal of Chemistry*, *100(2), 104-113.*

Additionally, the C191S mutant of Icl 1 was found to have a similar affinity for itaconate (K_d for wild type and mutant = 112 ± 11 μM and 155 ± 29 μM, respectively), which does not create an adduct, suggesting that the covalent bond is not the main cause of inhibition. On the other hand, Cys_{191} is necessary for the significant inhibition of 3-nitropropionate and bromopyruvate to occur [81, 82].

Immunoresponsive gene 1 (Irg1) is highly expressed in mammalian macrophages during inflammation and its product, itaconate, is produced by decarboxylating cis-aconitate, an intermediate of the Krebs cycle. According to gene expression profiling studies of murine macrophages and microglial cells, Immunoresponsive gene 1 (Irg1) is one of the most strongly up-regulated genes in proinflammatory circumstances, such as bacterial infections [84-86].

Table 2.1: Several Icl isoforms showing inhibition constants towards itaconate	
Isoform	**K_i (µM)**
Bacteria	
Corynebacterium glutamicum	5.05
Mycobacterium tuberculosis	Icl 1: 120, AceA/Icl 2: 220
Pseudomonas indigofera	0.9
Eukaryota	
Caenorhabditis elegans	19
Tetrahymena pyriformis	3.5
Aspergillus nidulans	40
Ricinus communis L. cv. Zanzibariensis	11.9
Ashbya gossypii	170
Linum usitatissimum L.	17
Pinus densiflora Sieb et Zucc	2.8
Leishmania amazonensis	4500
Ascaris suum	7.3
Fomitopsis palustris	68
Reference: Taken from Duncan, D., & Auclair, K. (2022). Itaconate: an antimicrobial metabolite of macrophages. *Canadian Journal of Chemistry, 100(2), 104-113.*	

IRG1 was classified into the MmgE/PrpD family based on sequence homology; this family included several proteins for which enzymatic functions in microbes were discovered. So, IRG1 functions as an enzyme that catalyzes the decarboxylation of cis-aconitate to produce the antimicrobial metabolite itaconic acid.

2.4. Itaconate catabolism by bacteria

Itaconate can be metabolized by mammals and some microorganisms. In the 1960s, the breakdown mechanism of itaconate was identified for *Salmonella* spp., *Micrococcus* spp., *Pseudomonas* spp., and mammalian mitochondria [87-90]. Succinyl-CoA:itaconate CoA transferase (Ict) first activates itaconate to its corresponding CoA ester. Itaconyl-CoA hydratase (Ich) then hydrates itaconyl-CoA to (*S*)-citramalyl-CoA and (*S*)-citramalyl-CoA lyase (Ccl) cleaves it into acetyl-CoA and pyruvate **(Figure 2.5)**.

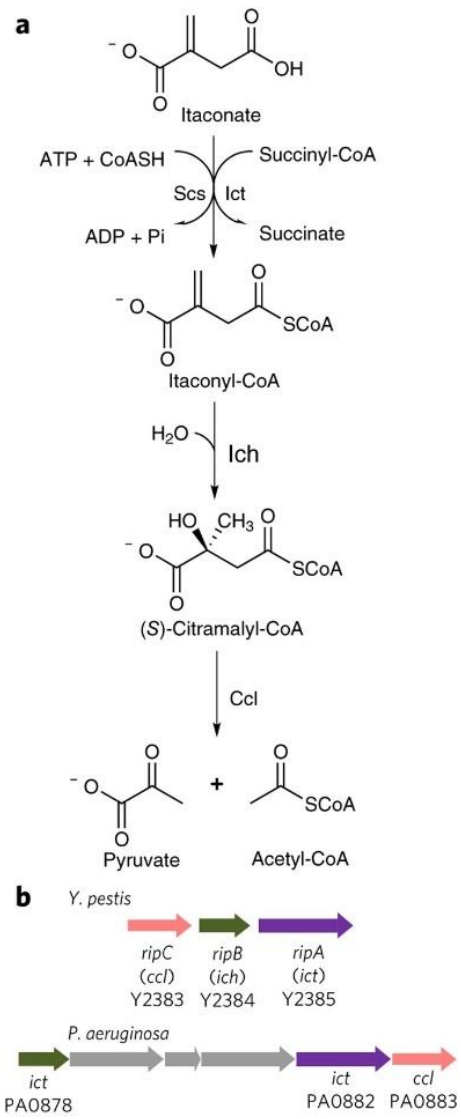


Figure 2.5: Pathway of itaconate catabolism and its corresponding gene clusters in *P. aeruginosa* and *Y. pestis*. (a) pathway for itaconate catabolism in bacteria and mammalian mitochondria. Two mechanisms for itaconate activation are depicted: One involves CoA and ATP via the nonspecific action of succinyl-CoA synthetase (observed in mammalian mitochondria), and the other requires succinyl-CoA via Ict (observed in *Pseudomonas* spp. and most bacteria). (b) Clusters of genetic constituents encoding enzymes for itaconate catabolism in *P. aeruginosa* and *Y. pestis*. Adapted from Sasikaran, J., Ziemski, M., Zadora, P. K., Fleig, A., & Berg, I. A. (2014). Bacterial itaconate degradation promotes pathogenicity. *Nature chemical biology, 10(5), 371-377.*

Yersinia pestis contains a "three-gene" operon called *ripABC* (required for intracellular proliferation) which is essential for growth in activated macrophages [91]. In post-activated macrophages, mutants containing deletions in either *ripA* or *ripB* cannot proliferate. Gene clusters linked to virulence contain the rip operon and adjacent operons. Together with genes encoding the multifunctional, virulence-associated metal-binding siderophore yersiniabactin,

these operons are located at the pigmentation locus of *Yersinia pestis*. Following an attack by macrophages, its transcription is considerably increased and proteomic analysis showed that natural resistance-associated macrophage protein 1 particularly modulates RipA and RipB protein levels, which are then dramatically raised after macrophage involvement [92, 93]. Furthermore, mutations in *ripB* were reported to decrease *S. enterica* var. *Typhimurium* infection in a mouse model while *S. enterica ripA* (previously identified as *cat2*) mutants are impaired in their capacity to infiltrate and survive in chicken macrophages [94-96]. The three enzymes involved in the metabolism of itaconate are encoded by the rip operon: a class III CoA transferase referred to as Ict (RipA encoded by *y2385*), an (*R*)-specific enoyl-CoA hydratase referred to as Ich (RipB encoded by *y2384*) and a C-C-lyase referred to as Ccl (RipC encoded by *y2383*). Furthermore, it was shown that RipA catalyzes the butyryl-CoA:acetate CoA transferase reaction and the observed K_m values were extremely high (above 100 mM), which strongly suggests that RipA serves another physiological role [97]. Structural investigation of Ccl from *Y. pestis* (RipC) revealed that it is a potential protein which only binds CoA or CoA derivatives. Furthermore, structural and computational studies suggested that a homologous protein from *Mycobacterium tuberculosis* (CitE; 36% identity, 51% similarity) functions as a citryl-CoA lyase [98]. The connection between the *M. tuberculosis* gene and citramalyl-CoA is compelling, even though the context of gene points to a function via a separate route.

2.4.1. Enzymes encoded by the *Y. pestis rip* operon

To investigate the hypothesis that RipA, RipB, and RipC are involved in the itaconate catabolic pathway, *ict* (*ripA*), *ich* (*ripB*) and *ccl* (*ripC*) genes from *Y. pestis* were synthesized and heterologously expressed in *Escherichia coli* and purified the recombinant proteins. With these three enzymes, the itaconate degradation pathway was fully reconstructed in vitro and RipA, RipB and RipC were assigned to have their proposed functions as Ict (indicated as *Y. pestis* Ict (*Yp*Ict)), Ich (*Yp*Ich) and Ccl (*Yp*Ccl) respectively. The first enzyme in this pathway is *Yp*Ict which does not require any divalent metal ions for activity. Itaconate is the actual physiological substrate of *Yp*Ict but it also showed poor catalytic efficiencies on some other dicarboxylic acids. In addition to succinyl-CoA, other CoA esters that could be utilized as CoA donors include acetyl-CoA, propionyl-CoA, and butyryl-CoA. It may be crucial to be able to activate itaconate using a variety of CoA donors to ensure that harmful itaconate is quickly trapped as itaconyl-CoA once it enters the cell.

*Yp*Ich, the second enzyme of the pathway, accomplished two fractional activities: itaconyl-CoA isomerase producing mesaconyl-C4-CoA and mesaconyl-CoA hydratase producing the product, (*S*)-citramalyl-CoA. It converts both substrates, itaconyl-CoA and mesaconyl-CoA, into the product (*S*)-citramalyl-CoA **(Table 2.2)**. Studies showed that Ich belongs to (*R*)-specific enoyl-CoA hydratases consisting of the so-called **'hotdog fold'** [99]. Itaconyl-CoA undergoes isomerization into mesaconyl-C4-CoA through proton abstraction at C2 and re-addition at C5. The (*S*) stereoisomer of citramalyl-CoA is formed when OH^- is added at position C3 and protonated at position C2 of mesaconyl-CoA. The catalytic activity of *Yp*Ich did not require any divalent metal ions.

Table 2.2: Catalytic characteristics of recombinant Ich from *Y. pestis* and *P. aeruginosa*

	Y. pestis Y2384 (*Yp*Ich)			*P. aeruginosa* PA0878 (*Pa*Ich)		
Substrate	V_{max} (U mg^{-1} protein)	K_m (mM)	K_{cat}/K_m (s^{-1} mM^{-1})	V_{max} (U mg^{-1} protein)	K_m (mM)	K_{cat}/K_m (s^{-1} mM^{-1})
Itaconyl-CoA	8.24 ± 0.88	0.12 ± 0.03	52.2	12.21 ± 0.52	0.034 ± 0.005	415.0
Mesaconyl-C4-CoA	66.36 ± 9.56	0.51 ± 0.24	98.8	1247 ± 125.8	0.31 ± 0.12	4650.8
(*S*)-Citramalyl-CoA	2.80 ± 0.56	0.49 ± 0.23	4.4	50.32 ± 3.56	0.14 ± 0.025	415.4

Reference: Taken from Sasikaran, J., Ziemski, M., Zadora, P. K., Fleig, A., & Berg, I. A. (2014). Bacterial itaconate degradation promotes pathogenicity. *Nature chemical biology, 10(5), 371-377*.

The third enzyme in this pathway, *Yp*Ccl, requires divalent metal ions such as Mg^{2+} and Mn^{2+} to show their catalytic efficiency. Three orders of magnitude slower was the cleavage of (3*S*)-citryl-CoA, the suggested substrate of *M. tuberculosis* CitE. Other β-hydroxy acyl-CoA molecules like (3*S*)-malyl-CoA or β-methylmalyl-CoA, and 3-hydroxy-3-methylglutaryl-CoA were not accepted by *Yp*Ccl. Several phylogenetically related enzymes can also cleave (*S*)-citramalyl-CoA, including (3*S*)-malyl-CoA/β-methylmalyl-CoA/(*S*)-citramalyl-CoA lyase and the β-subunit of citramalate lyase, which resembles the corresponding subunit of the ATP-independent citrate lyase (CitE), albeit much less efficiently than by *Yp*Ccl [100-103].

*Yp*Ict can employ acetyl-CoA, one of the byproducts of itaconate breakdown, to activate itaconate. Therefore, the initial concentration of succinyl-CoA does not affect the amount of itaconate converted in the process. Large amounts of itaconate may be detoxified with as little as a catalytic amount of a CoA ester.

Crystal structures of *Y. pestis* Ict were reported, while it was described as a 4-hydroxybutryate CoA-transferase (PDB ID: 4N8K, 4N8H, 4N8J, 4N8L, 4N8I) and coenzyme A transferase (PDB ID: 3S8D, 3QLK, 3QLI) in the PDB [97, 104]. It was proposed that the reaction occurs by two double displacements (two ping-pong reactions) **(Figure 2.6)**. To create an acetyl anhydride, the acetyl group of AcCoA is first transferred to the catalytic Glu_{429} residue of Ict. The thiol of CoA then attacks this acetyl anhydride nucleophilically to produce an enzyme-CoA adduct. Following that, one of the carboxylate groups of the itaconate attacks this thioester, resulting in an itaconate enzyme adduct. Eventually, a nucleophile attacks the thiol of CoA, producing itaconyl-CoA and regenerating Glu_{429} [104].

Crystal structures of Ccl were reported from different pathogenic species such as *M. tuberculosis* Rv4589c (annotated as CitE; PDB ID: 1U5H, 6AQ4), *Homo sapiens* CLYBL (PDB ID: 5VXS, 5VXC, 5VXO), *Y. pestis* RipC (PDB ID: 3QLL), as well as the non-pathogenic bacteria *Ralstonia eutropha* (PDB ID: 3QQW), *Burkholderia xenovorans* (PDB ID: 3R4I) and *Deinococcus radiodurans* (PDB ID: ISG1) [98, 105-107].

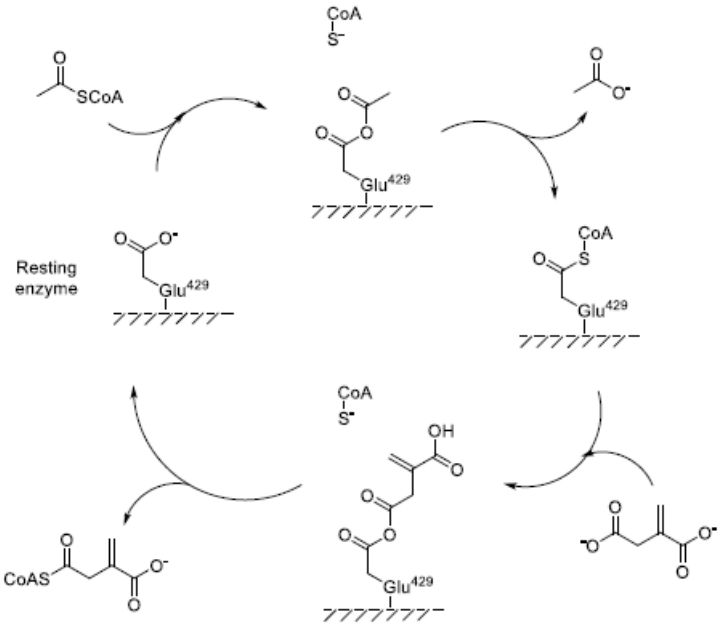


Figure 2.6: Proposed mechanism of *Y. pestis* Ict. Adapted from Duncan, D., & Auclair, K. (2022). Itaconate: an antimicrobial metabolite of macrophages. *Canadian Journal of Chemistry, 100(2), 104-113.*

Because of their structural similarity, these enzymes form a homotrimer with an α8β8-TIM barrel fold [98, 106, 107]. Glu_{39}, Asp_{40}, Arg_{71}, Glu_{129}, Asp_{156}, and Pro_{192} are among the probable active site residues of *Yp*Ccl isoform that have a substantial amount in common with the similar enzyme *Haloferax volcanii* malate synthase H.

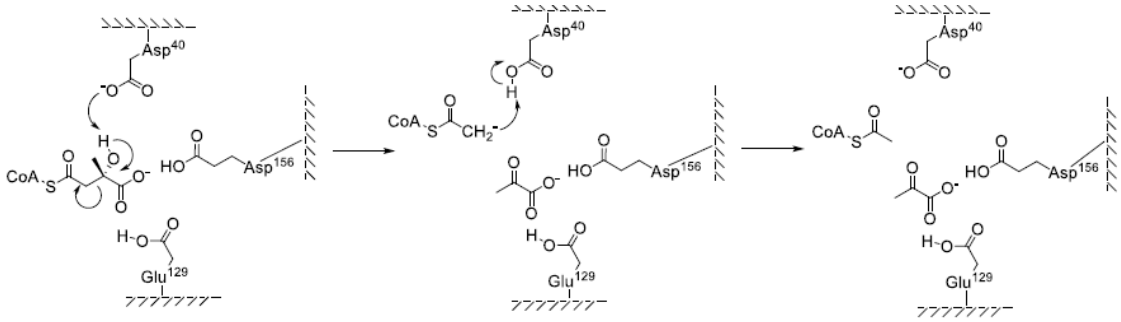


Figure 2.7: Proposed mechanism of *Y. pestis* Ccl. Adapted from Duncan, D., & Auclair, K. (2022). Itaconate: an antimicrobial metabolite of macrophages. *Canadian Journal of Chemistry, 100(2), 104-113.*

Torres et al. proposed a mechanism where Asp_{40} was situated near the hydroxyl group of (*S*)-citramalyl-CoA and may aid its deprotonation, whereas Glu_{129} and Asp_{156} may be placed near the carboxylate group, likely positioning the molecule [97] **(Figure 2.7)**.

2.4.2. Enzymes of itaconate metabolism in *P. aeruginosa*

There is only one *Yp*Ccl homolog (PA0883, also known as *Pa*Ccl; 35%/50% identity/similarity) in the entire genome of *P. aeruginosa*, even though some *Pseudomonas* species are capable of metabolizing itaconate. In contrast, *Yp*Ict and *Yp*Ich homologs could not be found using a typical BLASTp search. Studies showed that the growth of *P. aeruginosa* utilizing itaconate as a sole substrate exhibited significant activity of all three of the enzymes which further metabolize itaconate in the cell extract. These enzymes were downregulated while bacteria were grown on succinate. Itaconate could not be the sole carbon source for a *Pa*Ccl mutant [108]. To examine the activity of enzymes of this mutant involved in itaconate degradation pathway, the mutant was grown on succinate in the presence of itaconate. Studies showed that activity of Ict and Ich under these conditions was detected but not of Ccl, thus confirming that *Pa*Ccl is the supposed Ccl [53]. The proposed "six-gene operon" encoding *Pa*Ccl includes the genes for a putative (*R*)-specific enoyl-CoA hydratase (PA0878) and a class III CoA transferase (PA0882). **(Figure 2.5b)**. These proteins,

(*R*)-specific enoyl-CoA hydratase (PA0878) and class III CoA transferase (PA0882) are distinct from those of *Y. pestis* homologues mentioned before. Sasikaran et al. cloned and heterologously expressed the putative *P. aeruginosa ict* (PA0882), *ich* (PA0878) and *ccl* (PA0883) genes in *E. coli* and demonstrated that these three gene products were required for the conversion of itaconate to pyruvate and acetyl-CoA [53]. Three other genes in this putative 'six gene operon' of *P.* aeruginosa encode an MmgE-PrpD family protein (PA0881), a glyoxalase family protein (PA0880) and a putative acyl-CoA dehydrogenase (PA0879).

As a CoA acceptor, *Pa*Ict exhibited strong activity toward itaconate. Moreover, methylsuccinate was activated by the *Pa*Ict, and to some extent, (*S*)-citramalate. The *Pseudomonas* homologue *Pa*Ict uses exclusively succinyl-CoA as a CoA donor, in contrast to *Yersinia* CoA transferase (*Yp*Ict). *Pa*Ict showed no activity with acetyl-CoA, propionyl-CoA or butyryl-CoA. The presence of divalent metal ions did not increase the enzymatic activity of *Pa*Ich. Itaconyl-CoA and mesaconyl-C4-CoA are both converted to (*S*)-citramalyl-CoA by *Pa*Ich **(Table 2.2)**. Hence, *Pa*Ich/*Yp*Ich was characterized as an itaconyl-CoA isomerase/mesaconyl-CoA hydratase. However, *Pa*Ccl exhibited strong substrate specificity to (*S*)-citramalyl-CoA. Moreover, it facilitated the 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) lyase reaction with a noticeably lower catalytic efficiency (k_{cat}/K_m). Both *R* and *S* stereoisomers are present in commercially available HMG-CoA. *Pa*Ccl breaks down half of (*R,S*)-HMG-CoA to acetyl-CoA and acetoacetate when excess (*R,S*)-HMG-CoA is present, suggesting that only one isomer was utilized for catalysis. To further investigate whether *Pa*Ccl performs catalysis with (*R*)-HMG-CoA or with (*S*)-HMG-CoA, liuE encoding *S*-specific HMG-CoA lyase (PA2011) of *P. aeruginosa* was heterologously overproduced in *E. coli* and purified [109]. Most of the HMG-CoA was converted into products by this enzyme. Remarkably, it was shown that spatial position of the hydroxyl group about CoA in the (*S*)-citramalyl-CoA and the (*R*)-HMG-CoA is similar. Though the turnover with this substrate was lower than 0.1% when compared to citramalyl-CoA as the substrate, *Pa*Ccl can catalyze the (3S)-citryl-CoA lyase process in a manner similar to that of *Yp*Ccl. Divalent metal ions (Mg^{2+}, Mn^{2+} or Co^{2+}) were needed for *Pa*Ccl activity.

Unlike *Yp*Ict, *Pa*Ict is unable to utilize acetyl-CoA to activate itaconate. Based on another study, 5 mM itaconate almost disappeared in a reconstitution experiment when substoichiometric levels of succinyl-CoA (1 mM) were present. It can be explained by the fact that *Pa*Ict can catalyze the transfer of CoA from (*S*)-citramalyl-CoA to succinate while simultaneously producing citramalate, so when Ccl was absent, itaconate disappearance still

happened. Citramalate is converted back to citramalyl-CoA by adding Ccl, and subsequently to acetyl-CoA and pyruvate. The ratio of produced pyruvate/acetyl-CoA to the ingested itaconate/citramalate was 1:1. Therefore, *P. aeruginosa* employs a different yet effective method for quick itaconate breakdown and detoxification.

2.5. Hotdog fold of (*R*)-specific enoyl-CoA hydratases

The 'Hotdog' fold was initially discovered in the crystal structure of β-hydroxydecanoyl thiolester dehydratase (FabA) in *E. coli* [110]. In this structure, the hotdog fold is created by a long core helix that resembles a sausage wrapped by a highly curved six-stranded β-sheet resembling a bun **(Figure 2.8)**. This hotdog fold was found to be present in several protein structures. These proteins participate in multiple processes such as the dehydration step of fatty acid elongation, thioester hydrolysis, transcription regulation in fatty acid metabolism and degradation of phenylacetic acid and environmental pollutants. The hotdog fold appears to be intended for the binding of coenzyme A in cellular activities involving fatty acids and related compounds.

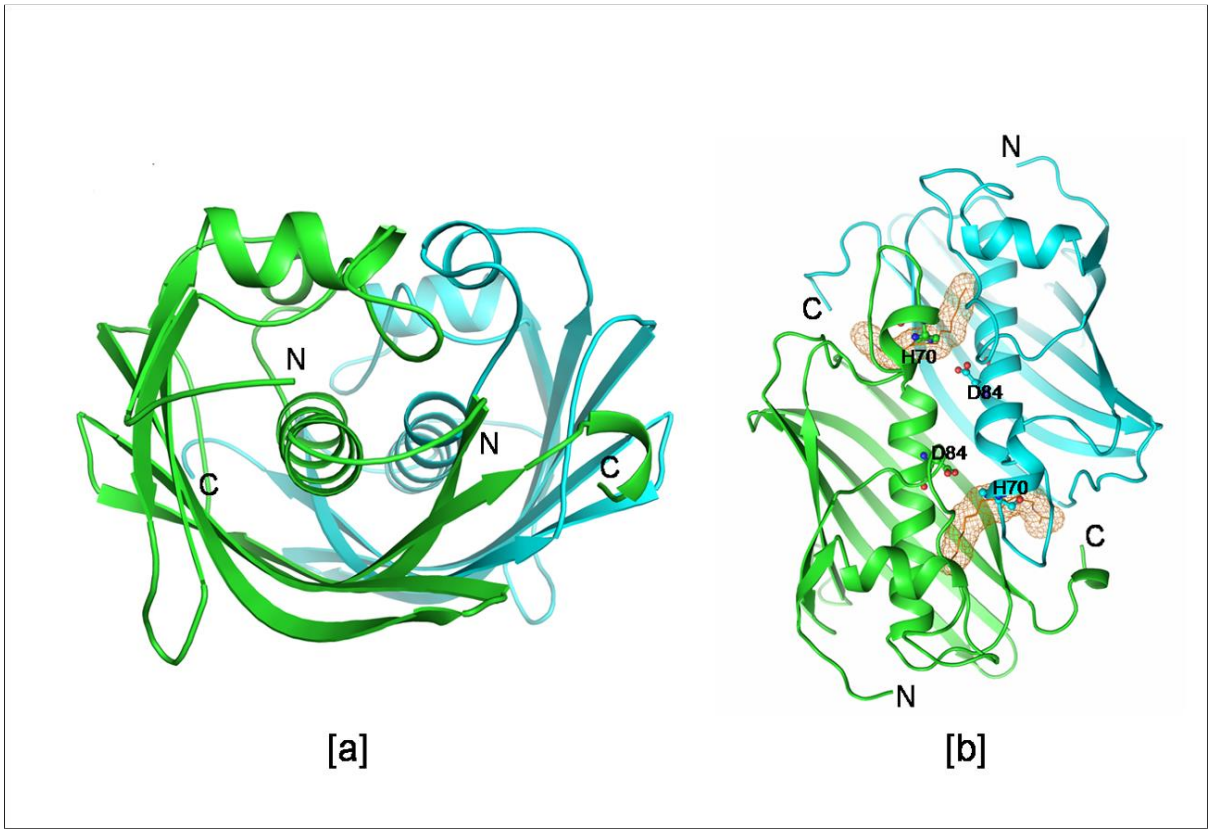


Figure 2.8: Dimer of FabA from *E. coli*. (a) top view of hotdog fold (chain A: green; chain B: cyan). (b) side view of hotdog fold showing inhibitor 3-decanoyl-NAC bound in active site, mesh representation). Catalytic residues are in ball and stick representation. Adapted from Pidugu, L. S.,

Maity, K., Ramaswamy, K., Surolia, N., & Suguna, K. (2009). Analysis of proteins with the 'hotdog' fold: Prediction of function and identification of catalytic residues of hypothetical proteins. *BMC structural biology, 9, 1-16.*

The binding site of FabA showed a well-formed deep tunnel present at the subunit interface. The tunnel is formed by residues contributed by both subunits and is hydrophobic. His_{70} and Asp_{84} are two catalytic residues lining this tunnel. A homodimer known as FabA consists of two tunnels related by a two-fold symmetry. The substrate binding site is situated along the surface of the protein. Based on sequence clustering and accessible structural and biochemical data, they divided a vast number of proteins from prokaryotes, archaea, and eukaryotes in this superfamily into 17 subfamilies. For every subfamily, consensus sequence motifs were recognized. Although all of these proteins have the hotdog fold, only 10% to 20% of their sequences are identical across subfamilies. Five or six strongly curled parallel/antiparallel β-strands are wrapped around a long α-helix to form each subunit of a hotdog fold. Based on structural investigation, The fundamental structural unit of the hotdog fold proteins is either a double hotdog comprised of a single polypeptide chain consisting of two hotdog domains or a homodimer of two such subunits. The hotdog family of proteins is classified into different sub-families.

Generally, itaconyl-coA hydratase (Ich) belongs to (*R*)-specific enoyl-CoA hydratases or MaoC dehydratase-like subfamily which is summarized below.

2.5.1. (*R*)-specific enoyl-CoA hydratases or MaoC dehydratase-like subfamily

In *fadB* mutant *E. coli* strains, the *maoC* gene encodes an enoyl-CoA hydratase that contributes to the transfer of (*R*)-3-hydroxyacyl-CoA from the fatty acid oxidation pathway to polyhydroxyalkanoate (PHA) production pathways. It is present with the *maoA* gene in *E. coli* as an operon [111]. It was recognized by its similarity to the (*R*)-specific enoyl-CoA hydratase (PhaJ1) of *P. aeruginosa*. Various bacteria produce PHAs, or polyesters of (*R*)-hydroxyalkanoic acids, as an internal carbon and energy storage material during high carbon availability. Precursors for PHA biosynthesis include enoyl-CoA, 3-ketoacyl-CoA, (*S*)-3-hydroxyacyl-CoA, and intermediates of fatty acid metabolism [112]. Crystal structure of the (*R*)-specific enoyl-CoA hydratase (PhaJ) from *Aeromonas caviae* (PDB ID: 1IQ6), a eukaryotic hydratase 2 from *Candida tropicalis* (PDB ID: 1PN2) and a human enzyme (PDB code: 1S9C) showed that it also contains a hotdog fold/domain **(Figure 2.9)** [99].

The enoyl-CoA hydratase activity of *E. coli* MaoC is most likely attributed to its C-terminal hotdog domain. Moreover, MaoC contains an N-terminal short-chain dehydrogenase domain that, in the presence of NADP as a cofactor, catalyzes the dehydrogenation of various aromatic and aliphatic aldehydes. Among the four human 17β HSDs (17β-hydroxysteroid dehydrogenase), the type 4 17β HSD catalyzes the redox processes at the site C17 of steroid molecules in the last step of androgen and estrogen biosynthesis [113, 114]. The eukaryotic enzyme was discovered to be a crucial component of a peroxisomal multifunctional protein (MFE-1 in mammals and MFE-2 in fungi) [115-118]. The loop corresponding to the catalytic histidine of the *E. coli* dehydratase FabA is longer by approximately 35 residues in these structures. In 1IQ6, an extra segment is known as the 'overhanging segment' that includes the catalytic residues Asp_{31} and His_{36}. The bacterial enzyme assembles as a dimer, with one hotdog fold on each subunit. Two catalytic sites are present in these structures while eukaryotic and human hydratases form dimers where each monomer consists of two hotdog fold/domains and each monomer possesses one catalytic site. It has been proposed that gene duplication gave rise to the double hotdog fold/domain of eukaryotic and human hydratase 2. It was thought that these structures diverged during evolution to make space for large, fatty enoyl-CoAs at the expense of one catalytic site [99, 119, 120].

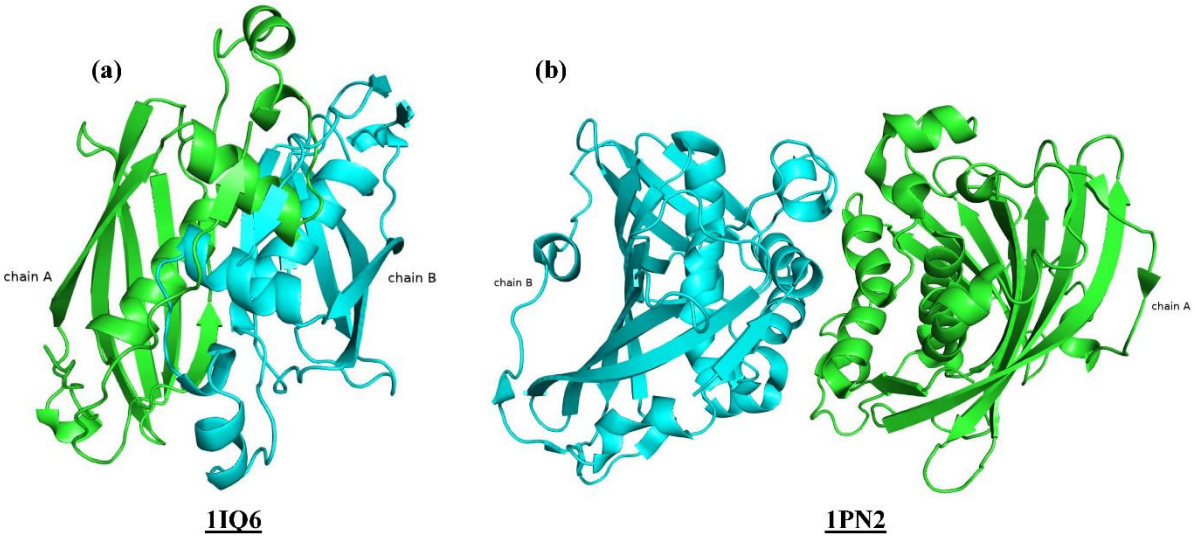


Figure 2.9: Structures of (*R*)-specific enoyl-CoA hydratases from *Aeromonas caviae* (PDB ID: 1IQ6) and human (PDB ID: 1PN2). Structural representations were made through Pymol.

All these proteins showed hydratase activities with different substrate specificities. Depending on the length of the fatty acyl chain of enoyl-CoAs, the activity of different (*R*)-hydratases (PhaJ) such as PhaJ1–PhaJ4 of *P. aeruginosa* varied [121]. Only short-chain enoyl-CoAs (C4–C6) were shown to be active by PhaJ1, but PhaJ2-PhaJ4 can operate on longer enoyl-CoAs (C8–C12).

Crystal structures of other enoyl-CoA (*R*)-hydratases such as FAS I synthase of *Thermomyces lanuginosus* (PDB ID: 2UV9 and 2UVA) and yeast (PDB ID: 2UV8) are reported at 3.1Å resolution [122, 123]. These two enzymes present as α6β6 dodecamers, in contrast to the homodimeric mammalian enzymes. The distinct triple hotdog fold is seen in the dehydratase domain situated on the beta subunit. The first and third hotdog folds form pseudodimers and share structural similarities with the double hotdog fold-containing eukaryotic hydratase 2 enzyme. The second hotdog domain is inserted in the long loop connecting the first and third hotdog domains in the same manner; the domains in the double hotdog fold structures are connected. In contrast to the normal hotdog domain discussed here, the central helix of this domain is substantially shorter and oriented differently.

2.6. Transcription factors controlling virulence in bacteria

DNA-binding transcription factors (TFs) are essential for regulating the pathogenicity and adaptability of pathogenic microorganisms. The genome of *P. aeruginosa* contains 371 putative TFs from 29 families including GntR family (23 TFs), OmpR family (24 TFs), TetR family (23 TFs), Arc family (56 TFs), LysR family (113 TFs), and LuxR family (30 TFs) [124]. Patients with cystic fibrosis (CF) experience a rapid decline in lung function and, consequently, a higher risk of death from infection in their respiratory system [125, 126]. Many studies have been conducted on the pathogenicity of *P. aeruginosa* infection in CF patients, specifically focusing on biofilm development and quorum sensing (QS) regulated virulence [127-129]. However, the contribution of *P. aeruginosa* nutrients that aid in high-cell density (HCD) reproduction during lung infection has received less attention. Prior research demonstrated that *P. aeruginosa* can replicate HCD to a concentration of >109 CFU/ml in the lungs of CF patients [55, 130]. Replication of HCD requires a lot of energy and efficient nutrient acquisition and metabolism. However, data demonstrated that lipids and amino acids produced from proteins or polypeptides are the nutrients in the lung environment that support *P. aeruginosa* HCD growth and maintenance in vivo [131-133]. An in vitro study showed that *P. aeruginosa* displays directed twitching motility to phosphatidylcholine

(PC) and phosphatidylethanolamine (PE) [134]. The essential lung surfactant that covers the lungs of mammals is made up of 90% lipids and 10% protein. In contrast, phosphatidylcholine (PC) accounts for about 80% of lung surfactant lipids [135, 136]. Therefore, PC, the most prevalent lipid in lung surfactants, might significantly contribute to the in vivo proliferation of HCD bacteria.

A sputa study from CF patients showed that *P. aeruginosa* produced lipases and phospholipase C (heat-labile hemolysin) that breakdown exogenous PC into three components: a phosphorylcholine headgroup, two long-chain fatty acids (LCFAs) and glycerol [55]. The choline (bet), fatty acid degradation (Fad), and glycerol (glp) pathways, respectively, can further metabolize these three metabolites. *P. aeruginosa* has a well-characterized metabolism of glycerol and choline [137, 138]. However, further research is needed to completely understand how *P. aeruginosa* breaks down LCFA and the genes involved in this process.

The enzymes responsible for the breakdown of fatty acids in *E. coli* were thoroughly studied. These enzymes are 3-ketoacyl-CoA thiolase (FadA), 3-hydroxyacyl-CoA dehydrogenase (FadB), fatty acyl-CoA synthetase (FadD), an acyl-CoA dehydrogenase (FadE) and an enoyl-CoA hydratase [139, 140]. In response to long-chain acyl-CoAs, the transcriptional regulator FadR, a member of the GntR family, coordinately controls genes in enteric bacteria that are involved in the breakdown and production of fatty acids [141, 142]. Furthermore, FadR plays a role in controlling virulence factors in pathogenic bacteria, including *Salmonella, Vibrio cholerae*, and enterohemorrhagic *E. coli* (EHEC) [143, 144]. Six fadD and five fadBA homologs are present in *P. aeruginosa*, which are identified as fadD 1, 2, 3, 4, 5, 6 and fadBA 1, 2, 3, 4, 5, respectively. Among those genes, *fad*D1, 2, 4, and *fad*BA1, 4, and 5 are important for the fitness of bacteria and the breakdown of fatty acids in a mouse lung infection model.

2.6.1 General features of TetR Family Regulators (TFRs)

One substantial and significant family of single-component signal transduction systems is the TetR family of regulators (TFRs). The sensory and output functions are found on the same polypeptide in one-component systems, whereas they are found on different polypeptides in two-component systems. One-component systems are far more common in prokaryotes than two-component systems [145]. The most well-known function of these family members is to act as regulators of antibiotic efflux pumps. Members that have been characterized are known

to interact with a wide variety of ligands and regulate multiple aspects of bacterial physiology. An N-terminal DNA binding domain and a large C-terminal domain make up all TetR family regulators (TFRs). The proteins act as dimers and are almost entirely helical in shape. The C-terminal domains typically bind with one or several ligands, which further modifies the capacity of the regulator to bind its cognate DNA. Many TFRs function as repressors, just like TetR. Nevertheless, there are significant outliers that function as activators or in other capacities unrelated to transcription. TetR, the protein component that controls and induces tetracycline resistance [146]. In the 1980s, the tetR sequence and many of the molecular features of the control of tetracycline resistance were elucidated [147-150]. We now understand that TetR is the repressor of the tetracycline efflux pump that TetA encodes. When tetracycline is not present, two TetR dimers attach to operator sequences that overlap in the intergenic space between the differently transcribed *tetR* and *tetA* genes. Tetracycline binds to TetR directly when it is present, locking it in a conformation that is unsuitable for DNA binding [151]. This enables *tetR* and *tetA* transcription.

More than 240 TFRs were partially described, and 25% of the TFR family members share their well-established function as regulators of antibiotic efflux [152]. Other TFRs that we are aware of include those that act as global or local regulators (AmtR), repressors and activators (DhaS and PvrA), and interact with protein or small-molecule ligands (SlmA).

Pseudomonas aeruginosa PAO1 and the multidrug-resistant taxonomic outlier PA7 have 36 TFRs in common, and a study of their TFRs revealed that TFRs specific to each strain might be involved in the variations in virulence seen between these two strains. The PAO1 strain of *P. aeruginosa* encodes five TFRs which are absent in PA7 (PA1241, PA1290, PA2020, PA2766, and PA2931), whereas PA7 strain encodes two TFRs that are not present in PAO1 (PSPA7_2630 and PSPA7_4004). The MexXY antibiotic resistance efflux pump is known to be regulated by MexZ encoded by PA2020 [153].

Numerous TFRs bind palindromic DNA operator sequences, which are frequently repeated. However, it is frequently challenging to anticipate operator locations for TFRs with unknown functions. A palindrome may be absent in many instances or exist upstream of genes encoding TFRs or projected targets that do not interact with the corresponding TFR in other circumstances. Occasionally, these might serve as binding locations for additional transcription factors.

2.6.2. TFRs in lipid metabolism

The biosynthesis of fatty acids and polyketide antibiotics shares several similarities. As previously mentioned, TFRs play a role in controlling resistance to a variety of polyketide antibiotics, such as tetracycline. TFRs are also important regulators of the metabolism of sterols and other lipid molecules, including fatty acids. *C. glutamicum* produces the lipid biosynthesis regulator FasR. The expression of 17 genes including *fas*A, *fas*B, *acc*B, *acc*C, and *acc*D1 were varied in a *fas*R mutant. Furthermore, it was shown that the *fasR* mutant had differential expression of two additional TFRs, one of which, Clg1640, might potentially be involved in fatty acid metabolism. Clg1640 is present in *Pseudonocardia autotropica* in a group that also has FadR. FadR regulates an operon involved in the breakdown of fatty acids in *P. autotropica* [154]. Though many TFRs are known to be involved in fatty acid breakdown, FasR is currently the only known TFR engaged in fatty acid production. Long-chain fatty acyl-CoAs are recognized by FadR from *B. subtilis*, which also controls five operons necessary for fatty acid breakdown [155]. Fad35R and Mce3R are two TFRs known to be involved in lipid metabolism in *M. tuberculosis*. When fatty acid derivatives are present, Fad35R regulates the production of an acyl-CoA synthetase encoded by Fad35 [156]. The transcription of the virulence-related mce3 locus and other genes necessary for fatty acid breakdown is repressed by Mce3R. DesT from *P. aeruginosa* and FabR from *E. coli* control distinct pathways that are part of the same general physiological process such as fatty acid saturation **(Fig. 2.10)**. Unsaturated fatty acid synthesis depends on *fabA* and *fabB*, whose expression is controlled by FabR [157]. The genes that encode FabA, FabB, and FabR are dispersed throughout the chromosome, and in contrast to most TFRs, FabR is not autoregulatory [158]. Saturated thioesters were shown to disrupt the FabR-DNA interaction while unsaturated thioesters such as acyl-ACP or acyl-CoA were found to increase binding [158]. Similar ligand binding patterns are seen in DesT, where DNA binding is facilitated by unsaturated acyl-CoAs and inhibited by saturated acyl-CoAs [159]. DesT controls the expression of *desB* and *desC* which are transcribed from *desT* in different directions [159]. One acyl-CoA desaturase and one reductase are encoded by the *desB* and *desC* genes, respectively. While FabR controls the synthesis of unsaturated fatty acids, DesT controls the gene products needed to desaturate preformed acyl chains.

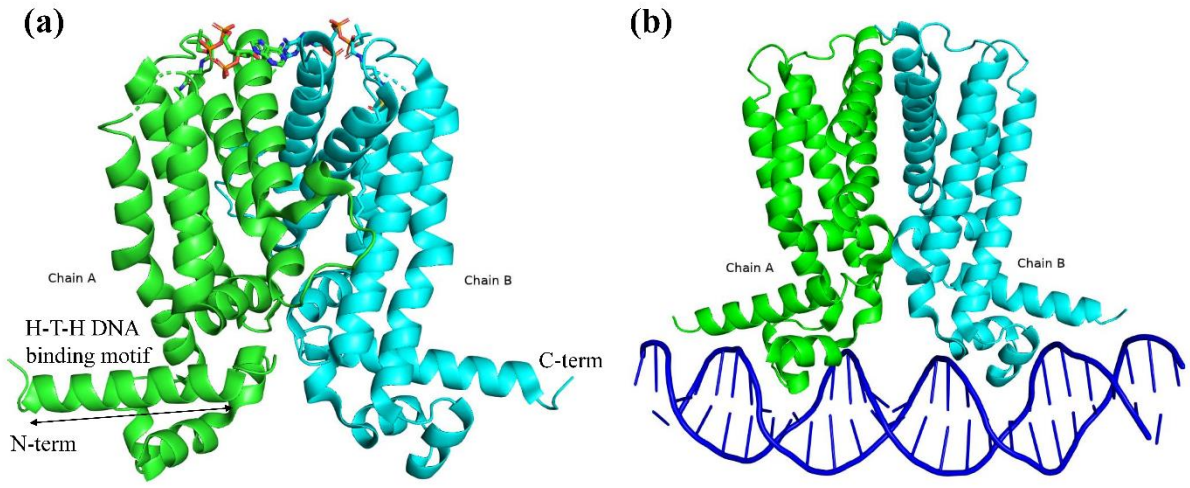


Figure 2.10: Structure of TetR family regulator (DesT). (a) Dimer of DesT in complex with palmitoyl-CoA. Both chain A (green) and chain B (cyan) consist of conserved N-terminal H-T-H DNA binding motifs. (b) DesT bound with its cognate DNA duplex (blue). Structural representations were made through Pymol.

PvrA (the Pseudomonas virulence regulator) is a transcriptional regulator that is a member of the TetR family, which typically functions as a repressor of gene expression. PvrA stimulates the production of the *plcH* phospholipase gene similarly to *gbdR*, but only in response to fatty acyl-CoAs like palmitoyl-CoA [160, 161]. PvrA is an important component of lung surfactant and one of the major carbon sources for *P. aeruginosa* during lung infection. It links virulence to the metabolism of phosphatidylcholine and long-chain fatty acids. PvrA stimulates the synthesis of pyocyanin and rhamnolipids via increasing transcription of the PQS biosynthetic operons of pqsABCDE and phnAB. This is in addition to directly regulating the expression of several metabolic genes and aprA alkaline protease. In addition, it suppresses the expression of phaG, which codes for a protein that transforms 3-hydroxyacyl-acyl carrier protein (ACP), the substrate needed to produce rhamnolipids, into polyhydroxyalkanoate energy-storing molecules [160]. PvrA thus functions as a switch that directs the synthesis of rhamnolipids rather than energy storage from the utilization of 3-hydroxyacyl-ACP. As a result, it was demonstrated that a pvrA mutation dramatically decreased *P. aeruginosa* colonization and survival in mouse lungs.

Chapter 3

General Methodology

3. General Methods

3.1. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

According to Laemmli method, protein samples were separated based on their molecular mass in SDS-PAGE [162].

Stock solutions:

30% (w/v) acrylamide mix: Dissolve 29.2 gm of acrylamide (SRL; 3x crystal Cat. No.- 61346) and 0.8 gm N, N- methylene -bis-acrylamide (SRL 3x crystal; Cat. No.- 38516) in 30 ml of distilled water. The volume was made upto 100 ml with distilled water.

Separating buffer or resolving buffer: By dissolving 18.13 grams of Tris base (SRL; Cat. No.-71033) in 90 ml of distilled water, 1.5 M Tris-HCL, pH 8.8 was created; 0.4 gm of SDS (SRL; Cat. No.- 54468) was added; the pH was then adjusted to 8.8 with 3M HCL. Distilled water was used to get the final volume upto 100 ml.

Stacking buffer: 6 grams of Tris base were dissolved in 80 ml of distilled water to create 0.5 M Tris-HCL, pH 6.8.; 0.4 gm of SDS was added. Next, using 3M HCL to bring the pH down to 6.8, the final volume was increased to 100 ml using distilled water.

Ammonium persulfate solution (APS): Dissolve 0.3 gm of ammonium persulfate (SRL Cat. No.- 65553) in 1.5 ml of distilled water.

TEMED- N,N,N,N-Tetramethyl Ethylenediamine (SRL Cat. No.- 84666).

Working solutions:

Table 3.1: Resolving gel buffer solution

Components	Composition				
	8%	**10%**	**12%**	**15%**	**20%**
Distilled water	3.86 ml	3.33 ml	2.8 ml	2 ml	1.66 ml
30% (w/v) acrylamide mix	2.14 ml	2.67 ml	3.2 ml	4 ml	5.34 ml
Resolving buffer	2 ml	2 ml	2 ml	2 ml	2 ml
20% APS*	30 µl	30 µl	30 µl	30 µl	30 µl
TEMED*	18 µl	18 µl	18 µl	18 µl	18 µl

Table 3.2: Stacking gel buffer solution	
Components	**Composition**
	5%
Distilled water	1.112 ml
30% (w/v) acrylamide mix	200 µl
Stacking buffer	187 µl
20% APS*	6 µl
TEMED*	6 µl

*APS and TEMED were introduced to every solution before polymerization process.

SDS gel electrophoresis buffer: 25 mM Tris base (SRL, Cat. No.- 71033), 190 mM glycine (SRL, Cat. No.- 25853) and 0.1% SDS (SRL, Cat. No.- 54468) was made by dissolving:

- Tris-base - 3.0 gm
- Glycine - 14.3 gm
- SDS - 1.0 gm

Distilled water was added to get a final volume of 1000 ml.

SDS sample gel loading buffer: 0.1 M Tris-HCL, 0.025 M dithiothreitol (DTT), 50% glycerol, 10% SDS and 0.025% bromophenol blue was made by dissolving:

- Stacking buffer solution – 0.5 ml
- DTT – 48 mg
- Glycerol – 1.25 ml
- SDS – 240 mg
- Water – 0.75 ml
- Bromophenol blue (12.5%) – 5 µl

This buffer was stored at -20° C.

Procedure:

Polymerization: gel electrophoresis was carried out in vertical gel electrophoresis apparatus (Tarsons). The whole casting assembly was set up at first. Stacking gel solution was loaded into the apparatus first and kept for around 20 minutes for solidifying. The resolving gel solution at required percentage was then added with a 0.75 mm thick 10-well teflon comb and kept for solidifying for around 20 minutes. Later, the comb was removed and the casting assembly was placed in an electrophoresis buffer chamber which was further filled up with gel electrophoresis buffer.

Sample preparation: Complete denaturation was achieved by boiling protein samples in a water bath for three minutes after mixing them with one-fifth of their volume of SDS sample gel loading buffer.

Sample loading and electrophoresis: Different protein samples were loaded into individual wells with a standard protein marker. Then electrophoresis was carried out at a constant voltage of 100 volts and 16 ampere current until the bromophenol blue dye front reached the baseline.

Protein staining and destaining:

Staining solution: 0.3 grams of brilliant blue R250 were dissolved in 100 ml of destaining solution to create 0.3% staining solution. The solution was then filtered and stored in a bottle.

Destaining solution: 50% methanol, 10% acetic acid, 40% water.

Procedure: After electrophoresis, gels were placed in staining solution and gently shaken for 15 minutes. Then excess staining solution was discarded and gels were destained by destaining solution until the background stain was sufficiently clear to visualize the protein bands. After that, gels were transferred to distilled water.

3.2. Protein concentration estimation

The concentration of purified protein was estimated by Folin lowrey method as described in the protocol given by Waterborg et al. [163]. Estimation of protein concentration was also done by taking absorbance at 280 nm. 2 µl of purified protein sample mixed with 998 µl of buffer and absorbance was read at 280 nm in a spectrophotometer with a reference sample.

Concentration was estimated by the equation given below:

$$C = (A/\varepsilon l) \times Df \times 10^6$$

Where,

C = Concentration of protein in µM

A = Absorbance at 280 nm

ε = Molar absorption coefficient of protein

l = Path length (1 cm in all cases)

D*f* = dilution factor (here 200 in all cases)

3.3. Agarose gel electrophoresis

- Ethidium bromide (EtBr) solution: A stock solution containing 10 mg/ml of ethidium bromide was obtained from Sigma (Cat. No. E1510).
- DNA gel loading dye (5X): Add 25 mg bromophenol blue (Sigma; Cat. No.- B5525), 25 mg Xylene Cyanol FF (Sigma; Cat. No.- X4126), and 3.3 ml glycerol in distilled water. The final volume was adjusted upto 10 ml and stored at -20°C.
- 50X Tris-acetate EDTA (TAE) running buffer: Add 242 gm of Tris base and 57.1 ml of acetic acid in 700 ml distilled water. The final volume was adjusted up to 1000 ml. No need to adjust the pH of the solution.

Procedure: Agarose gel electrophoresis was used to visualize the DNA. 1% (w/v) powdered agarose was dissolved in 1X TAE buffer upon gentle heating. Then, it was cooled to around 60°C and 0.2-0.5 μg/ml of EtBr (10mg/ml) solution was added followed by proper mixing. The agarose solution was immediately poured onto a pre-set electrophoresis chamber including a comb. The whole setup was kept for around 30 minutes to solidify.

The comb was removed from the agarose gel and gel was immersed in electrophoresis running buffer. DNA samples were mixed with DNA gel loading dye, and loaded into individual wells. A 1-5 V/cm voltage was applied, and wells were placed near the cathode so DNA could migrate towards anode. The gel was run until the baseline was reached by xylene cyanol and bromophenol blue. The gel was taken out from the tank and examined with a UV transilluminator.

3.4. Purification of PCR products

The PCR clean up kit (Qiagen; Cat No.- 28104) was used for this purpose.

Kit contents: PCR clean-up columns with collection tube, Nucleic acid binding buffer and Washing buffer.

Procedure:

1. 5 volumes of nucleic acid binding buffer were added to PCR reaction sample and mixed properly.
2. The mix was transferred to a PCR clean-up column, and centrifuged for 1 min at 13,000 rpm.
3. After adding 750 μl of wash buffer to the column, it was centrifuged at 13,000 rpm for 1 minute. After discarding the flow through, centrifugation was done at 13,000 rpm for 1 minute.

4. 30 µl of elution buffer (10 mM Tris; pH 8.0) was added to the column and kept for 5 minutes. Eluant containing PCR product was centrifuged for 1 min at 13,000 rpm and collected in a fresh tube. The PCR purified product was then stored at -20°C.

3.5. Purification of DNA fragments by gel extraction method

The gel extraction kit (Qiagen; Cat. No. 28704) was used for this purpose.

Kit contents: GenElute binding column G, Gel solubilization solution, Washing buffer, and elution buffer.

Procedure:

1. The desired DNA fragment was removed from the agarose gel using a sterile and sharp scalpel.
2. The gel slice was weighed. One volume of gel (100 mg gel ~ 100 µl) was mixed with three volumes of gel solubilization buffer. Incubated the gel slice at 50-55°C for 10 minutes and vortexed the tube every 2-3 minutes intervals.
3. 1 gel volume of 100% isopropanol was added and mixed properly. Sample was moved to the binding column and centrifuged at 13,000 rpm for 1 minute.
4. After discarding the flow through, 750 µl of wash buffer solution was added, and centrifugation was run for one minute at 13,000 rpm. After discarding the flow through, any remaining wash solution was eliminated by centrifuging it once more for a minute at 13,000 rpm.
5. 30-35 µl of elution buffer was added and incubated for 5 minutes. Eluant was collected in a fresh tube.

3.6. Preparation of competent *E. coli* cells

E. coli DH5α or any other strain was grown overnight in Luria broth at 37°C. 1% of this inoculum was added to sterile 100 ml Luria broth. The growth of the cells was maintained at 37°C until the $O.D._{600}$ reached 0.3-0.4. Cells were then pelleted down at 4,000 rpm for 15 minutes at 4°C. All following steps were accomplished at 4°C. Cells were resuspended in 100 mM $MgCl_2$ with gentle mixing and centrifuged for 15 minutes at 3,500 rpm. Cells were then resuspended in 100 mM $CaCl_2$ with gentle mixing and incubated at 4°C for 20 minutes. Cells were pelleted down at 3,500 rpm for 15 minutes and resuspended with solution made up of 85 mM $CaCl_2$ and 15% glycerol. Cell pellets were dissolved properly and centrifuged for 15 minutes at 3,000 rpm. Resuspended cells with 2 ml of 85 mM $CaCl_2$ and 15% glycerol

solution and mixed gently. Later, ~50 µl of cell resuspension was aliquoted in a microfuge tube and kept at -80°C for long-term storage.

3.6.1. Transformation of competent *E. coli* cells

For each transformation, DNA (60-70 ng) was added to a tube containing ~50 µl of competent cells and kept in ice for ~15 minutes. The tube was then placed in a water-bath preheated at 42°C for 90 seconds followed by snap-chilling in ice for 5 minutes. The vial was then filled with 1 ml of Luria broth, grown at 37°C shaker for 1 hour. Cell suspension was pelleted down at 5,000 rpm for 15 minutes and 50-100 µl of culture was then plated on Luria agar plates containing appropriate antibiotics. The plate was then kept in an incubator at 37°C in an inverted position for overnight.

3.7. Isolation of Plasmid DNA

Plasmid DNA was isolated by Qiagen Spin Miniprep kit (Cat. No. 27104).
Procedure:
1. 1-3 ml of overnight culture was centrifuged at 7,000 rpm for 5 minutes in a microfuge tube.
2. After discarding the supernatent, 250 µl of cell resuspension solution was added and mixed properly.
3. 250 µl of lysis solution was added to the tube and inverted it for 3-4 times. Then, 350 µl of neutralization solution was added, inverted the tube for 3-4 times. Centrifuged the tube for 13,000 rpm for 10 minutes.
4. The supernatant was placed in a Qiagen binding column and centrifuged for 1 minute at 13,000 rpm.
5. After adding 750 µl of wash solution to the column, it was centrifuged for 1 minute at 13,000 rpm. To remove any residual wash solution, centrifuged the column again at 13,000 rpm for 1 minute.
6. 30-40 µl of elution buffer was added to the tube and collected in a fresh tube after centrifugation at 13,000 rpm for 1 minute. For later usage, DNA was kept in storage at -20°C.

3.8. Protein expression and purification methods:

To express the desired protein, the positive clone of the recombinant plasmid was introduced into *E. coli* BL21 competent cells.

3.8.1. Induction and solubility test

To check the expression of desired protein, a single colony was inoculated into a 5 ml primary culture and was grown at 37°C overnight in a shaking condition. Another 6 ml secondary culture was given from primary culture and was grown at 37°C until $O.D._{600}$ reached 0.5-0.6. 1 ml of culture was taken in a fresh microfuge tube denoted as an 'uninduced' culture for control and 1 mM of IPTG was given in the rest of the culture for induction. Both cultures were grown at 37°C until $O.D._{600}$ reached 0.5-0.6. Cultures were harvested at 7,000 rpm for 5 minutes. 5 ml of sonication buffer was added to both uninduced and induced cultures and resuspended properly. Both cultures were sonicated for cell lysis and centrifuged at 12,000 rpm for 45 minutes. SDS-PAGE gel was used to analyze the protein expression and solubility of the supernatant and pellet fractions.

3.8.2. Immobilized metal affinity chromatography (IMAC)

N-terminal or C-terminal His_6-tagged proteins were purified first through Ni^{2+}-NTA (nitrilotriacetic acid) affinity chromatography method. Overnight large culture (500-1000 ml) grown at 37°C were harvested at 7,000 rpm for 5 minutes. All necessary following steps were performed at 4°C. For lysis, the cell pellet was resuspended in sonication buffer and subjected to sonication. The supernatant was loaded onto pre-equilibrated column containing Ni^{2+}-NTA beads and incubated at 4°C for 15 minutes with intermittent mixing. Flow through was passed out from the column, wash 1 buffer was added and flow through was passed out. In addition, wash 2 buffer was also added and flow through was passed out. Finally, elution buffer was added to the column and incubated at 4°C for 15 minutes. After the eluate was collected in a new tube, an SDS-PAGE gel analysis was used to determine its purity.

3.8.3. Size-exclusion chromatography

This step is required for further polishing of desired protein. After IMAC purification, protein sample was loaded onto Superdex 200 16/600 prep-grade column which was pre-equilibrated with gel filtration buffer. Desired protein and any other contaminant proteins were separated based on their size and further analyzed for purity by SDS-PAGE gel.

Chapter 4

Structural and Functional Insights of Itaconyl-CoA hydratase (Ich) with a Novel N-terminal Hotdog fold

4.1. Introduction

In the current worldwide scenario, an increasing number of multidrug-resistant (MDR) organisms are a major health concern for treating the severity of diseases in our society. Among them is *Pseudomonas aeruginosa*, an opportunistic gram-negative bacteria that causes acute or persistent infections in immunocompromised people suffering from sepsis, burns, trauma, cancer, and chronic lung illnesses [164, 165]. To cause infection, it secretes some virulence factors either extracellularly or by specialized secreting machinery into host cells. These virulence factors then modulate the host immune response by targeting some essential proteins or immunomodulatory substances in the host cells [166]. It ultimately led to downregulating the defensive mechanism and maintaining bacterial proliferation and dissemination in eukaryotic cells [167, 168].

To overcome infection by these virulence factors, phagocytic cells like macrophages act as a secondary line of defense against invading pathogens in eukaryotic host cells. Over many secretory substances, they secrete some antimicrobial products at the infection site, inhibiting specific biological processes in pathogenic bacteria. Of them, itaconic acid (methylene succinic acid), an unsaturated 1,4-dicarboxylic acid is also known as a potent inhibitor of isocitrate lyase of bacterial glyoxylate cycle which is required for the assimilation of acetyl-CoA as a carbon source for bacteria upon degradation from fatty acids [69, 79, 169]. Many pathogenic bacteria like *Pseudomonas* spp., *Yersinia* spp., *Micrococcus* spp., and *Salmonella* spp. used itaconate as a sole carbon source and catabolized it into acetyl-CoA and pyruvate for their survival in human host cells [88-90]. In this itaconate catabolic pathway, succinyl-CoA:itaconate CoA transferase (Ict) first activates itaconate to its corresponding CoA-ester (itaconyl-CoA), then (*R*)-specific itaconyl-CoA hydratase (Ich) hydrates itaconyl-CoA to (*S*)-citramalyl-CoA and at last (*S*)-citramalyl-CoA lyase cleaves (*S*)-citramalyl-CoA into acetyl-CoA and pyruvate **(Figure 4.1)** [53]. Later, Zrieq et al. called this gene product PA14_52910 (*Pa*Ich) as *Pseudomonas aeruginosa* effector candidate 1 (Pec1) based on showing virulence in eukaryotic host cells [170]. Thus, previous literature showed the crucial role of *Pa*Ich in itaconate degradation pathway as part of bacterial defense weaponry against macrophages.

Despite this significance, structural and mechanistic insights into *Pa*Ich and hydration of itaconyl-CoA remain undetermined. Here, in this present study, we report the first three-dimensional structure of *Pa*Ich at 1.98 Å resolution. The crystal structure of *Pa*Ich shows that

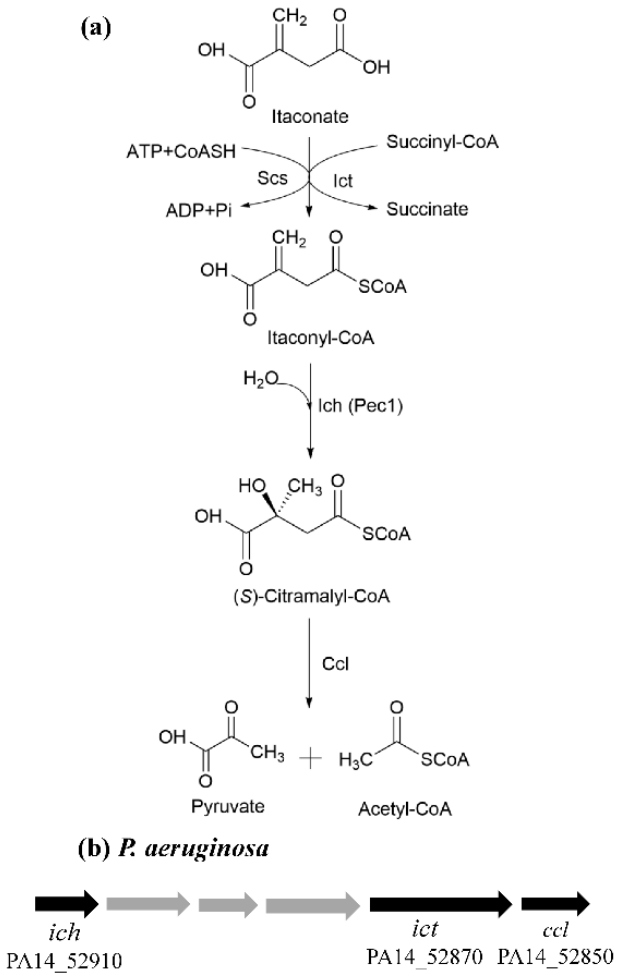


Figure 4.1: Schematic representation of itaconate catabolic pathway and corresponding genetic constituents of *P. aeruginosa*. (a) Itaconate degradation pathway showing conversion of itaconate to pyruvate and acetyl-CoA. Itaconyl-CoA hydratase (Ich) from *P. aeruginosa* also known as Pec1 catalyzes the conversion of itaconyl-CoA to citramalyl-CoA by addition of H2O. (b) Corresponding gene clusters of itaconate degradation pathway.

it is dimeric in nature confirmed by also in-solution by mass spectrometry analysis. The structure resembles MaoC-like hydratases/dehydratases consisting of N- and C-terminal domains connected by a long stretch of flexible loop. In addition, we found that it shows a significant binding affinity towards acetoacetyl-CoA like in crotonase and mitochondrial enoyl-CoA hydratase of bovine and rat liver respectively [171, 172].

4.2. Materials and Methods

4.2.1. Cloning, expression and purification

Itaconyl-CoA hydratase (*Pa*Ich) was cloned, heterologously produced and purified to extreme homogeneity prior to crystallization. Full-length gene was amplified using Nde1 and BamH1

as restriction sites and the following primers used in this study are 5′-TGCCATATGAGTGAGTCCGCTTTCGCC-3′ as forward primer and 5′-ATTGGATCCTTAGTCGAATTCCACGTCGCC-3′ as reverse primer. The PCR product was cloned into N-terminal His_6-tagged pET28a vector (Novagen) using the same restriction site. This recombinant plasmid was then transformed into chemically competent BL21 (DE3) cells (Thermo Fisher Scientific). Culture was grown in 500 ml of LB supplemented with 50 mg/ml (final conc.) kanamycin at 37°C in a continuous shaking condition until the O.D. reached ~0.6-0.8. After that, 0.5 mM IPTG was added to the culture for induction and shaken overnight at 37°C. Cells were harvested at 6000 rpm for 10 min and pellets were dissolved in re-suspension buffer (50mM Tris; pH 8.0, 150 mM NaCl, 5 mM Imidazole and 5% glycerol) with 1 mM PMSF prior to sonication. Lysed cells were then centrifuged at 12,000 rpm for 1 hour and the supernatant was then loaded into Ni^{2+}-NTA column which was pre-equilibrated with equilibration buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 10 mM Imidazole and 5% Glycerol). 1 column volume of wash 1 buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 35 mM Imidazole and 5% Glycerol) followed by 3 ml of wash 2 buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 50 mM Imidazole and 5% Glycerol) was added. After washing, protein was eluted with elution buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 250 mM Imidazole and 5% Glycerol). Eluted protein was dialyzed against dialysis 1A buffer (50 mM Tris; pH 8.0, 150 mM NaCl, and 5% Glycerol) for 3-4 hrs. and then it was shifted to dialysis 1B buffer (50 mM Tris; pH 8.0, 150 mM NaCl, and 5% Glycerol) for another 9-10 hours. The protein was then loaded into Hiload 16/60 Superdex 75 prep grade column pre-equilibrated with gel filtration buffer (50 mM Tris; pH 8.0, 150 mM NaCl, and 2% Glycerol). Eluted protein was collected and concentrated in a 30 kDa cut-off spin concentrator until the concentration reached ~13 mg/ml. The purity of protein was then checked by SDS-PAGE analysis.

4.2.2. Crystallization, data collection and processing

Purified protein was crystallized in sitting drop vapour-diffusion method. Initial crystal hit was obtained in 20% PEG 3350, 200 mM KNO_3 of Wizard 3 screen and crystals appeared in a thin plate-like morphology stacked one after another. The whole setup was incubated at 20°C. Single and large crystals were grown in 3-4 weeks and obtained in 22% PEG 3350, 220mM KNO_3. Datasets for both native and heavy atom derivatives were collected in Bruker D8 Venture using Cu K_α source (~1.54 Å) and Photon III CCD detector was used here for collecting frames **(Table 4.1)**. The data were integrated and scaled in Proteum 3 software and were found to belong to space group P1 with unit cell parameters *a*=65.312 Å, *b*= 66.197 Å,

c= 77.462 Å, α= 102.45°, β= 95.07° and γ=101.06°. Matthew's coefficient of 2.57Å^3/Da assumed that there are four copies present in each asymmetric unit and the solvent content is 52.2%. Initially, the coordinates of heavy atom positions and occupancies were calculated in Phenix by Hyss method [173, 174]. Estimation of phasing power was calculated by Autosol and initially, the model was built by Autobuild in Phenix respectively [175-177]. Successive model building and refinement were done in Coot and Phenix respectively [178]. Structure-based sequence alignment was done by DALI analysis [179]. All structural representations were made through Pymol. The atomic coordinates for *Pa*Ich have been deposited in the Protein Data Bank **(PDB ID: 8HUC)**.

Table 4.1: Data collection and Refinement Statistics

Dataset	Native	KAuCN	$HgCl_2$	C_2H_5HgCl
Data collection and phasing statistics				
Wavelength (Å)	1.54	1.54	1.54	1.54
Resolution range (Å)[a]	31.76-1.98 (2.05-1.98)	27.60-1.66 (1.71-1.66)	27.89-3.27 (3.39-3.27)	27.60-3.01 (3.12-3.01)
Spacegroup	P1	P1	P1	P1
Unit cell dimensions (a,b,c=Å; α,β,γ=°)	*a*=65.31, *b*= 66.19, *c*= 77.46, α= 102.45, β= 95.07, and γ=101.06	a=65.43, b=66.18, c=77.60, α=102.46, β=95.04, and γ=101.08	a=65.39, b=66.40, c=77.43, α=102.49, β=95.13, and γ=101.29	a=65.38, b=66.08, c=77.34, α=102.78, β=95.09, and γ=101.20
Total reflections[a]	1608435 (54330)	1279481 (9533)	550217 (45323)	528740 (47828)
Unique reflections[a]	83791 (7025)	137054 (7165)	18987 (1830)	24258 (2396)
Redundancy	19.2	9.3	29.0	21.8
Completeness (%)[a]	98.2 (82.6)	93.0 (48.6)	99.6(96.0)	99.8(98.2)
Overall I/σ[a]	18(1.5)	9.8(2.6)	20(8.7)	28(12)
R_{merge} (%)[a]	12.7(12.3)	17.5(69.5)	18.7(40.6)	10.4(20.5)
Refinement statistics				
Resolution range (Å)	31.76-1.98			
Number of used refelections	77191			
R_{work} (%)	18.22			
R_{free} (%)	22.06			
Total number of atoms	17191			
Protein	16186			
Water/glycerol/KNO_3	982/7/14			
Average B-factors (Å^2)	34.0			
Root mean square deviations				

Bonds (Å)	0.012
Angles (°)	1.168
Ramachandran plot	
Most favoured region (%)	97.74
Ramachandran outliers (%)	0.20

a Values in parentheses refer to the highest resolution shell.

4.2.3. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed using Microcal VP-ITC 200 (Malvern) with a cell volume of 350 µl. Acetoacetyl-CoA (syringe) at 45 µM was titrated against 1.75 µM of *Pa*Ich (cell) with a stirring speed of 300 rpm using titration buffer (25 mM Tris; pH 8.0, 150 mM NaCl) at 25°C. A total of 28, 10 µl injections were made every 180 seconds. The raw data was analyzed in ORIGIN software [180].

4.3. Results

4.3.1. Biochemical analysis of *Pa*Ich

Full-length N-terminal His_6-tagged *Pa*Ich (~32 kDa) was heterologously produced in *E. coli* and purified to apparent homogeneity using affinity chromatography method. The purity of eluant after affinity chromatography method was further analyzed by SDS-PAGE. The protein eluted as a dimer (~64 kDa) in size exclusion chromatography method. The biologically active form of *Pa*Ich presented as a dimer which was also confirmed by mass spectrometry analysis **(Figure 4.2)**.

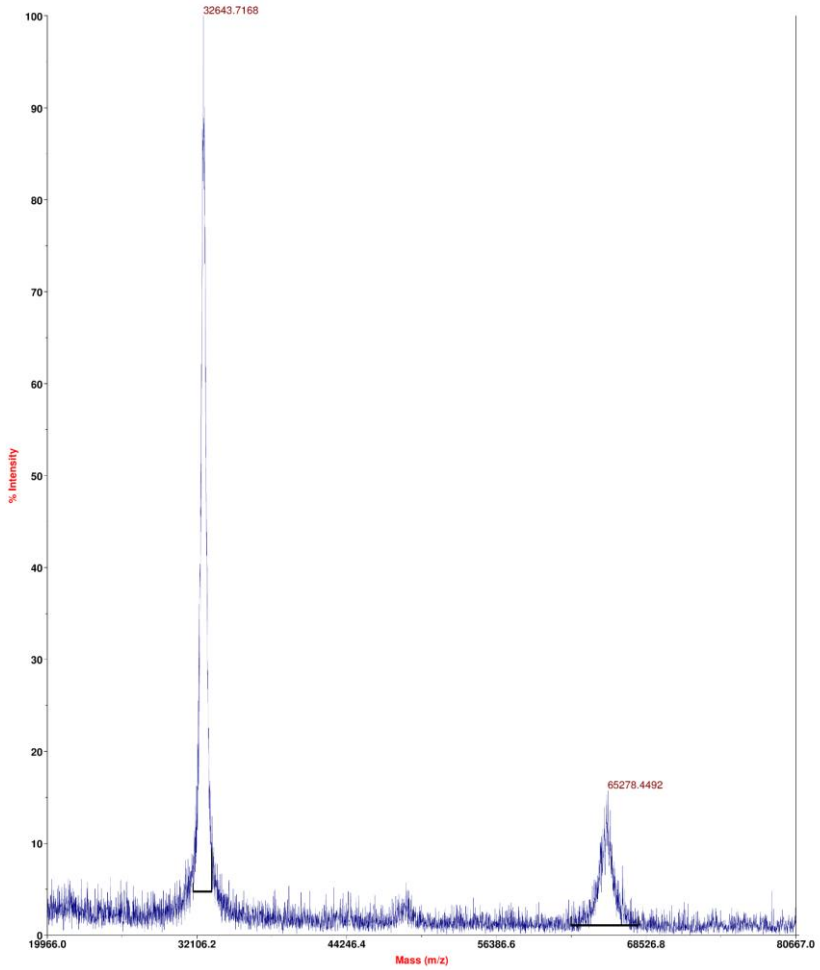


Figure 4.2: Mass-spectrometry analysis of *Pa*Ich. The molecular weight of *Pa*Ich was calculated from two respective peaks (value shown in red) using the formula:

a. p = m/z where, p_1 = 32643.716 and p_2 = 65278.450

b. $p_1 = (M_r+z_1)/z_1$ m = total mass of an ion, z = total charge, M_r = average mass of protein

c. $p_2 = [M_r+(z_1-1)]/(z_1-1)]$ hence, Calculated mass = 65285.432 kDa

4.3.2. Overall structure of *Pa*Ich

Here, we report the first crystal structure of itaconyl-CoA hydratase from *P. aeruginosa* (*Pa*Ich) at 1.98 Å resolution. Initially, the phase retrieved by molecular replacement method was not successful because of very low sequence similarity (<20%) with other hydratases/dehydratases classes of enzymes. Eventually, multiple isomorphous replacement with anomalous scattering (MIRAS) method was used to solve the structure using heavy-atom derivatives (Au and Hg). The more detailed structural statistics of native and heavy-atom derivatized crystals are summarized in **Table 4.1**. The crystallographic asymmetric unit

(ASU) of *Pa*Ich consists of four copies (chains A, B, C and D) of a monomeric unit. Therefore, two dimers (A-C and B-D) are present in one asymmetric unit, where each dimer represents one biologically active form **(Figure 4.3a, b)**. In ASU, two dimers are oriented at 90° to each other and each monomer of this dimer is related to each other by 2-fold symmetry. The overall structure of *Pa*Ich resembles the structure of MaoC-like hydratases/dehydratases (PDB ID: 1PN2, 3KH8, and 5I7N) composed of two domains, the N-terminal half and the C-terminal half connected by an intervening bridge **(Figure 4.3c)**. Moreover, C-terminal domain of *Pa*Ich was found to be similar to other (*R*)-specific enoyl-CoA hydratases/dehydratases whereas N-terminal domain was shown to be different. The R.M.S.D. values of structural superposition of N-terminal domain of *Pa*Ich with other (*R*)-specific enoyl-CoA hydratases/dehydratases such as 1PN2, 3KH8, and 5I7N are 2.03, 0.97 and 1.11 Å respectively.

Each domain of *Pa*Ich consists of 2-4 α-helices and 5-6 continuous parallel β-strands comprising a β-sheet scaffold. This β-sheet scaffold of both N- and C-terminal domains provide the main architectural framework for each monomer. The N-terminal end of *Pa*Ich consists of an overhanging segment of 3-7 residues followed, by first strand of β-sheet scaffold. Each β-sheet scaffold of N- and C-terminal domains is represented as a bun of the so-called 'hotdog fold' which was seen in other hotdog folds containing (*R*)-specific enoyl-CoA hydratases/dehydratases (PDB ID: 1PN2, 3KH8, and 5I7N) [110, 119, 181-184]. A long central helix α13 (Gly_{207}-Ala_{221}) represented a sausage which was properly wrapped up by the C-terminal β-sheet scaffold constituting the C-terminal hotdog fold. In contrast, a very short helical segment α3 (Trp_{47}-Ala_{50}) which we termed an 'eaten sausage', is not properly bound by a β-sheet scaffold comprising the N-terminal hotdog fold. The structural framework of C-terminal hotdog fold is very similar to other DHF (PDB ID: 1PN2, 3KH8, and 4E3E) and also SHF-containing (*R*)-specific enoyl-CoA hydratases/dehydratases (PDB ID: 1IQ6, and 5CPG). In contrast to C-terminal hotdog fold, N-terminal hotdog fold of *Pa*Ich has distinguished characteristic features [99, 119, 185, 186]. A long flexible loop (Phe_{51}-Arg_{79}) (average b-factor 45.08 $Å^2$) connects helix α3 (eaten sausage) to the succeeding β5 strand of N-terminal domain. An extra short helix α4 (Ala_{58}-Gly_{60}) is observed in this loop which is found in all monomers (chain B, C and D) in the asymmetric unit except in chain A. It suggests that a helix-loop or loop-helix transition occurred in that region. However, this particular Arg_{69}-Arg_{79} region of that extended loop (Phe_{51}-Arg_{79}) shows a high b-factor of ~81.62 $Å^2$ comprising 163 atoms. Similar to other DHF-containing hydratases, α2 (Val_{25}-

Phe$_{32}$) helix of the N-terminal domain is stacked on α10 helix of the C-terminal domain in *Pa*Ich.

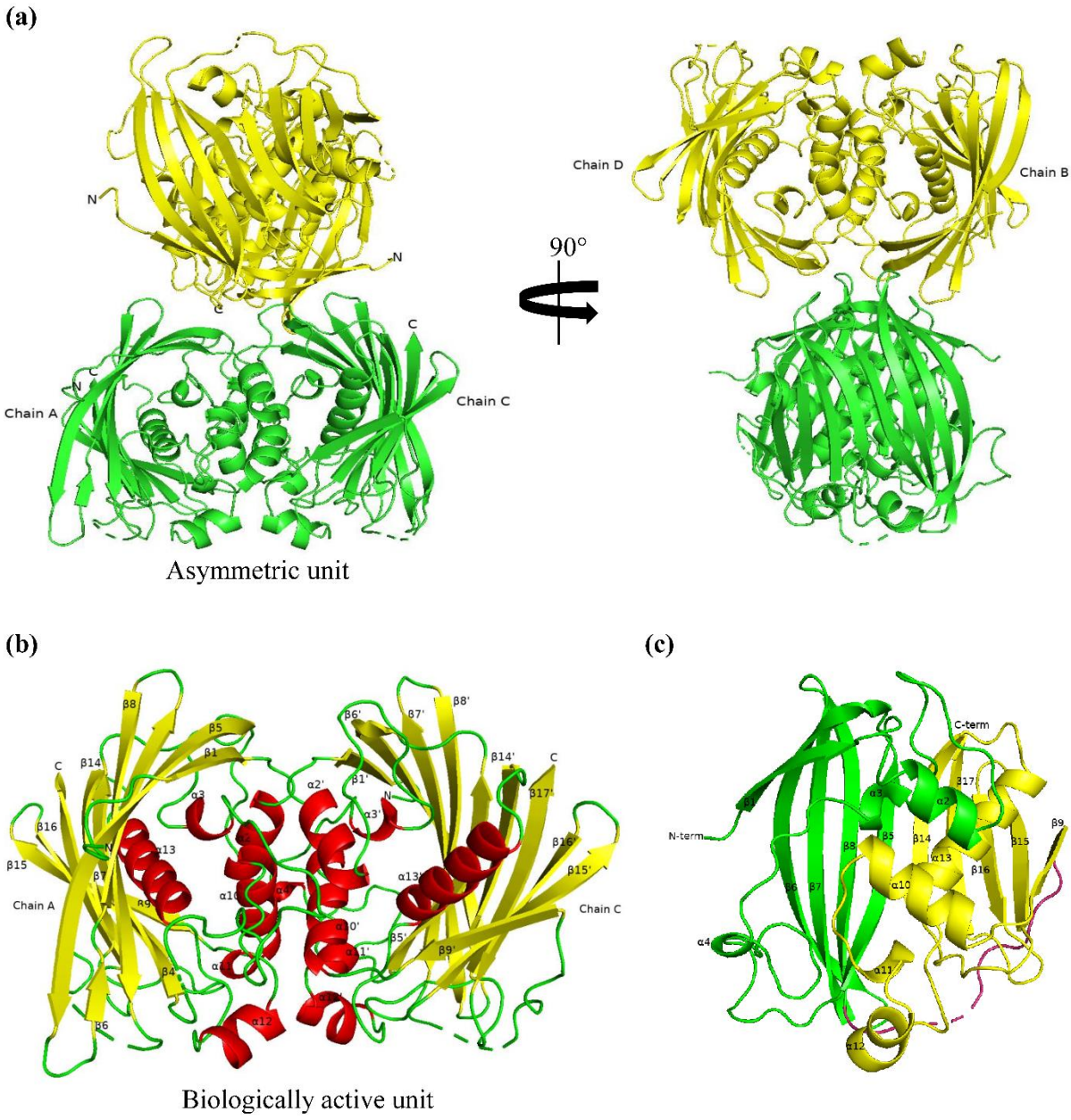


Figure 4.3: Overall structure of itaconyl-CoA hydratase from *P. aeruginosa* (*Pa*Ich). (a) The whole asymmetric unit of *Pa*Ich consists of a tetrameric assembly. Chain A and chain C formed one dimer (green) whereas chain B and chain D made another dimer (yellow). (b) Overall structural view of a dimeric assembly (chain A-chain C) is represented here as a functional unit of *Pa*Ich. Each α-helix and parallel β-pleated sheet are coloured red and yellow respectively and they are labelled sequentially in order of their arrangement. (c) Cartoon representation of monomeric unit (chain C) showing the N-terminal and C-terminal domain comprised of a double hotdog fold. The N-terminal

and C-terminal domains are coloured green and yellow respectively and the intervening highly flexible segment is coloured warm pink. All structural representations were made through Pymol v2.5.4.

4.3.3. Dimeric unit

As previously stated, the asymmetric unit of *Pa*Ich has two dimers where chains A-C represent one dimer and chains B-D form another dimer. The *Pa*Ich dimeric unit is like a crab's shell with a dimension of ~77.5 x 60.8 x 54.2 Å. It was remarkably observed that each monomer consists of two β-sheets with a variable number of strands and their arrangement present in this dimeric assembly. However, no significant R.M.S.D. value was observed between the monomers. Each monomer consists of 275 amino acid residues consisting of 6 α-helices and 11 strands of parallel β-sheet in chain A while 7 helices and 10 strands of β-sheet are in its complementary chain C **(Figure 4.3b)**. Each helix and β-strand of a monomer was sequentially numbered based on their arrangement **(Figure 4.3b, c)**. These strands of β-sheet range from 3 to 13 residues in length. The order of their arrangement in chain A is 1-4-5-6-7-8-9-14-15-16-17 while the β4 strand is not present in chain C (R.M.S.D. value over Cα atoms between chains A and C is 0.207) **(Figure 4.3b)**. Nevertheless, β-sheet consistency of A-C was not observed in another dimer B-D. However, the arrangement of β-strands in chain B is 1-5-6-7-8-9-10-15-16-17-18-19 order while in subunit D it is found in 1-5-6-7-8-9-14-15-16-17 order. So, it was deduced from structural analysis that β-sheet arrangement between two dimers of an asymmetric unit is different (R.M.S.D value over Cα atoms between A-C and B-D is 0.271) while monomers such as chain C and chain D from different dimer consists of a similar arrangement. An extra β6 strand (Ala_{118}-Leu_{120}) is formed in chain A by splitting up the β7 strand of chain C (Gly_{117}-Gln_{130}). Although it was not seen in counterpart chain A, the C-terminal end of β5 followed by N-terminal end of subsequent β6 strand of chain C produces a bend in the opposite direction from the typical hotdog fold. Unlike dimer A-C, the β5 and β16 strands of chain D split in counter chain B to form two new strands, β5-β6 and β16-β17, respectively. Like other DHF-containing hydratases, two domains of *Pa*Ich are connected by an extremely flexible solvent-exposed loop. The electron density of some residues (Thr_{148}-Glu_{156}) was not found in that region because of the high B-factor (average b-factor ~90$Å^2$).

Most of the residues involved in dimeric interfaces are hydrophobic in nature which ultimately leads to overall dimeric stability. The buried region of ~252 $Å^2$ was contributed by

each monomer. The total surface area of *Pa*Ich is about 25978 $Å^2$. Helices α2, α10, α12 and the C-terminal half of the first strand of β-sheet from both monomers participate in the dimeric stability. It was observed that α2 (Val_{25}-Phe_{32}) helix of chain A interacts with the loop region (Phe_{51}-Glu_{57}) of its counterpart chain C. On the contrary, α2 from chain C makes contact with part of a large disordered loop region (Gln_{53}-Ala_{58}) of chain A to provide conformational stability to that region. ε-amide of Lys_{26} in chain A makes a hydrogen bond contact with γ-carbonyl oxygen of Asp_{54} in chain C while side chain of Lys_{26} in chain C makes two H-bonding interactions with Gln_{53} and Asp_{54} of chain A respectively. Another residue from chain C, Arg_{27}, forms H-bonds with Ala_{50} and Val_{180} in its dimeric counterpart. The α10 helices of the two monomers forming dimers are positioned 90° apart. The centrally located α10 helices (Pro_{171}-Thr_{181}) of hotdog fold from both subunits are buried deep inside this dimeric interface and around 75% of residues of those helices are hydrophobic in nature. α10 helix of each monomer interacts with a short loop (Phe_{182}-Gly_{184}) from its dimeric counterpart to participate in this dimeric stability. Being exposed to solvent, both α12 helices in dimer are situated parallelly to each other. The N-terminal part of small helix-12 (Tyr_{189}-Pro_{192}) and C-terminal end of β1 (His_{17}-Gln_{19}) from both subunits interact hydrophobically with each other. Furthermore, this dimer receives structural stability from the C-terminal end of β1 (His_{17}-Gln_{19}) of both chains. It was noted that an H-bond is formed between the δ-O carbonyl amide of Gln_{19} from chain A and the δ-NH2 of Gln_{19} from chain C. In addition, the γ-carboxylic group of Asp_{18} and carbonyl oxygen of Gln_{19} in its counterpart chain A also participate in H-bonding interactions with the γ-carbonyl amide of Asn_{23} in its counterpart subunit C and vice versa. Two pi-pi interactions were also observed between the aromatic side chains of Phe_{175} and Tyr_{189} from both subunits.

4.3.4. Comparisons between N-terminal domain of *Pa*Ich and other MaoC family hydratases

As previously stated that N-terminal domain possesses distinct characteristic features which were not seen in other hydratases, whereas C-terminal domain is found to be similar throughout all hydratases. A basic feature of all hotdog fold-containing hydratases indicated a 9-17 residues long central helix represented as sausage properly wrapped by β-sheet as a bun [99, 119, 185, 187]. A new kind of hotdog fold in structure of *Pa*Ich breaks the universality of basic features of other hotdog folds containing (*R*)-specific enoyl-CoA hydratases.

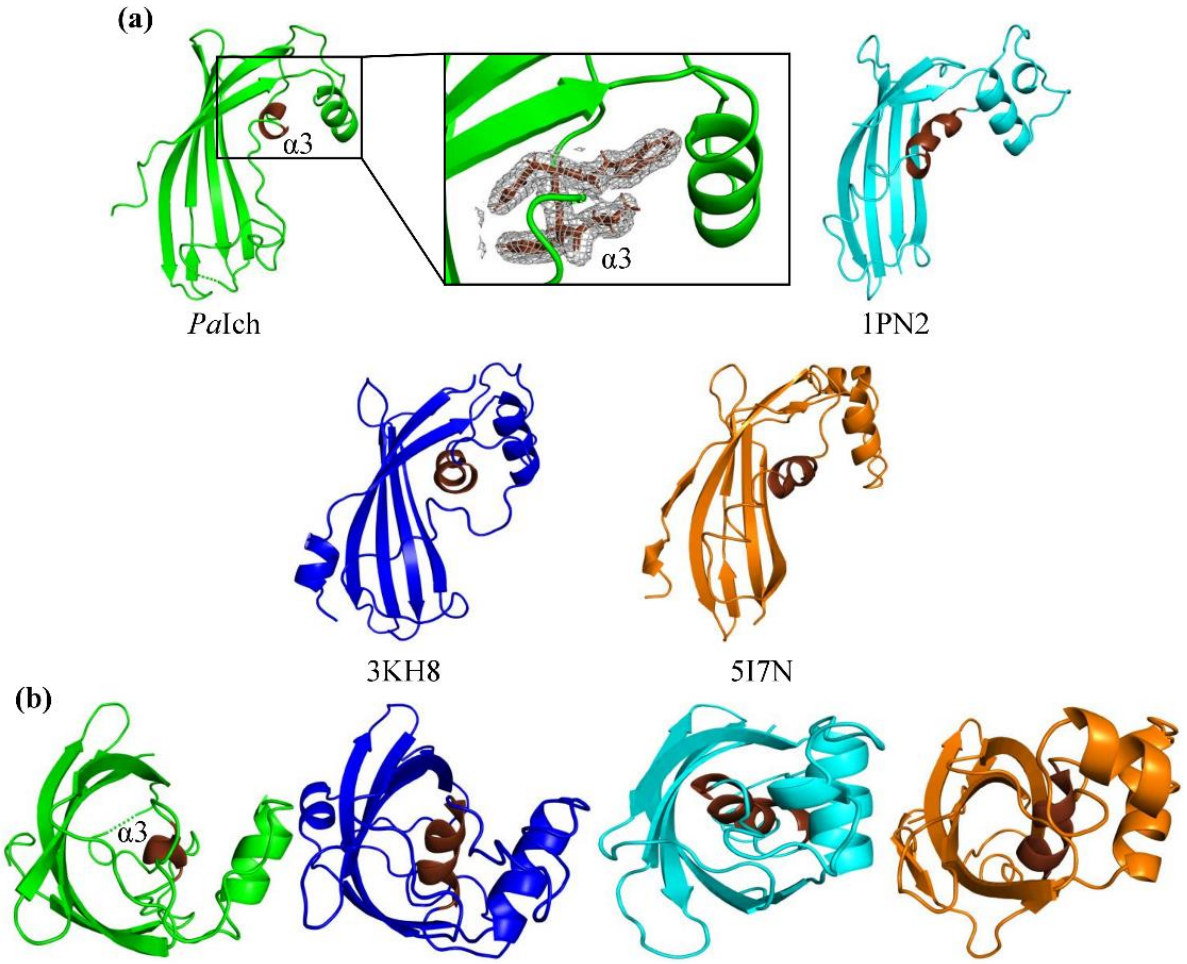


Figure 4.4: A unique N-terminal hotdog fold of *Pa*Ich. (a) Cartoon representation of side view of N-terminal hotdog fold. The β-sheets represented here as buns of *Pa*Ich from *P. aeruginosa*, eukaryotic hydratase 2 from *C. tropicalis* (PDB ID: 1PN2) and MaoC-like dehydratases from *P. capsici* (PDB ID: 3KH8) and *M. abscessus* (PDB ID: 5I7N) are shown in green, cyan, blue and orange colour respectively whereas all central helices of those β-sheets of so-called 'hotdog fold' represented as sausage are shown in chocolate colour. Inset: Final $2F_o$-F_c map showing that central helix 'α3' (chocolate) of *Pa*Ich is in right fit (contoured at 1.0 σ). (b) Top view of N-terminal hotdog fold of *Pa*Ich (green) showing that central helix (α3; chocolate) is slipped away from the β-sheet scaffold (bun) while central helices (sausage) of other (*R*)-specific enoyl-CoA hydratases (PDB ID: 1PN2, 3KH8, and 5I7N) are properly wrapped by their respective hotdog folds. All structural representations were made through Pymol v2.5.4.

It is first observed in the N-terminal domain of *Pa*Ich that a very short 4-residues helical segment α3 (Trp_{47}-Ala_{50}) which we named an 'eaten sausage' is not properly placed inside the core of its hotdog fold **(Figure 4.4a, b)**. Top view representation of *Pa*Ich showed that α3 characterized as an eaten sausage slipped away from the β-sheet scaffold **(Figure 4.4b)**. The

extended loop region (~7 residues) succeeding helix α3 is the missing region of sausage seen in other hydratases. The reason for the appearance of very short helical segment α3 is due to α2 and α10 helices in its dimeric counterpart **(Figure 4.5a, b)**. α2 and α10 helices of each monomer interact with the extended loop region (Phe_{51}-Gly_{65}) of the other monomer **(Figure 4.5a)**. Mainly, α2 rather than α10 helix takes part in this interaction and α2 is found to be situated near this extended loop region (Phe_{51}-Gly_{65}). This extended loop region (Phe_{51}-Gly_{65}) contains around 69% hydrophobic residues. Interactions between α2/α10 helices of one monomer and the extended loop region of its counterpart subunit were found to be mostly hydrophobic. These hydrophobic interactions also provide conformation stability to dimer. Conversely, very low polar contacts were made between α3 and the surrounding β-sheet scaffold which ultimately caused relaxed grasping in N-terminal hotdog fold. His_{48} and Trp_{49} are making an H-bonding interaction with the side chain carboxylic group of Glu_{138}. So, due to poor grasping of β-sheet scaffold, α2/α10 helices pulled out the α3 helix (eaten sausage) towards them **(Figure 4.5a)**. Another notable characteristic revealed that the most disordered region (His_{66}-Arg_{79}) of the extended loop (Phe_{51}-Arg_{79}) was located far away from the β-sheet core frame, which is not the case in other hydratases. Electron density of ~5 residues was not found due to high b-factor (~82 $Å^2$ over 163 atoms) of this disordered region (His_{66}-Arg_{79}). Nevertheless, in other hydratases, this area is situated near the inside wall of the β-sheet scaffold.

4.3.5. Structural comparisons between Ich from *P. aeruginosa* (*Pa*Ich) and Ich (RipB) from *Y. pestis*

Here, a model structure of Ich (RipB) of *Y. pestis* was generated using MaoC domain containing dehydratase from *C. auranticus* as a model (PDB ID: 4E3E) by SWISS-PDB **(Figure 4.5c)**. The structure of *Yp*Ich homologue possesses a single hotdog fold whereas *Pa*Ich consists of two hotdog folds connected by a long linker region. The C-terminal hotdog fold of *Pa*Ich is similar to hotdog fold of *Yp*Ich like in other SHF-containing hydratases **(Figure 4.5d)**. A five-stranded β-sheet enclosing 13 residues long central α-helix (Ser_{119}-Val_{131}) forms the hotdog fold of *Yp*Ich. A 70-residues long disordered N-terminal stretch followed by the first strand of β-sheet forms the N-terminal region of *Yp*Ich which is not similar in case of *Pa*Ich.

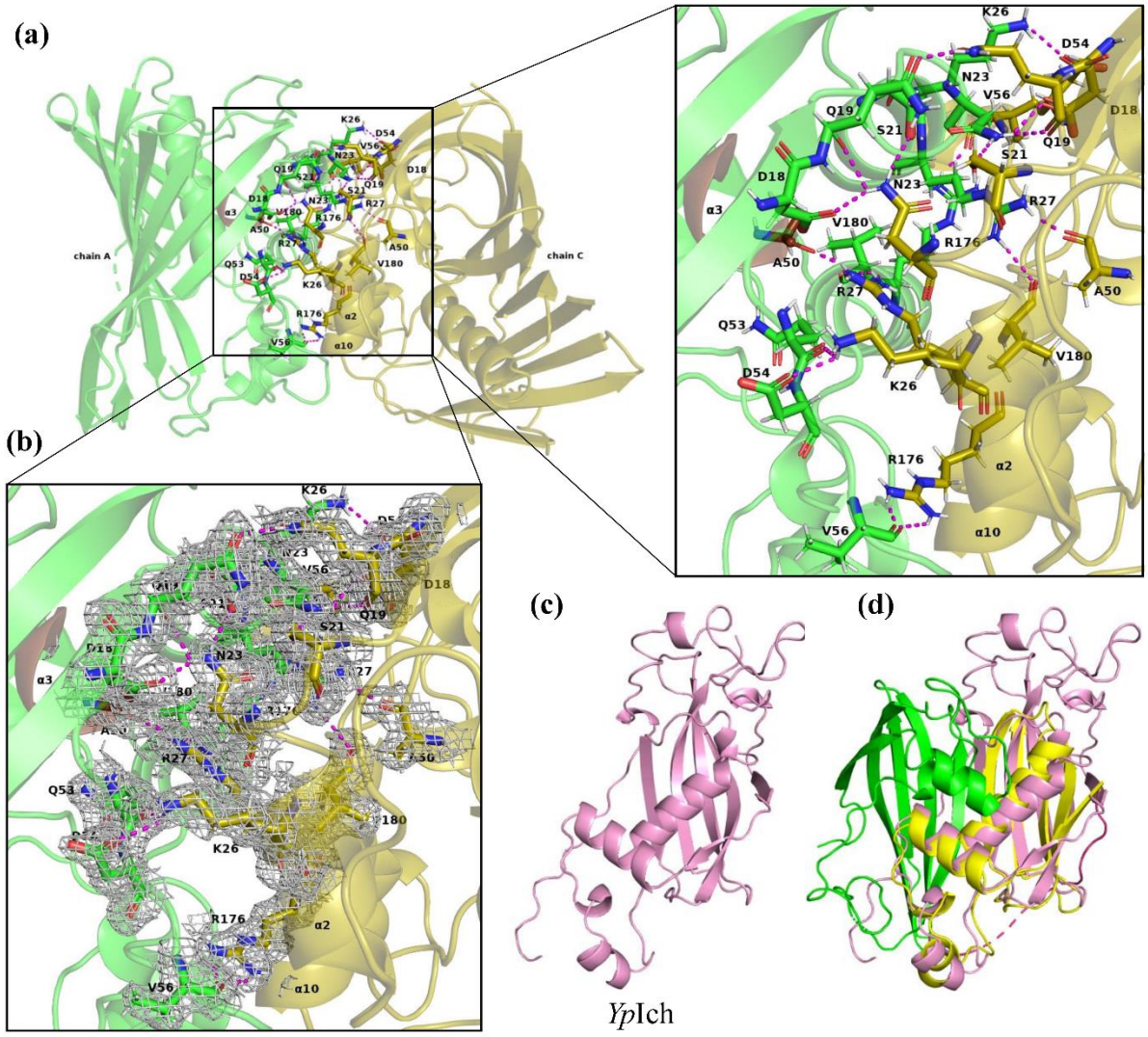


Figure 4.5: Dimeric interfaces between two monomers (chainA-chainC) and model structure of *Yp*Ich from *Y. pestis*. (a) α3 represented as an eaten sausage of monomer A (green) was pulled out from hotdog fold by α2/α10 helices of dimeric counterpart C (olive). Inset: Residues involved in H-bonding interactions between both monomers. Only polar contacts (magenta) are shown here. (b) Inset: Final $2F_o$-F_c map around residues involved in dimeric interfaces showed that they were in right fit (contoured at 1.0 σ). (c) Side view representation of model structure of Ich (pink) from *Y. pestis* (*Yp*Ich) consists of a single hotdog fold. (d) Structural superposition between *Yp*Ich and *Pa*Ich showed that SHF of *Yp*Ich showed maximum similarity with C-terminal domain (yellow) of *Pa*Ich with DHF. All structural representations were made through Pymol v2.5.4.

4.3.6. Substrate binding tunnel and active site

Unlike two active sites located at both top and bottom of dimeric interfaces of SHF-containing (*R*)-specific enoyl-CoA hydratases, it was observed that only one active site region is present in DHF-containing (*R*)-specific enoyl-CoA hydratases [99, 182, 183]. Superposition of *Pa*Ich with its most closely related homologue of *Ct* hydratase 2 (PDB ID: 1PN2) showed a similar conserved hydratase motif present in the structure **(Figure 4.6a)**.

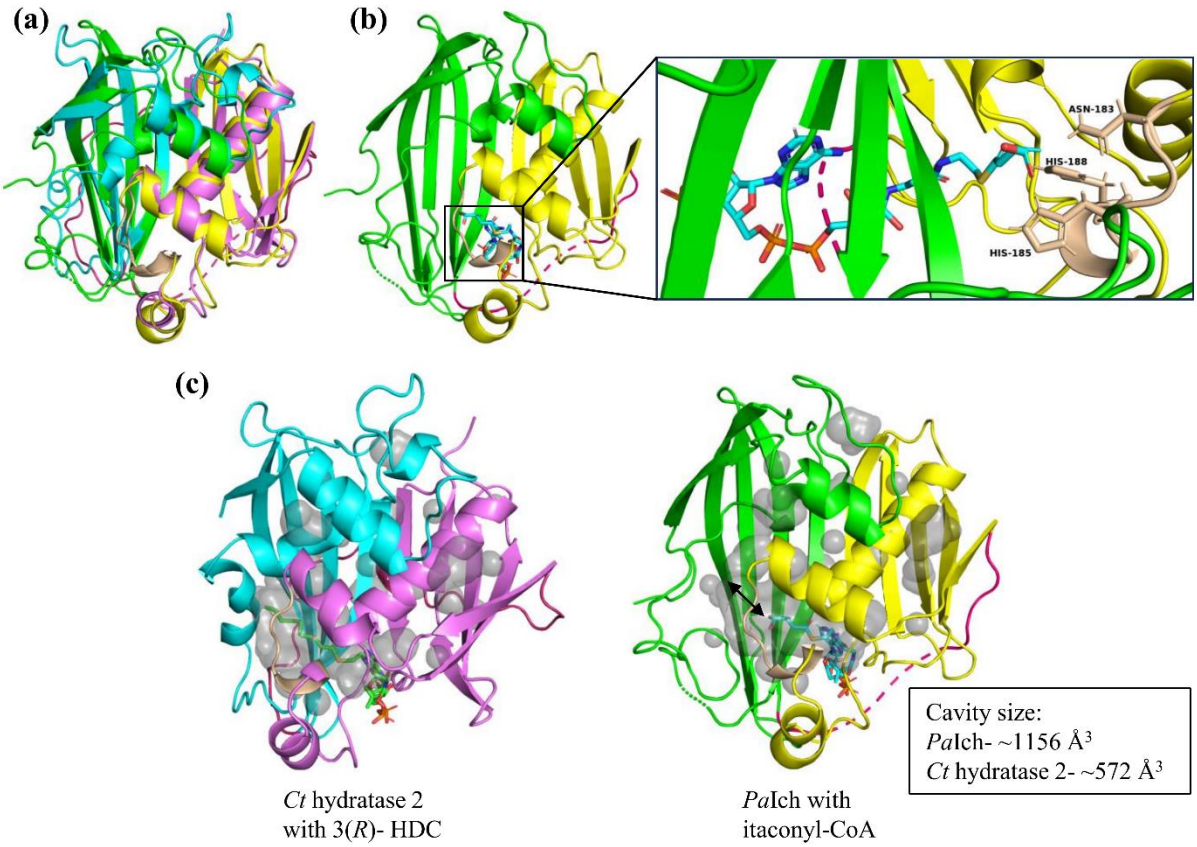


Figure 4.6: The enzymatic active site and substrate binding tunnel of *Pa*Ich. (a) Superposition of *Pa*Ich (N-terminal domain: green and C-terminal domain: yellow) with *Ct* hydratase 2 (N-terminal domain: cyan and C-terminal domain: violet) showed the similar conserved hydratase motif (wheat) responsible for catalysis. (b) A complex of *Pa*Ich with itaconyl-CoA was modelled from its most closely related homologue *Ct* hydratase 2 bound 3(*R*)-HDC complex due to structural similarity of 3(*R*)-HDC with itaconyl-CoA. Inset: The three putative catalytic residues are Asn_{183}, His_{185}, and His_{188} (yellow) taking part in hydratase reaction. (c) Surface representation of active site pocket (grey) of *Pa*Ich consists of an extra region marked as an arrow beyond its C_5 acyl chain of substrate (itaconyl-CoA). It predicts that *Pa*Ich can bind longer acyl chain derivatives of CoA while in *Ct* hydratase 2 complex, length of the tunnel (grey) is limited to its substrate's acyl chain length. All structural representations were made through Pymol v2.5.4.

This conserved hydratase motif is not only present in *Ct* hydratase 2 but it was consistently seen in other DHF-containing dehydratases from *P. capsici*, *M. absecssus* and *C. auranticus* (PDB ID: 3KH8, 5I7N, and 4E3E) [187]. Based on structural superposition, this hydratase motif is present at the interface of double hotdog fold of *Pa*Ich. This conserved hydratase motif in *Pa*Ich is denoted as α11 and is responsible for hydration of itaconyl-CoA. All SHF- and DHF-containing hydratases (PDB IDs: 3KH8, 5I7N, and 4E3E) showed no active site cleft in the C-terminal domain core region due to its higher compactness. Multiple sequence alignment by clustalW and structure-based sequence alignment by DALI analysis of *Pa*Ich with other DHF-containing (*R*)-specific enoyl-CoA hydratases/dehydratases also identified the conserved active site residues in this hydratase motif responsible for catalysis **(Figure 4.7)**. Active site architecture of (3*R*)-hydroxydecanoyl-CoA (HDC) bound complex of *Ct* hydratase 2 (PDB ID: 1PN4) also confirmed the catalytic residues responsible for this hydration. Asp_{182}, Asn_{184}, and Gln_{187} are three catalytic residues in *Ct* hydratase 2 responsible for hydration of 3(*R*)-HDC [119]. Multiple sequence alignment analysis showed that Asn_{183}, His_{185}, and His_{188} are three putative catalytic residues in *Pa*Ich responsible for hydration to the C3 position of itaconyl-CoA **(Figure 4.7)**. Due to similarity of 3(*R*)-HDC with itaconyl-CoA, a model complex of PaIch with itaconyl-CoA was generated based on *Ct* hydratase 2 bound 3(*R*)-HDC complex (PDB ID:1PN4) **(Figure 4.6b)**. *Pa*Ich catalyzes the conversion of substrate itaconyl-CoA to citramalyl-CoA upon addition of a water molecule, which is further hydrolyzed by (*S*)-citramalyl-CoA lyase (Ccl), the last enzyme of the itaconate degradation pathway.

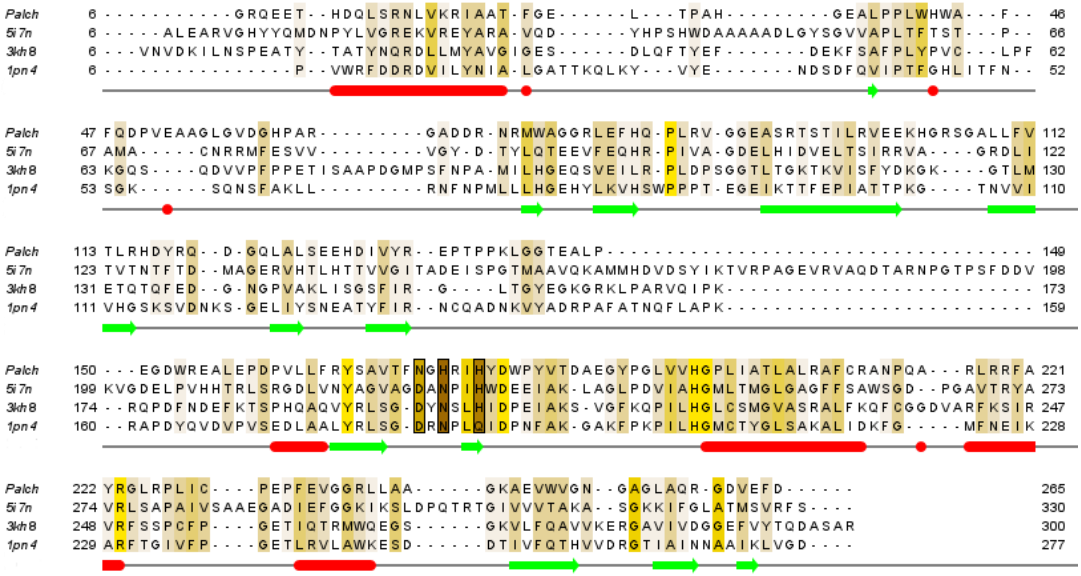


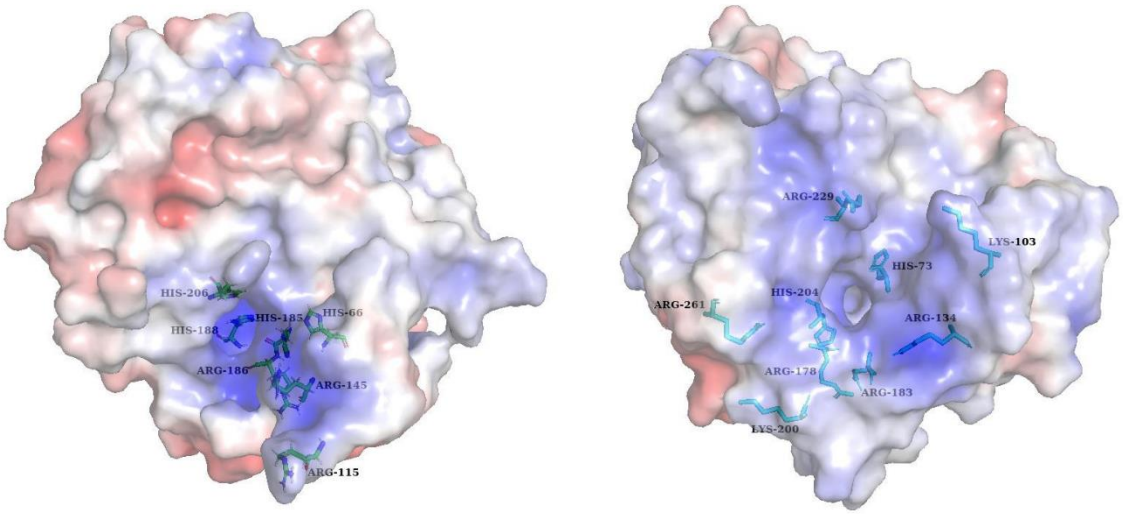
Figure 4.7: Multiple sequence alignment of *Pa*Ich. Multiple sequence alignment of *Pa*Ich with other DHF-containing (*R*)-specific enoyl-CoA hydratases by ClustalW (PDB ID: 5I7N, 3KH8, 1PN4) and pictorial representation was made through Jalview. Residues of these alignment are coloured based on their coservation. Secondary structure prediction analysis showed cylinder (red) and arrow (green) as α-helix and β-sheet respectively. Similar hydratase motif consisting of conserved active site residues marked under square box (black).

The position of side chain of putative catalytic residues was also found to be similar to the side chain of active site residues in *Ct* hydratase 2. It is suggested that amide group of Gly_{207} forms an H-bonding interaction with carbonyl oxygen of acyl chain of itaconyl-CoA by creating an oxyanion hole which ultimately initiates the process of catalysis. The phosphorylated ADP-moiety of model complex of itaconyl-CoA bound *Pa*Ich is located outside of the substrate binding tunnel in a bent conformation, like the 3(*R*)-HDC bound *Ct* hydratase 2 complex, whereas the acyl chain is positioned inside the central cavity of double hotdog fold. The side chain of Lys_{112} forms another polar contact with this ADP-moiety, and an H-bonding interaction is seen between the -NH2 group of the adenine moiety and the carbonyl oxygen of Gly_{234}.

The size of active site pocket showed to be different in *Pa*Ich, whereas in other DHF-containing (*R*)-hydratases it is similar to the size of their substrate. Remarkably, it was first observed that a deep slender substrate binding tunnel greater than the expected length of C_5 acyl chain of itaconyl-CoA was seen in *Pa*Ich. In contrast, the range of substrate binding tunnels is limited to the acyl chain length of substrate 3(*R*)-HDC of *Ct* hydratase 2 (PDB ID: 1PN2) complex. The most remarkable feature found was that the tunnel consisting of C10 acyl chain of HDC in *Ct* hydratase 2 is smaller than the tunnel consisting of C5 acyl chain of itaconyl-CoA in *Pa*Ich. No difference in the length of the tunnel was observed between apo form (PDB ID: 1PN2) and 3(*R*)-HDC bound complex of *Ct* hydratase 2 (PDB ID: 1PN4). In other hydratases/dehydratases, the path of this tunnel was restricted by the central helix (sausage) followed by a long-disordered loop region of N-terminal hotdog fold. In contrast, no such restriction was observed in *Pa*Ich. α2 and α10 helices of chain C pulled out the central helical segment α3 (Trp_{47}-Ala_{50}) represented as eaten sausage followed by subsequent loop region (Phe_{51}-Gly_{65}) of chain A and vice versa, which ultimately led to clearance of the path of the tunnel in *Pa*Ich. So, it may be proposed that active site pocket of *Pa*Ich can bind longer acyl chain derivatives of CoA than C_5 acyl chain of itaconyl-CoA. The entrance of the

tunnel is identical to other DHF-containing hydratases/dehydratases, despite the tunnel lengths differing in *Pa*Ich with other hydratases.

Electrostatic potential map of *Pa*Ich and the *Ct* hydratase 2 also showed the positively charged residues lining the active site pocket region **(Figure 4.8)**. Compared to *Ct* hydratase 2, the tunnel mouth of *Pa*Ich is somewhat wider. His_{66}, His_{185}, Arg_{186}, His_{188} and His_{206} of *Pa*Ich contribute the positively charged surface to the inner lining of the tunnel whereas Arg_{145} contributes to the mouth position. The residues positioned at the entry of the tunnel are not so well conserved between them and other (*R*)-specific enoyl-CoA hydratases. Leu_{235}, Arg_{115}, Arg_{145}, and Leu_{119} are situated at the mouth of this tunnel whereas Ile_{233}, Leu_{72}, Phe_{132}, and Phe_{198} are found to be positioned in that region of *Ct* hydratase 2. The β4 and β14 strands of chain A in *Pa*Ich are parallelly positioned at the two-domain interface, likely similar to other hydratases/dehydratases. To create the mouth of the tunnel, the C-terminal end of β4 splayed apart from the N-terminal end of β14 strand. The *Pa*Ich tunnel entrance was found to differ significantly from those of other DHF-containing hydratases. The N-terminal end of β6 strand of *Ct* hydratase 2 (PDB ID: 1PN2) was more bent toward the core region of N-terminal domain compared to the corresponding β4 strand of *Pa*Ich. At the opening of mouth in *Pa*Ich, Asn_{80}-Trp_{83} of β4 strand produced a slight kink interiorly to make a path for entry of the itaconyl-CoA whereas, in *Ct* hydratase 2, Leu_{71}-His_{73} made such a kink. Therefore, a larger arc was needed to allow longer acyl chain substrates like C10-acyl chain of 3(*R*)-HDC in *Ct* hydratase 2 (PDB ID: 1PN4) rather than a shorter arc for C5-acyl chain of itaconyl-CoA in *Pa*Ich.



*Pa*Ich *Ct* hydratase 2

Figure 4.8: Electrostatic surface potential map of *Pa*Ich and *Ct* hydratase 2. Key residues (stick representation) are involved in positive charge distribution in active site pockets.

4.3.7. Affinity towards acetoacetyl-CoA

Previous studies showed that certain CoA derivatives function as inhibitors for (*R*)-specific enoyl-CoA hydratases [188]. Of them, acetoacetyl-CoA showed inhibitory effects on crotonase and mitochondrial enoyl-CoA hydratase (PDB ID: 1DUB) in bovine and rat liver, respectively [171, 172]. Due to structural similarity of itaconyl-CoA with acetoacetyl-CoA, we first tried to find the binding affinity of *Pa*Ich with acetoacetyl-CoA. However, our findings revealed a significant binding affinity of acetoacetyl-CoA with *Pa*Ich. The dissociation constant (K_d) value of *Pa*Ich with acetoacetyl-CoA was found to be 0.214 µM using isothermal titration calorimetry (ITC) experiments **(Figure 4.9)**.

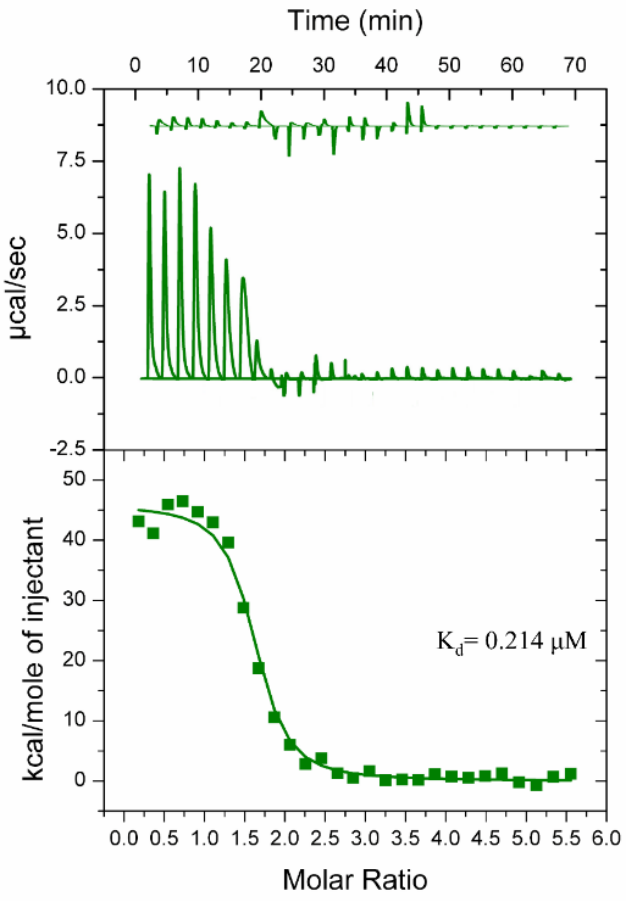


Figure 4.9: Isothermal titration calorimetry (ITC) analysis between *Pa*Ich and acetoacetyl-CoA. Acetoacetyl-CoA showed a significant binding affinity of K_d = 0.214 µM. The ITC experiments were carried out at a constant temperature of 25°C. A control experiment was carried out where *Pa*Ich and buffer (25 mM Tris; pH 8.0 and 150 mM NaCl) were given in cell and syringe respectively. The binding isotherm for control experiment was shown above the experimental data. The raw data were analyzed through ORIGIN.

4.4. Discussion

Over the past few years, the significance of the itaconate degradation pathway in pathogenic bacteria has become increasingly apparent. Previously, structural insights into itaconyl-CoA remained obscure, but two other enzymes, Ict and Ccl, of the itaconate degradation pathway have been investigated so far. In this work, we have determined the first crystal structure of *Pa*Ich which catalyzes the conversion of itaconyl-CoA to citramalyl-CoA and the functional insights about its active sites and respective substrate binding tunnel. The 1.9 Å crystal structure of *Pa*Ich resembles MaoC-like hydratase/dehydratase subfamily proteins consisting of double hotdog folds connected by a long flexible loop. Structural model of *Yp*Ich sheds light on its basic structure and shows that it consists of only one domain similar to the C-terminal domain of *Pa*Ich. So, it is deduced that the structure of *Yersinia* homologue *Yp*Ich shares 50% structural similarities with *Pa*Ich but both show similar functions. The N-terminal domain of *Pa*Ich is supposed to have evolved via gene duplication from single hotdog fold of *Yersinia* or any other hydratases/dehydratases. The number and arrangement of strands in each monomer of a dimeric unit showed differences, which confers pliability that is not restricted within helix/loop region but is also present in main frame β-sheet. We assume that each monomer modulates the main frame architecture of hotdog fold of another monomer while keeping itself fixed in a particular state.

The new N-terminal hotdog fold consisting of a short helical segment α3 pulled out by α2/α10 helices of its dimeric counterpart creates a long substrate binding tunnel. So, two protomers may cooperatively influence the binding of longer acyl chain CoA substrate which is not seen in other hydratases. The sizes of active site pocket of *Ct* hydratase 2 and *Pa*Ich measured by PyVol v1.7.8 analysis are ~572 Å3 and ~1156 $Å^3$ respectively. This kind of short helix represented as eaten sausage in *Pa*Ich, most likely arose due to helix-loop transition of 6-10 residues which might be facilitated by the pulling effect of α2/α10 helices of counterpart chain. Despite slippage of α3 helix from the β-sheet scaffold, the N-terminal hotdog fold still holds similar characteristic features of bun as other hotdog folds. It may be concluded that all MaoC-like hydratases/dehydratases consist of a similar β-sheet scaffold (bun) which is conserved throughout all domains of life while central helix α3 in *Pa*Ich is found differently as eaten sausage.

It has been found previously that *Pa*Ich of itaconate degradation pathway catalyzes the conversion of itaconyl-CoA to (*S*)-citramalyl-CoA through an unstable intermediate,

mesaconyl-C4-CoA. It was reported that it carries out two partial activities as itaconyl-CoA-isomerase and mesaconyl-CoA-hydratase where itaconyl-CoA-isomerase first transforms 3-cis itaconyl-CoA to 2-trans mesaconyl-C4-CoA and later mesaconyl-CoA-hydratase converts this transient intermediate to (*S*)-citramalyl-CoA upon addition of a water [53, 189]. At first, a proton is abstracted from C2 of itaconyl-CoA and re-added to C5 by carbonyl amide of Gly_{207} of itaconyl-CoA-isomerase/*Pa*Ich to produce mesaconyl-C4-CoA. Later, three putative catalytic residues Asn_{183}, His_{185} and His_{188} of mesaconyl-CoA-hydratase/*Pa*Ich made an H-bonding network with a catalytic water molecule which further catalyzes the reaction from mesaconyl-C4-CoA to (*S*)-citramalyl-CoA by a nucleophilic attack of OH- at C3 and addition of a proton to C2 **(Figure 4.10)**.

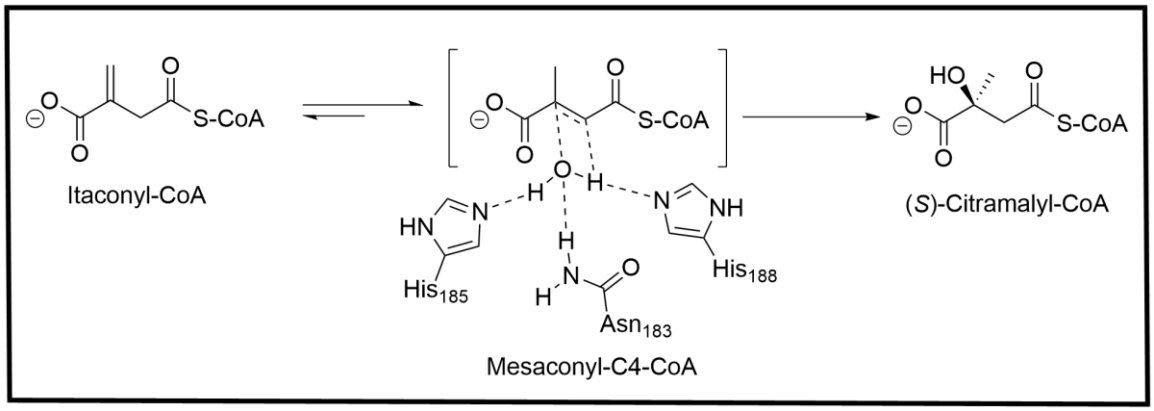


Figure 4.10: Proposed catalytic mechanism of *Pa*Ich. It converts itaconyl-CoA to citramalyl-CoA through a transient intermediate mesaconyl-C4-CoA. Asn_{183}, His_{185} and His_{188} are three catalytic residues making an H-bonding network with a water molecule which further adds OH^- ion at C3 and a proton at C2 to produce the final product (*S*)-citramalyl-CoA.

So, this study on structural as well as mechanistic insights of *Pa*Ich connects the missing link of itaconate degradation pathway in virulent organisms like *Pseudomonas aeruginosa*. So, the structural information about active site pocket with putative catalytic residues will help to find more potent drug candidates against *Pa*Ich in future. More importantly, future drug discovery based on structural insights from *Pa*Ich will ultimately replenish the defensive role of macrophages and prevent the abusive usage of antibiotics.

Chapter 5

Structural and Molecular Docking Studies of PvrA – a TetR family Transcriptional Regulator

5.1. Introduction

P. aeruginosa regulates global gene expression during infection in the host to adapt to the host environment and escape from host immune attacks. *P. aeruginosa* possesses a number of regulatory genes that are crucial for regulating other virulence genes during host infection. However, many of these regulatory factors contributing to bacterial virulence remain undetermined. Upregulation of these regulatory genes during infection indicated that they were involved in bacterial virulence in the host environment. To confirm their role in pathogenesis, several mutants of each individual gene were obtained from the PA14 transposon (Tn) mutant library [190]. Deletion mutants including ΔPA2957, ΔPA0253, ΔPA3616, and ΔPA0707 were created in the *P. aeruginosa* PA14 background for those that were unavailable in the Tn mutant library. Mice were infected with each of these deletion mutants which showed that ΔPA2957 strain caused significantly lower bacterial colonization compared to wild-type PA14 cells. In addition, mutation of PA2957 did not alter microbial growth rate in LB. Therefore, it suggested that virulence attenuation might be due to reduced virulence gene expression. So, Pan et al. designated this PA2957 gene as *pvrA* (*Pseudomonas* virulence regulator A) [161]. PvrA belongs to TetR family transcriptional regulator which regulates several vitulent genes. Notably, potential regulatory genes of PvrA such as *plcH*, *fadD1*, *fadD6* and *PA0508* aid in the utilization of phosphatidylcholine (PC) from the host, mainly from lung tissues [132, 133, 191]. PC is broken down into phosphorylcholine, fatty acids and glycerol by phospholipase C, which is encoded by the plcH [55]. Fatty acyl-CoA synthetase is encoded by both *fadD1* and *fadD6*, and PA0508 is quite similar to *fadE* of *E. coli* [134, 192]. Additionally, other studies revealed that PvrA can bind to aprA, glcB, and maeB promoter regions. A crucial enzyme in the glyoxylate shunt, malate synthase, is encoded by the gene glcB, whereas the malic enzyme is encoded by maeB controlling the amount of malate in the glyoxylate cycle and initiating gluconeogenesis from the glyoxylate shunt [193, 194]. The *aprA* encodes an extracellular alkaline metalloproteinase that contributes to bacterial virulence [195, 196].

To further investigate the role of PvrA as a transcriptional regulator in bacterial virulence, EMSA studies were performed to confirm the binding of the identified promoter [161]. Band shifts obtained through EMSA analysis also confirmed the promoter regions of *plcH*, *fadD1*, *fadD6*, *PA0508*, *glcB*, *maeB* and *aprA*. In the cytoplasm of bacteria, fatty acids are first converted into fatty acyl-CoA, which is then catabolized by the oxidation process in bacterial use of fatty acids [191]. The long-chain fatty acyl-CoA in *E. coli* binds to FadR and

influences its regulatory function [141, 142, 197]. Therefore, Pan et al. showed that palmitoyl-CoA had significant binding affinity to PvrA obtained through ITC analysis.

5.2. Materials and methods

5.2.1. Cloning, expression and purification

PvrA was cloned, heterologously produced and purified to extreme homogeneity prior to crystallization. Full-length gene was amplified using Nde1 and BamH1 as restriction sites and the following primers used in this study are 5′-CGCCGGAATTCGATGCAGAAAGAGCCT CGCAAAGTTCGTG- 3′ as forward primer and 5′- CGCCCAAGCTTTTAGGCGCTGGG AGCGTCGCCTTC- 3′ as reverse primer. The PCR product was cloned into N-terminal His_6-tagged pETduet 1 vector (Novagen) using the same restriction site. Culture was grown in 2 litres of LB supplemented with 100 mg/ml (final conc.) ampicillin at 37°C in a continuous shaking condition until the O.D. reached ~0.6-0.8. Then the culture was induced with 0.5 mM IPTG at 37°C for overnight in a shaking condition. Cells were harvested at 6000 rpm for 10 min and pellets were dissolved in re-suspension buffer (25 mM Tris; pH 8.0, 150 mM NaCl, and 5 mM Imidazole) with 1mM PMSF prior to sonication. Lysed cells were then centrifuged at 12,000 rpm for 1 hour and the supernatant was then loaded into Ni^{2+}-NTA column which was pre-equilibrated with equilibration buffer (25 mM Tris; pH 8.0, 150 mM NaCl, and 10 mM Imidazole). 2 column volume of wash 1 buffer (25 mM Tris; pH 8.0, 150 mM NaCl, and 35 mM Imidazole) followed by 3 ml of wash 2 buffer (25 mM Tris; pH 8.0, 150 mM NaCl, and 50 mM Imidazole) was added. After washing, protein was eluted with elution buffer (25 mM Tris; pH 8.0, 150 mM NaCl, and 250 mM Imidazole). The protein was then loaded to Hiload 16/60 Superdex 75 prep grade column pre-equilibrated with gel filtration buffer (25 mM Tris; pH 8.0, 150 mM NaCl). Eluted protein was collected and concentrated in a 30 kDa cut-off spin concentrator until the concentration reached ~7 mg/ml. The purity of the protein was then checked by SDS-PAGE analysis.

5.2.2 Crystallization, data collection and processing

Purified protein was crystallized in sitting drop vapour-diffusion method. Initial crystal hit was obtained in 30% PEG 3000, 100 mM Tris pH 7.0, and 200 mM NaCl of Wizard 2 screen and crystals appeared in a thin plate-like morphology stacked one after another. The whole setup was incubated at 20°C. Single and large crystals were grown within 1 week and obtained in 30% PEG 3000, 100 mM Tris pH 8.4 and 200 mM NaCl. Datasets for both native

and heavy atom derivatives were collected in Bruker D8 Venture using Cu K_α source (~1.54 Å) and Photon III CCD detector was used here for collecting frames **(Table 5.1).** The data were integrated and scaled in Proteum 3 software and were found to belong to space group C 2 2 21 with unit cell parameter *a*=93.30 Å, *b*= 96.58 Å, *c*= 104.79 Å and α,β,γ= 90°. Initially, the coordinate of heavy atom position and occupancy was calculated in Phenix by Hyss method [173, 174]. Estimation of phasing power was calculated by Autosol and initially, the model was built by Autobuild in Phenix respectively [175-177]. Successive model building and refinement were done in Coot and Phenix respectively [178]. Structure-based sequence alignment was done by DALI analysis [179]. All structural representations were made through Pymol. The atomic coordinates for *Pa*Ich have been deposited in the Protein Data Bank **(PDB ID: 8I2K)**.

Table 5.1: Data collection and Refinement Statistics

Dataset	Native	K_2PtCl_6	C_2H_5HgCl
Data collection and phasing statistics			
Wavelength (Å)	1.54	1.54	1.54
Resolution range (Å)[a]	31.96-2.30	26.48-3.30	27.89-3.27
	(2.38-2.30)	(3.42-3.30)	(3.39-3.27)
Spacegroup	C 2 2 21	C 2 2 21	C 2 2 21
Unit cell dimensions (a,b,c=Å; α,β,γ=°)	*a*=93.30, *b*= 96.58, *c*= 104.79 and α,β,γ= 90°	a=93.52, b=97.23, c=105.41 and α,β,γ= 90°	A=94.22, b=97.12, c=105.47 and α,β,γ= 90°
Total reflections[a]	395833 (28990)	51240 (3633)	148101 (10854)
Unique reflections[a]	21122 (2059)	7459 (717)	13533 (1293)
Redundancy	18.7	6.9	10.9
Completeness (%)[a]	99.2 (98.60)	99.90 (99.20)	99.8 (98.60)
Overall I/σ[a]	28 (7.3)	10.0 (2.9)	21 (7.0)
R_{merge} (%)[a]	10.1 (44.7)	17.7 (53.4)	8.7 (23.3)
Refinement statistics			
Resolution range (Å)	24.40-2.30		
Number of used refelections	20894		
R_{work} (%)	19.18		
R_{free} (%)	25.40		
Total number of atoms	6765		
Protein	65		
Water/glycerol/KNO_3	982/7/14		

Average B-factors (Å^2)	34.0		
Root mean square deviations			
Bonds (Å)	0.012		
Angles (°)	1.233		
Ramachandran plot			
Most favoured region (%)	94.79		
Ramachandran outliers (%)	0.52		

5.2.3. Molecular docking analysis

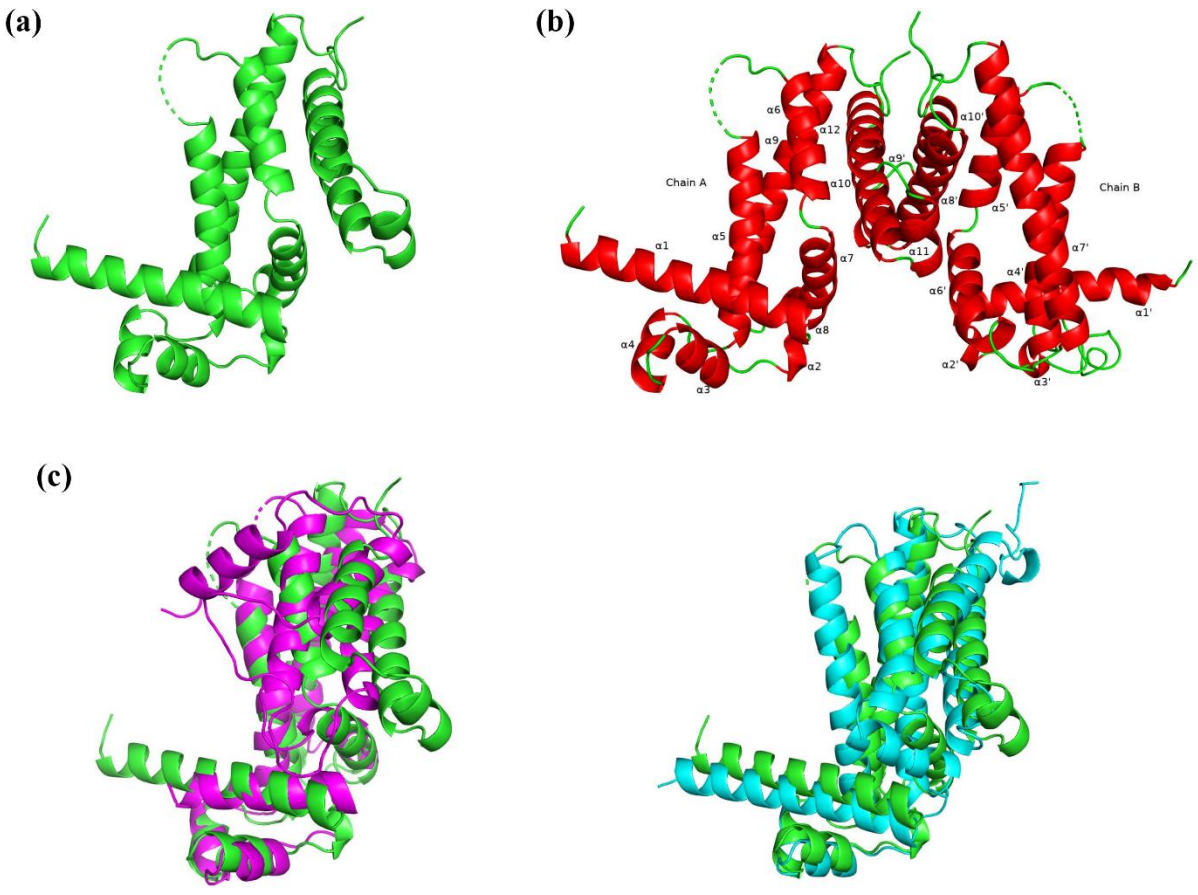
Molecules were obtained through receptor-based screening from chemical library. Molecular docking analysis was performed by Gilde in Maestro (Schrodinger). The energy grid was generated using binding site residues involved in interaction with its substrate, palmitoyl-CoA.

5.3. Results

5.3.1. Structural analysis of PvrA

Full-length N-terminal His_6-tagged PvrA (~26 kDa) was heterologously produced in *E. coli* and purified to apparent homogeneity using affinity chromatography method. The purity of the protein was further analyzed by SDS-PAGE. The protein was eluted as a dimer (~52 kDa) through size exclusion chromatography method and this higher order assembly was further confirmed by mass spectrometry analysis. Here, we solved the crystal structure of PvrA from *P. aeruginosa* at 2.30 Å resolution. Initially, molecular replacement was not successful in obtaining phase by using various structures from similar TetR family transcriptional regulators as a model. Eventually, the structure was solved by multiple isomorphous replacements with anomalous scattering (MIRAS) method using heavy-atom derivatives (Pt and Hg). The more detailed structural statistics of native and heavy-atom derivatized crystals are summarized in **Table 5.1**. In ASU, each monomer is oriented 180° toward the other. The overall structure of PvrA resembles the structures of other TetR family transcriptional regulators such as AmtR (PDB ID: 5DXZ), PfmR (PDB ID: 3VPR) and AcrR (PDB ID: 2QOP) **(Figure 5.1a)**. The crystallographic asymmetric unit (ASU) of PvrA consists of two copies (chain A and B) which is the actual biological conformation also confirmed by size exclusion chromatography method (**Figure 5.1b)**. Despite the vast sequence divergence observed between these TetR family regulators (TFRs), structural studies showed that all

TFRs family members possess a common N-terminal H-T-H DNA binding domain. Each monomer is comprised of around 10-12 helices where no β-sheet is present. The N-terminal end of both monomers starts with a long 22-24 residues H-T-H DNA binding motif. Compared to chain A, the region comprised of nearly 5 residues (Val_8-Arg_{12}) of the H-T-H motif of chain B is slightly bent downward, providing flexibility for DNA binding. Despite the low R.M.S.D. value (0.397 Å) between those two chains A and B, helix-loop/loop-helix transitions occurred between α3 and α5 helices. The most closely related structures of PvrA are AmtR and AcrR from *Corynebacterium glutamicum* and *E. coli* respectively. The R.M.S.D. values were found to be 4.72 and 5.33 Å between PvrA and AmtR/AcrR respectively (**Figure 5.1c)**.



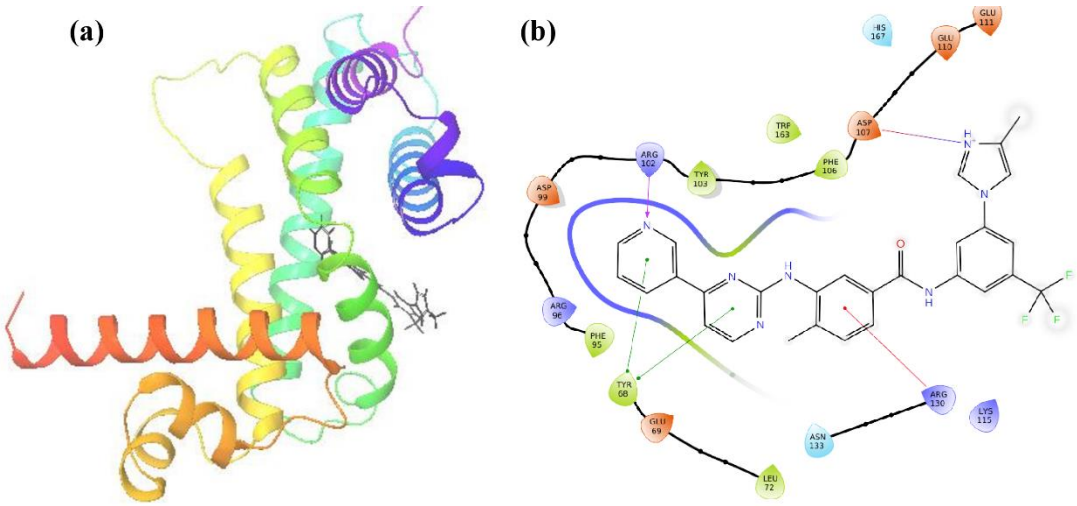
(Unpublished)

Figure 5.1: Structure of PvrA and Superposition of PvrA with closely related homologues. (a) Monomeric structure of PvrA (green). (b) Dimeric conformation of PvrA present in vivo conditions (chain A and chain B). Secondary structures are sequentially labelled where only α-helices (red) are present. (c) structural superposition of PvrA with two other closely related TFRs such as AmtR (magenta) and AcrR (cyan) respectively. All structural representations were made through Pymol.

The overall architecture of both TetR family transcriptional regulators is similar to PvrA but with some minor variations. Like PvrA, both AmtR and AcrR consist of 10-11 α-helices. However, a short 16-residue DNA binding motif was seen in AmtR compared to a 26-residue long DNA binding motif observed in AcrR. In dimeric assembly, two monomers are highly attached by mostly hydrophobic interactions. Mainly, helices 10, 11 and 12 of chain A participate in overall dimeric stability with helices 8′ and 10′ from its counterpart chain B. ~80% residues of helix 10 of chain A and helix 8′ from its counterpart chain B are hydrophobic in nature. Some polar contacts are also seen between these two monomers. Two hydrogen bond interactions are seen between side chains of Tyr_{156} of chain B and Glu_{183} and Glu_{184} from chain A respectively. The side chain amide group of Arg_{198} of chain A makes another H-bond with side chain carbonyl oxygen of Asn_{201} of chain B. Moreover, side chain carbonyl oxygen of Asn_{201} in chain A also makes a polar contact with side chain amide group of Arg_{203} in chain B. Peptide oxygen of Asn_{201} of chain A makes an H-bond with side chain amide group of Arg_{198} in chain B. Some pi-pi interactions also participate in this overall dimeric stability. One pi-pi interaction is present between side chain Phe_{188} of chain A and side chain Phe_{157} of chain B. Another pi-pi interaction exists between side chain Phe_{191} of chain A and side chain His_{158} from chain B and vice versa.

5.3.2. Interaction of Nilotinib with PvrA

The substrate binding region of PvrA is similar to other two closely related structures of AmtR and AcrR. At first, we tried to find out that PvrA has some binding affinity to other compounds rather than its substrate, palmitoyl-CoA, through computational analysis. So, we found some compounds through receptor-based screening of drug libraries by DrugRep [198]. The docking analysis was also performed by Maestro (Schrodinger). Of several compounds, nilotinib showed the highest interaction with PvrA (**Figure 5.2a)**. Topological diagram of nilotinib-bound PvrA showed the putative residues involved in this interaction. Tyr_{68}, Arg_{102}, Asp_{107}, and Arg_{130} are putative residues responsible for binding to PvrA **(Figure 5.2b)**. One salt-bridge interaction was observed between Arg_{102} and pyridine group of nilotinib molecule. Two pi-cation interactions were observed between Tyr_{68} and pyridine/pyrimidine groups of the ligand. Arg_{130} was also involved in hydrophobic interaction with nilotinib. Moreover, Asp_{107} made an H-bond interaction with imidazole moiety of nilotinib.



(Unpublished)

Figure 5.2: Interaction of PvrA with Nilotinib. (a) Cartoon representation of PvrA bound with nilotinib by docking analysis. (b) topological representation of interaction of PvrA with nilotinib molecule. Pi-cation, salt-bridge, hydrophobic and H-bond interactions are shown in green, pink, red and reddish blue.

5.4. Discussion

Several regulatory genes controlling many virulent genes during infection in the host are yet to be determined. PvrA is one such TetR family transcriptional regulatory factor highly expressed during infection and controls the expression of many virulent genes such as *plc*H, *fad*D1, *fad*D6, and PA0508. These virulent genes are highly required for bacterial survival by aiding in the utilization of PC and long-chain fatty acids in the host. In this study, we report the crystal structure of PvrA from *P. aeruginosa* at 2.30 Å resolution. The overall dimeric structure of PvrA resembles the structures of other TetR family transcriptional regulators like AmtR (PDB ID: 5DXZ), PfmR (PDB ID: 3VPR) and AcrR (PDB ID: 2QOP). Like other TetR family regulators, a common N-terminal H-T-H DNA binding motif is also seen in PvrA structure. Structural flexibility is also observed in N-terminal DNA binding motif of chain B. A short region of ~5 residues (Val_8-Arg_{12}) of chain B is slightly bent downwards compared to chain A. Most of the hydrophobic residues in the interface region participate in dimeric stability. Previous studies showed that fatty acyl-CoA acts as a modulator of PvrA which further influences its regulatory function [161]. To confirm the role of fatty acyl-CoA in controlling the function of PvrA, ITC analysis of fatty acyl-CoA such as palmitoyl-CoA

showed a significant binding towards PvrA. On the contrary, we first attempted to see if any drug molecules may show binding affinity towards PvrA through computational analysis. Among some compounds obtained through receptor-based screening from drug libraries by DrugRep, nilotinib showed highest possible interaction with PvrA by molecular docking analysis by Maestro. So, finding new drugs against these virulence-associated factors, such as PvrA, which regulates the expression of many virulent genes simultaneously will be a new pathway to control pathogenesis for invading pathogens like *P. aeruginosa* in future.

Chapter 6

Concluding Remarks

Concluding remarks:

Gram-negative *Pseudomonas* bacteria are widely distributed in the environment. This human pathogen is highly opportunistic and can cause a variety of severe acute and chronic illnesses, especially in people with weakened immune systems. Since it is the primary cause of morbidity and mortality in individuals with cystic fibrosis (CF) and one of the most common nosocomial bacteria affecting hospitalized patients, it has received special attention due to its inherent resistance to a broad spectrum of medicines. The increasing prevalence of antibiotic resistance in *P. aeruginosa*, along with its extensive distribution in hospitals, has presented a great challenge for the development of antibacterial drugs. Several virulence factors or effector molecules secreted by this pathogen during infection in the host are the primary cause of pathogenesis. Along with these virulence factors, some immunomodulatory substances also act as biomolecular weapons which modulate host immune response by targeting some essential molecules in the host, and this process is very much required for their survival during active infection. So, a strategy for the development of new drugs against these immunomodulatory virulence factors or effector molecules can prevent the dissemination of bacteria in the host and the usage of multiple drugs for treating infected patients, which ultimately restores the defensive roles of phagocytic cells such as macrophages and aids in killing them.

Among many secretory substances of phagocytic cells, such as macrophages against these invading pathogens, they secrete C_5 dicarboxylic acid itaconate, a key defense player having a role as an inhibitor of isocitrate lyase of the bacterial glyoxylate cycle. This glyoxylate cycle is very necessary for bacterial survival during infection in the host. Previous studies showed that catabolism of itaconate occurred by three-step enzymatic conversion by *P. aeruginosa* where itaconate is first converted to its corresponding CoA ester, itaconyl-CoA by succinyl-CoA:itaconate CoA transferase (Ict), then itaconyl-CoA is hydrated to (*S*)-citramalyl-CoA by (*R*)-specific itaconyl-CoA hydratase (Ich) and at last (*S*)-citramalyl-CoA is cleaved by (*S*)-citramalyl-CoA lyase (Ccl) to acetyl-CoA and pyruvate. Two enzymes, Ict and Ccl of this pathway were extensively investigated but any idea about the second enzyme, Ich, remained undetermined. The work presented in this thesis primarily focuses on structural and functional insights into the itaconyl-CoA hydratases (Ich) of itacnoate catabolic pathway. This study provides the first atomic-level insights into *Pa*Ich which fulfils the missing nexus of itaconate degradation pathway in *P. aeruginosa*. Here, we report the first three-dimensional crystal structure of *Pa*Ich at 1.98 Å resolution. In crystallogrpahic asymmetric

unit (ASU), the structure of *Pa*Ich consists of four copies of a monomeric unit (chains A, B, C, and D), where each dimer represents its actual biological confirmation, which was also confirmed by mass spectrometry analysis. In ASU, chains A and C make one dimer and chains B and D form another dimer. The overall structure of *Pa*Ich resembles the structure of MaoC-like hydratases/dehydratases (PDB ID: 1PN2, 3KH8, and 5I7N) composed of two domains, the N-terminal half and the C-terminal half connected by a long intervening bridge. Structural analysis of *Pa*Ich showed that C-terminal domain was very similar to other (*R*)-specific enoyl-CoA hydratases whereas N-terminal domain was shown to be different. The most noticeable feature observed in this dimeric assembly was that each monomer consists of two types of β-sheets with variable number of strands as well as their arrangement. It confers pliability that is not restricted within helix/loop region but is also present in main frame β-sheet.

The most remarkable feature found in this *Pa*Ich structure was a unique N-terminal hotdog fold which is not seen in other (*R*)-specific enoyl-CoA hydratases (PDB ID: 1PN2, 3KH8, and 5I7N). It is first noticed that a very short ~4 residues (Trp_{47}-Ala_{50}) α-helical segment, denoted as 'α3' which we termed an 'eaten sausage' slipped away from its core of N-terminal hotdog fold. This kind of hotdog fold is not seen in C-terminal and also other hotdog fold containing hydratases. α2 and α10 helices from each monomer are the main reason for the appearance of this kind of short helical segment 'α3' of its dimeric counterpart. Actually, α2 and α10 helices interact with the extended loop region (Phe_{51}-Gly_{65}) succeding 'α3' helical segment from its dimeric counterpart. Due to very poor contacts made between 'α3' and the surrounding β-sheet scaffold, both helices α2 and α10 pulled out the 'α3' towards them.

Superposition of *Pa*Ich with its most closely related homologue of *Ct* hydratase 2 (PDB ID: 1PN2) showed a similar conserved hydratase motif present in the structure. By multiple sequence alignment and structure-based sequence alignment of PaIch with other DHF-containing hydratases, we have found that Asn_{183}, His_{185}, and His_{188} are three putative catalytic residues of *Pa*Ich, responsible for hydration of the C3 position of its substrate, itaconyl-CoA. Although a conserved hydratase motif with active site residues is seen in all (*R*)-specific enoyl-CoA hydratases, the size of active site pocket was shown to be different in *Pa*Ich. A deep, slender substrate binding tunnel greater than the expected length of C5 acyl chain of itaconyl-CoA *Pa*Ich may propose the binding of longer acyl chain of CoA derivatives than itaconyl-CoA.

Apart from that, PvrA is a transcriptional regulator of *P. aeruginosa* which is highly expressed during bacterial infection in the host environment. Previous studies showed that it regulates the expression of many virulent genes such as *plch*, *fadD1*, *fadD6* and PA0508 which help in the utilization of phosphatidylcholine from lung tissues in cystic fibrosis patients. This virulence-associated factor such as PvrA binding with fatty acyl-CoA regulates the expression of several virulent genes during pathogenesis in the host. In this study, we report the crystal structure of PvrA from *P. aeruginosa* at 2.30 Å resolution. The overall structure of dimeric PvrA resembles the structures of other TetR family transcriptional regulators such as AmtR (PDB ID: 5DXZ), PfmR (PDB ID: 3VPR) and AcrR (PDB ID: 2QOP). Structural analysis of PvrA also showed that it possesses a conserved N-terminal H-T-H DNA binding motif similar to other DNA binding motifs in TetR family transcriptional regulators (TFRs). The region comprised of ~5 residues (Val_8-Arg_{12}) of H-T-H binding motif of chain B is slightly bent downwards compared to chain A.

.

Chapter 7

References

References:

1. Rahman, A., Sarker, M. T., Islam, M. A., Hossain, M. U., Hasan, M. & Susmi, T. F. (2023) Targeting Essential Hypothetical Proteins of Pseudomonas aeruginosa PAO1 for Mining of Novel Therapeutics: An In Silico Approach, *Biomed Res Int.* **2023**, 1787485.
2. Klockgether, J. & Tummler, B. (2017) Recent advances in understanding Pseudomonas aeruginosa as a pathogen, *F1000Res.* **6**, 1261.
3. de Sousa, T., Hebraud, M., Dapkevicius, M., Maltez, L., Pereira, J. E., Capita, R., Alonso-Calleja, C., Igrejas, G. & Poeta, P. (2021) Genomic and Metabolic Characteristics of the Pathogenicity in Pseudomonas aeruginosa, *Int J Mol Sci.* **22**.
4. Luengo, J. M. & Olivera, E. R. (2020) Catabolism of biogenic amines in Pseudomonas species, *Environ Microbiol.* **22**, 1174-1192.
5. Poulsen, B. E., Yang, R., Clatworthy, A. E., White, T., Osmulski, S. J., Li, L., Penaranda, C., Lander, E. S., Shoresh, N. & Hung, D. T. (2019) Defining the core essential genome of Pseudomonas aeruginosa, *Proc Natl Acad Sci U S A.* **116**, 10072-10080.
6. Ozer, E. A., Allen, J. P. & Hauser, A. R. (2014) Characterization of the core and accessory genomes of Pseudomonas aeruginosa using bioinformatic tools Spine and AGEnt, *BMC Genomics.* **15**, 737.
7. Mathee, K., Narasimhan, G., Valdes, C., Qiu, X., Matewish, J. M., Koehrsen, M., Rokas, A., Yandava, C. N., Engels, R., Zeng, E., Olavarietta, R., Doud, M., Smith, R. S., Montgomery, P., White, J. R., Godfrey, P. A., Kodira, C., Birren, B., Galagan, J. E. & Lory, S. (2008) Dynamics of Pseudomonas aeruginosa genome evolution, *Proc Natl Acad Sci U S A.* **105**, 3100-5.
8. Dangla-Pelissier, G., Roux, N., Schmidt, V., Chambonnier, G., Ba, M., Sebban-Kreuzer, C., de Bentzmann, S., Giraud, C. & Bordi, C. (2021) The horizontal transfer of Pseudomonas aeruginosa PA14 ICE PAPI-1 is controlled by a transcriptional triad between TprA, NdpA2 and MvaT, *Nucleic Acids Res.* **49**, 10956-10974.
9. Chandler, C. E., Horspool, A. M., Hill, P. J., Wozniak, D. J., Schertzer, J. W., Rasko, D. A. & Ernst, R. K. (2019) Genomic and Phenotypic Diversity among Ten Laboratory Isolates of Pseudomonas aeruginosa PAO1, *J Bacteriol.* **201**.
10. Mathee, K. (2018) Forensic investigation into the origin of Pseudomonas aeruginosa PA14 - old but not lost, *J Med Microbiol.* **67**, 1019-1021.
11. Diggle, S. P. & Whiteley, M. (2020) Microbe Profile: Pseudomonas aeruginosa: opportunistic pathogen and lab rat, *Microbiology (Reading).* **166**, 30-33.
12. Rossi, E., La Rosa, R., Bartell, J. A., Marvig, R. L., Haagensen, J. A. J., Sommer, L. M., Molin, S. & Johansen, H. K. (2021) Pseudomonas aeruginosa adaptation and evolution in patients with cystic fibrosis, *Nat Rev Microbiol.* **19**, 331-342.
13. Jurado-Martin, I., Sainz-Mejias, M. & McClean, S. (2021) Pseudomonas aeruginosa: An Audacious Pathogen with an Adaptable Arsenal of Virulence Factors, *Int J Mol Sci.* **22**.
14. Buehrle, D. J., Shields, R. K., Clarke, L. G., Potoski, B. A., Clancy, C. J. & Nguyen, M. H. (2017) Carbapenem-Resistant Pseudomonas aeruginosa Bacteremia: Risk Factors for Mortality and Microbiologic Treatment Failure, *Antimicrob Agents Chemother.* **61**.
15. Magill, S. S., Edwards, J. R., Bamberg, W., Beldavs, Z. G., Dumyati, G., Kainer, M. A., Lynfield, R., Maloney, M., McAllister-Hollod, L., Nadle, J., Ray, S. M., Thompson, D. L., Wilson, L. E., Fridkin, S. K., Emerging Infections Program Healthcare-Associated, I. & Antimicrobial Use Prevalence Survey, T. (2014) Multistate point-prevalence survey of health care-associated infections, *N Engl J Med.* **370**, 1198-208.
16. Fujitani, S., Sun, H. Y., Yu, V. L. & Weingarten, J. A. (2011) Pneumonia due to Pseudomonas aeruginosa: part I: epidemiology, clinical diagnosis, and source, *Chest.* **139**, 909-919.
17. Trouillet, J. L., Vuagnat, A., Combes, A., Kassis, N., Chastre, J. & Gibert, C. (2002) Pseudomonas aeruginosa ventilator-associated pneumonia: comparison of episodes due to piperacillin-resistant versus piperacillin-susceptible organisms, *Clin Infect Dis.* **34**, 1047-54.

18. Rosenthal, V. D., Al-Abdely, H. M., El-Kholy, A. A., AlKhawaja, S. A. A., Leblebicioglu, H., Mehta, Y., Rai, V., Hung, N. V., Kanj, S. S., Salama, M. F., Salgado-Yepez, E., Elahi, N., Morfin Otero, R., Apisarnthanarak, A., De Carvalho, B. M., Ider, B. E., Fisher, D., Buenaflor, M., Petrov, M. M., Quesada-Mora, A. M., Zand, F., Gurskis, V., Anguseva, T., Ikram, A., Aguilar de Moros, D., Duszynska, W., Mejia, N., Horhat, F. G., Belskiy, V., Mioljevic, V., Di Silvestre, G., Furova, K., Ramos-Ortiz, G. Y., Gamar Elanbya, M. O., Satari, H. I., Gupta, U., Dendane, T., Raka, L., Guanche-Garcell, H., Hu, B., Padgett, D., Jayatilleke, K., Ben Jaballah, N., Apostolopoulou, E., Prudencio Leon, W. E., Sepulveda-Chavez, A., Telechea, H. M., Trotter, A., Alvarez-Moreno, C., Kushner-Davalos, L. & Remaining, a. (2016) International Nosocomial Infection Control Consortium report, data summary of 50 countries for 2010-2015: Device-associated module, *Am J Infect Control.* **44**, 1495-1504.
19. Mogayzel, P. J., Jr., Naureckas, E. T., Robinson, K. A., Brady, C., Guill, M., Lahiri, T., Lubsch, L., Matsui, J., Oermann, C. M., Ratjen, F., Rosenfeld, M., Simon, R. H., Hazle, L., Sabadosa, K., Marshall, B. C. & Cystic Fibrosis Foundation Pulmonary Clinical Practice Guidelines, C. (2014) Cystic Fibrosis Foundation pulmonary guideline. pharmacologic approaches to prevention and eradication of initial Pseudomonas aeruginosa infection, *Ann Am Thorac Soc.* **11**, 1640-50.
20. Langan, K. M., Kotsimbos, T. & Peleg, A. Y. (2015) Managing Pseudomonas aeruginosa respiratory infections in cystic fibrosis, *Curr Opin Infect Dis.* **28**, 547-56.
21. Khan, T. Z., Wagener, J. S., Bost, T., Martinez, J., Accurso, F. J. & Riches, D. W. (1995) Early pulmonary inflammation in infants with cystic fibrosis, *Am J Respir Crit Care Med.* **151**, 1075-82.
22. Rosenfeld, M., Gibson, R. L., McNamara, S., Emerson, J., Burns, J. L., Castile, R., Hiatt, P., McCoy, K., Wilson, C. B., Inglis, A., Smith, A., Martin, T. R. & Ramsey, B. W. (2001) Early pulmonary infection, inflammation, and clinical outcomes in infants with cystic fibrosis, *Pediatr Pulmonol.* **32**, 356-66.
23. Flume, P. A. & Van Devanter, D. R. (2012) State of progress in treating cystic fibrosis respiratory disease, *BMC Med.* **10**, 88.
24. Furukawa, S., Kuchma, S. L. & O'Toole, G. A. (2006) Keeping their options open: acute versus persistent infections, *J Bacteriol.* **188**, 1211-7.
25. Turner, K. H., Everett, J., Trivedi, U., Rumbaugh, K. P. & Whiteley, M. (2014) Requirements for Pseudomonas aeruginosa acute burn and chronic surgical wound infection, *PLoS Genet.* **10**, e1004518.
26. MacDougall, C., Harpe, S. E., Powell, J. P., Johnson, C. K., Edmond, M. B. & Polk, R. E. (2005) Pseudomonas aeruginosa, Staphylococcus aureus, and fluoroquinolone use, *Emerg Infect Dis.* **11**, 1197-204.
27. Poole, K. (2012) Stress responses as determinants of antimicrobial resistance in Gram-negative bacteria, *Trends Microbiol.* **20**, 227-34.
28. Loebinger, M. R., Wells, A. U., Hansell, D. M., Chinyanganya, N., Devaraj, A., Meister, M. & Wilson, R. (2009) Mortality in bronchiectasis: a long-term study assessing the factors influencing survival, *Eur Respir J.* **34**, 843-9.
29. Woo, T. E., Lim, R., Surette, M. G., Waddell, B., Bowron, J. C., Somayaji, R., Duong, J., Mody, C. H., Rabin, H. R., Storey, D. G. & Parkins, M. D. (2018) Epidemiology and natural history of Pseudomonas aeruginosa airway infections in non-cystic fibrosis bronchiectasis, *ERJ Open Res.* **4**.
30. Tofas, P., Samarkos, M., Piperaki, E. T., Kosmidis, C., Triantafyllopoulou, I. D., Kotsopoulou, M., Pantazatou, A., Perlorentzou, S., Poulli, A., Vagia, M. & Daikos, G. L. (2017) Pseudomonas aeruginosa bacteraemia in patients with hematologic malignancies: risk factors, treatment and outcome, *Diagn Microbiol Infect Dis.* **88**, 335-341.
31. Cattaneo, C., Antoniazzi, F., Casari, S., Ravizzola, G., Gelmi, M., Pagani, C., D'Adda, M., Morello, E., Re, A., Borlenghi, E., Manca, N. & Rossi, G. (2012) P. aeruginosa bloodstream infections among hematological patients: an old or new question?, *Ann Hematol.* **91**, 1299-304.

32. Vidaillac, C. & Chotirmall, S. H. (2021) Pseudomonas aeruginosa in bronchiectasis: infection, inflammation, and therapies, *Expert Rev Respir Med.* **15**, 649-662.
33. Veesenmeyer, J. L., Hauser, A. R., Lisboa, T. & Rello, J. (2009) Pseudomonas aeruginosa virulence and therapy: evolving translational strategies, *Crit Care Med.* **37**, 1777-86.
34. Pier, G. B. (2007) Pseudomonas aeruginosa lipopolysaccharide: a major virulence factor, initiator of inflammation and target for effective immunity, *Int J Med Microbiol.* **297**, 277-95.
35. Hanessian, S., Regan, W., Watson, D. & Haskell, T. H. (1971) Isolation and characterization of antigenic components of a new heptavalent Pseudomonas vaccine, *Nat New Biol.* **229**, 209-10.
36. Bleves, S., Viarre, V., Salacha, R., Michel, G. P., Filloux, A. & Voulhoux, R. (2010) Protein secretion systems in Pseudomonas aeruginosa: A wealth of pathogenic weapons, *Int J Med Microbiol.* **300**, 534-43.
37. Matsumoto, T., Tateda, K., Furuya, N., Miyazaki, S., Ohno, A., Ishii, Y., Hirakata, Y. & Yamaguchi, K. (1998) Efficacies of alkaline protease, elastase and exotoxin A toxoid vaccines against gut-derived Pseudomonas aeruginosa sepsis in mice, *J Med Microbiol.* **47**, 303-8.
38. Kessler, E., Safrin, M., Abrams, W. R., Rosenbloom, J. & Ohman, D. E. (1997) Inhibitors and specificity of Pseudomonas aeruginosa LasA, *J Biol Chem.* **272**, 9884-9.
39. Bradshaw, J. L., Caballero, A. R., Bierdeman, M. A., Adams, K. V., Pipkins, H. R., Tang, A., O'Callaghan, R. J. & McDaniel, L. S. (2018) Pseudomonas aeruginosa Protease IV Exacerbates Pneumococcal Pneumonia and Systemic Disease, *mSphere.* **3**.
40. Engel, L. S., Hill, J. M., Caballero, A. R., Green, L. C. & O'Callaghan, R. J. (1998) Protease IV, a unique extracellular protease and virulence factor from Pseudomonas aeruginosa, *J Biol Chem.* **273**, 16792-7.
41. Elsen, S., Huber, P., Bouillot, S., Coute, Y., Fournier, P., Dubois, Y., Timsit, J. F., Maurin, M. & Attree, I. (2014) A type III secretion negative clinical strain of Pseudomonas aeruginosa employs a two-partner secreted exolysin to induce hemorrhagic pneumonia, *Cell Host Microbe.* **15**, 164-76.
42. Lamont, I. L., Beare, P. A., Ochsner, U., Vasil, A. I. & Vasil, M. L. (2002) Siderophore-mediated signaling regulates virulence factor production in Pseudomonasaeruginosa, *Proc Natl Acad Sci U S A.* **99**, 7072-7.
43. Balasubramanian, D., Schneper, L., Kumari, H. & Mathee, K. (2013) A dynamic and intricate regulatory network determines Pseudomonas aeruginosa virulence, *Nucleic Acids Res.* **41**, 1-20.
44. Michalska, M. & Wolf, P. (2015) Pseudomonas Exotoxin A: optimized by evolution for effective killing, *Front Microbiol.* **6**, 963.
45. Verma, N., Dollinger, P., Kovacic, F., Jaeger, K. E. & Gohlke, H. (2020) The Membrane-Integrated Steric Chaperone Lif Facilitates Active Site Opening of Pseudomonas aeruginosa Lipase A, *J Comput Chem.* **41**, 500-512.
46. Zhang, Z. & Zhang, X. (2021) Evolution of Subfamily I.1 Lipases in Pseudomonas aeruginosa, *Curr Microbiol.* **78**, 3494-3504.
47. Vasil, M. L., Graham, L. M., Ostroff, R. M., Shortridge, V. D. & Vasil, A. I. (1991) Phospholipase C: molecular biology and contribution to the pathogenesis of Pseudomonas aeruginosa, *Antibiot Chemother (1971).* **44**, 34-47.
48. Terada, L. S., Johansen, K. A., Nowbar, S., Vasil, A. I. & Vasil, M. L. (1999) Pseudomonas aeruginosa hemolytic phospholipase C suppresses neutrophil respiratory burst activity, *Infect Immun.* **67**, 2371-6.
49. Konig, B., Vasil, M. L. & Konig, W. (1997) Role of haemolytic and non-haemolytic phospholipase C from Pseudomonas aeruginosa in interleukin-8 release from human monocytes, *J Med Microbiol.* **46**, 471-8.
50. Morello, E., Perez-Berezo, T., Boisseau, C., Baranek, T., Guillon, A., Brea, D., Lanotte, P., Carpena, X., Pietrancosta, N., Herve, V., Ramphal, R., Cenac, N. & Si-Tahar, M. (2019) Pseudomonas

aeruginosa Lipoxygenase LoxA Contributes to Lung Infection by Altering the Host Immune Lipid Signaling, *Front Microbiol.* **10**, 1826.
51. Lu, J., Zhang, C., Leong, H. Y., Show, P. L., Lu, F. & Lu, Z. (2020) Overproduction of lipoxygenase from Pseudomonas aeruginosa in Escherichia coli by auto-induction expression and its application in triphenylmethane dyes degradation, *J Biosci Bioeng.* **129**, 327-332.
52. Kalms, J., Banthiya, S., Galemou Yoga, E., Hamberg, M., Holzhutter, H. G., Kuhn, H. & Scheerer, P. (2017) The crystal structure of Pseudomonas aeruginosa lipoxygenase Ala420Gly mutant explains the improved oxygen affinity and the altered reaction specificity, *Biochim Biophys Acta Mol Cell Biol Lipids.* **1862**, 463-473.
53. Sasikaran, J., Ziemski, M., Zadora, P. K., Fleig, A. & Berg, I. A. (2014) Bacterial itaconate degradation promotes pathogenicity, *Nat Chem Biol.* **10**, 371-7.
54. Cuthbertson, L. & Nodwell, J. R. (2013) The TetR family of regulators, *Microbiol Mol Biol Rev.* **77**, 440-75.
55. Son, M. S., Matthews, W. J., Jr., Kang, Y., Nguyen, D. T. & Hoang, T. T. (2007) In vivo evidence of Pseudomonas aeruginosa nutrient acquisition and pathogenesis in the lungs of cystic fibrosis patients, *Infect Immun.* **75**, 5313-24.
56. Martinez-Reyes, I. & Chandel, N. S. (2020) Mitochondrial TCA cycle metabolites control physiology and disease, *Nat Commun.* **11**, 102.
57. Michelucci, A., Cordes, T., Ghelfi, J., Pailot, A., Reiling, N., Goldmann, O., Binz, T., Wegner, A., Tallam, A., Rausell, A., Buttini, M., Linster, C. L., Medina, E., Balling, R. & Hiller, K. (2013) Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production, *Proc Natl Acad Sci U S A.* **110**, 7820-5.
58. Wu, R., Chen, F., Wang, N., Tang, D. & Kang, R. (2020) ACOD1 in immunometabolism and disease, *Cell Mol Immunol.* **17**, 822-833.
59. Nonnenmacher, Y. & Hiller, K. (2018) Biochemistry of proinflammatory macrophage activation, *Cell Mol Life Sci.* **75**, 2093-2109.
60. Denko, N. C. (2008) Hypoxia, HIF1 and glucose metabolism in the solid tumour, *Nat Rev Cancer.* **8**, 705-13.
61. Meiser, J., Kramer, L., Sapcariu, S. C., Battello, N., Ghelfi, J., D'Herouel, A. F., Skupin, A. & Hiller, K. (2016) Pro-inflammatory Macrophages Sustain Pyruvate Oxidation through Pyruvate Dehydrogenase for the Synthesis of Itaconate and to Enable Cytokine Expression, *J Biol Chem.* **291**, 3932-46.
62. Zhu, X., Long, D., Zabalawi, M., Ingram, B., Yoza, B. K., Stacpoole, P. W. & McCall, C. E. (2020) Stimulating pyruvate dehydrogenase complex reduces itaconate levels and enhances TCA cycle anabolic bioenergetics in acutely inflamed monocytes, *J Leukoc Biol.* **107**, 467-484.
63. Cordes, T. & Metallo, C. M. (2021) Itaconate Alters Succinate and Coenzyme A Metabolism via Inhibition of Mitochondrial Complex II and Methylmalonyl-CoA Mutase, *Metabolites.* **11**.
64. Bambouskova, M., Gorvel, L., Lampropoulou, V., Sergushichev, A., Loginicheva, E., Johnson, K., Korenfeld, D., Mathyer, M. E., Kim, H., Huang, L. H., Duncan, D., Bregman, H., Keskin, A., Santeford, A., Apte, R. S., Sehgal, R., Johnson, B., Amarasinghe, G. K., Soares, M. P., Satoh, T., Akira, S., Hai, T., de Guzman Strong, C., Auclair, K., Roddy, T. P., Biller, S. A., Jovanovic, M., Klechevsky, E., Stewart, K. M., Randolph, G. J. & Artyomov, M. N. (2018) Electrophilic properties of itaconate and derivatives regulate the IkappaBzeta-ATF3 inflammatory axis, *Nature.* **556**, 501-504.
65. Swain, A., Bambouskova, M., Kim, H., Andhey, P. S., Duncan, D., Auclair, K., Chubukov, V., Simons, D. M., Roddy, T. P., Stewart, K. M. & Artyomov, M. N. (2020) Comparative evaluation of itaconate and its derivatives reveals divergent inflammasome and type I interferon regulation in macrophages, *Nat Metab.* **2**, 594-602.
66. Qin, W., Zhang, Y., Tang, H., Liu, D., Chen, Y., Liu, Y. & Wang, C. (2020) Chemoproteomic Profiling of Itaconation by Bioorthogonal Probes in Inflammatory Macrophages, *J Am Chem Soc.* **142**, 10894-10898.

67. Williams, J. O., Roche, T. E. & McFadden, B. A. (1971) Mechanism of action of isocitrate lyase from Pseudomonas indigofera, *Biochemistry.* **10**, 1384-90.
68. Rittenhouse, J. W. & McFadden, B. A. (1974) Inhibition of isocitrate lyase from Pseudomonas indigofera by itaconate, *Arch Biochem Biophys.* **163**, 79-86.
69. McFadden, B. A. & Purohit, S. (1977) Itaconate, an isocitrate lyase-directed inhibitor in Pseudomonas indigofera, *J Bacteriol.* **131**, 136-44.
70. Nair, S., Huynh, J. P., Lampropoulou, V., Loginicheva, E., Esaulova, E., Gounder, A. P., Boon, A. C. M., Schwarzkopf, E. A., Bradstreet, T. R., Edelson, B. T., Artyomov, M. N., Stallings, C. L. & Diamond, M. S. (2018) Irg1 expression in myeloid cells prevents immunopathology during M. tuberculosis infection, *J Exp Med.* **215**, 1035-1045.
71. Naujoks, J., Tabeling, C., Dill, B. D., Hoffmann, C., Brown, A. S., Kunze, M., Kempa, S., Peter, A., Mollenkopf, H. J., Dorhoi, A., Kershaw, O., Gruber, A. D., Sander, L. E., Witzenrath, M., Herold, S., Nerlich, A., Hocke, A. C., van Driel, I., Suttorp, N., Bedoui, S., Hilbi, H., Trost, M. & Opitz, B. (2016) IFNs Modify the Proteome of Legionella-Containing Vacuoles and Restrict Infection Via IRG1-Derived Itaconic Acid, *PLoS Pathog.* **12**, e1005408.
72. Ruetz, M., Campanello, G. C., Purchal, M., Shen, H., McDevitt, L., Gouda, H., Wakabayashi, S., Zhu, J., Rubin, E. J., Warncke, K., Mootha, V. K., Koutmos, M. & Banerjee, R. (2019) Itaconyl-CoA forms a stable biradical in methylmalonyl-CoA mutase and derails its activity and repair, *Science.* **366**, 589-593.
73. Chen, M., Sun, H., Boot, M., Shao, L., Chang, S. J., Wang, W., Lam, T. T., Lara-Tejero, M., Rego, E. H. & Galan, J. E. (2020) Itaconate is an effector of a Rab GTPase cell-autonomous host defense pathway against Salmonella, *Science.* **369**, 450-455.
74. Schuster, E. M., Epple, M. W., Glaser, K. M., Mihlan, M., Lucht, K., Zimmermann, J. A., Bremser, A., Polyzou, A., Obier, N., Cabezas-Wallscheid, N., Trompouki, E., Ballabio, A., Vogel, J., Buescher, J. M., Westermann, A. J. & Rambold, A. S. (2022) TFEB induces mitochondrial itaconate synthesis to suppress bacterial growth in macrophages, *Nat Metab.* **4**, 856-866.
75. Zhang, Z., Chen, C., Yang, F., Zeng, Y. X., Sun, P., Liu, P. & Li, X. (2022) Itaconate is a lysosomal inducer that promotes antibacterial innate immunity, *Mol Cell.* **82**, 2844-2857 e10.
76. Spano, S. & Galan, J. E. (2012) A Rab32-dependent pathway contributes to Salmonella typhi host restriction, *Science.* **338**, 960-3.
77. Nathan, C. F. (1987) Secretory products of macrophages, *J Clin Invest.* **79**, 319-26.
78. Strelko, C. L., Lu, W., Dufort, F. J., Seyfried, T. N., Chiles, T. C., Rabinowitz, J. D. & Roberts, M. F. (2011) Itaconic acid is a mammalian metabolite induced during macrophage activation, *J Am Chem Soc.* **133**, 16386-9.
79. Hillier, S. & Charnetzky, W. T. (1981) Glyoxylate bypass enzymes in Yersinia species and multiple forms of isocitrate lyase in Yersinia pestis, *J Bacteriol.* **145**, 452-8.
80. Honer Zu Bentrup, K., Miczak, A., Swenson, D. L. & Russell, D. G. (1999) Characterization of activity and expression of isocitrate lyase in Mycobacterium avium and Mycobacterium tuberculosis, *J Bacteriol.* **181**, 7161-7.
81. Ray, S., Kreitler, D. F., Gulick, A. M. & Murkin, A. S. (2018) The Nitro Group as a Masked Electrophile in Covalent Enzyme Inhibition, *ACS Chem Biol.* **13**, 1470-1473.
82. Sharma, V., Sharma, S., Hoener zu Bentrup, K., McKinney, J. D., Russell, D. G., Jacobs, W. R., Jr. & Sacchettini, J. C. (2000) Structure of isocitrate lyase, a persistence factor of Mycobacterium tuberculosis, *Nat Struct Biol.* **7**, 663-8.
83. Kwai, B. X. C., Collins, A. J., Middleditch, M. J., Sperry, J., Bashiri, G. & Leung, I. K. H. (2021) Itaconate is a covalent inhibitor of the Mycobacterium tuberculosis isocitrate lyase, *RSC Med Chem.* **12**, 57-61.
84. Tangsudjai, S., Pudla, M., Limposuwan, K., Woods, D. E., Sirisinha, S. & Utaisincharoen, P. (2010) Involvement of the MyD88-independent pathway in controlling the intracellular fate of Burkholderia pseudomallei infection in the mouse macrophage cell line RAW 264.7, *Microbiol Immunol.* **54**, 282-90.

85. Degrandi, D., Hoffmann, R., Beuter-Gunia, C. & Pfeffer, K. (2009) The proinflammatory cytokine-induced IRG1 protein associates with mitochondria, *J Interferon Cytokine Res.* **29**, 55-67.
86. Basler, T., Jeckstadt, S., Valentin-Weigand, P. & Goethe, R. (2006) Mycobacterium paratuberculosis, Mycobacterium smegmatis, and lipopolysaccharide induce different transcriptional and post-transcriptional regulation of the IRG1 gene in murine macrophages, *J Leukoc Biol.* **79**, 628-38.
87. Adler, J., Wang, S. F. & Lardy, H. A. (1957) The metabolism of itaconic acid by liver mitochondria, *J Biol Chem.* **229**, 865-79.
88. Cooper, R. A., Itiaba, K. & Kornberg, H. L. (1965) The Utilization of Aconate and Itaconate by Micrococcus Sp, *Biochem J.* **94**, 25-31.
89. Martin, W. R., Frigan, F. & Bergman, E. H. (1961) Noninductive metabolism of itaconic acid by Pseudomonas and Salmonella species, *J Bacteriol.* **82**, 905-8.
90. Cooper, R. A. & Kornberg, H. L. (1964) The utilization of itaconate by Pseudomonas sp, *Biochem J.* **91**, 82-91.
91. Pujol, C., Grabenstein, J. P., Perry, R. D. & Bliska, J. B. (2005) Replication of Yersinia pestis in interferon gamma-activated macrophages requires ripA, a gene encoded in the pigmentation locus, *Proc Natl Acad Sci U S A.* **102**, 12909-14.
92. Eriksson, S., Lucchini, S., Thompson, A., Rhen, M. & Hinton, J. C. (2003) Unravelling the biology of macrophage infection by gene expression profiling of intracellular Salmonella enterica, *Mol Microbiol.* **47**, 103-18.
93. Shi, L., Adkins, J. N., Coleman, J. R., Schepmoes, A. A., Dohnkova, A., Mottaz, H. M., Norbeck, A. D., Purvine, S. O., Manes, N. P., Smallwood, H. S., Wang, H., Forbes, J., Gros, P., Uzzau, S., Rodland, K. D., Heffron, F., Smith, R. D. & Squier, T. C. (2006) Proteomic analysis of Salmonella enterica serovar typhimurium isolated from RAW 264.7 macrophages: identification of a novel protein that contributes to the replication of serovar typhimurium inside macrophages, *J Biol Chem.* **281**, 29131-40.
94. Zhao, Y., Jansen, R., Gaastra, W., Arkesteijn, G., van der Zeijst, B. A. & van Putten, J. P. (2002) Identification of genes affecting Salmonella enterica serovar enteritidis infection of chicken macrophages, *Infect Immun.* **70**, 5319-21.
95. Santiviago, C. A., Reynolds, M. M., Porwollik, S., Choi, S. H., Long, F., Andrews-Polymenis, H. L. & McClelland, M. (2009) Analysis of pools of targeted Salmonella deletion mutants identifies novel genes affecting fitness during competitive infection in mice, *PLoS Pathog.* **5**, e1000477.
96. Periaswamy, B., Maier, L., Vishwakarma, V., Slack, E., Kremer, M., Andrews-Polymenis, H. L., McClelland, M., Grant, A. J., Suar, M. & Hardt, W. D. (2012) Live attenuated S. Typhimurium vaccine with improved safety in immuno-compromised mice, *PLoS One.* **7**, e45433.
97. Torres, R., Swift, R. V., Chim, N., Wheatley, N., Lan, B., Atwood, B. R., Pujol, C., Sankaran, B., Bliska, J. B., Amaro, R. E. & Goulding, C. W. (2011) Biochemical, structural and molecular dynamics analyses of the potential virulence factor RipA from Yersinia pestis, *PLoS One.* **6**, e25084.
98. Goulding, C. W., Bowers, P. M., Segelke, B., Lekin, T., Kim, C. Y., Terwilliger, T. C. & Eisenberg, D. (2007) The structure and computational analysis of Mycobacterium tuberculosis protein CitE suggest a novel enzymatic function, *J Mol Biol.* **365**, 275-83.
99. Hisano, T., Tsuge, T., Fukui, T., Iwata, T., Miki, K. & Doi, Y. (2003) Crystal structure of the (R)-specific enoyl-CoA hydratase from Aeromonas caviae involved in polyhydroxyalkanoate biosynthesis, *J Biol Chem.* **278**, 617-24.
100. Erb, T. J., Berg, I. A., Brecht, V., Muller, M., Fuchs, G. & Alber, B. E. (2007) Synthesis of C5-dicarboxylic acids from C2-units involving crotonyl-CoA carboxylase/reductase: the ethylmalonyl-CoA pathway, *Proc Natl Acad Sci U S A.* **104**, 10631-6.
101. Zarzycki, J., Brecht, V., Muller, M. & Fuchs, G. (2009) Identifying the missing steps of the autotrophic 3-hydroxypropionate CO2 fixation cycle in Chloroflexus aurantiacus, *Proc Natl Acad Sci U S A.* **106**, 21317-22.

102. Khomyakova, M., Bukmez, O., Thomas, L. K., Erb, T. J. & Berg, I. A. (2011) A methylaspartate cycle in haloarchaea, *Science.* **331**, 334-7.
103. Dimroth, P., Buckel, W., Loyal, R. & Eggerer, H. (1977) Isolation and function of the subunits of citramalate lyase and formation of hybrids with the subunits of citrate lyase, *Eur J Biochem.* **80**, 469-77.
104. Torres, R., Lan, B., Latif, Y., Chim, N. & Goulding, C. W. (2014) Structural snapshots along the reaction pathway of Yersinia pestis RipA, a putative butyryl-CoA transferase, *Acta Crystallogr D Biol Crystallogr.* **70**, 1074-85.
105. Wang, H., Fedorov, A. A., Fedorov, E. V., Hunt, D. M., Rodgers, A., Douglas, H. L., Garza-Garcia, A., Bonanno, J. B., Almo, S. C. & de Carvalho, L. P. S. (2019) An essential bifunctional enzyme in Mycobacterium tuberculosis for itaconate dissimilation and leucine catabolism, *Proc Natl Acad Sci U S A.* **116**, 15907-15913.
106. Torres, R., Chim, N., Sankaran, B., Pujol, C., Bliska, J. B. & Goulding, C. W. (2012) Structural insights into RipC, a putative citrate lyase beta subunit from a Yersinia pestis virulence operon, *Acta Crystallogr Sect F Struct Biol Cryst Commun.* **68**, 2-7.
107. Shen, H., Campanello, G. C., Flicker, D., Grabarek, Z., Hu, J., Luo, C., Banerjee, R. & Mootha, V. K. (2017) The Human Knockout Gene CLYBL Connects Itaconate to Vitamin B(12), *Cell.* **171**, 771-782 e11.
108. Jacobs, M. A., Alwood, A., Thaipisuttikul, I., Spencer, D., Haugen, E., Ernst, S., Will, O., Kaul, R., Raymond, C., Levy, R., Chun-Rong, L., Guenthner, D., Bovee, D., Olson, M. V. & Manoil, C. (2003) Comprehensive transposon mutant library of Pseudomonas aeruginosa, *Proc Natl Acad Sci U S A.* **100**, 14339-44.
109. Chavez-Aviles, M., Diaz-Perez, A. L., Reyes-de la Cruz, H. & Campos-Garcia, J. (2009) The Pseudomonas aeruginosa liuE gene encodes the 3-hydroxy-3-methylglutaryl coenzyme A lyase, involved in leucine and acyclic terpene catabolism, *FEMS Microbiol Lett.* **296**, 117-23.
110. Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M. & Smith, J. L. (1996) Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site, *Structure.* **4**, 253-64.
111. Park, S. J. & Lee, S. Y. (2003) Identification and characterization of a new enoyl coenzyme A hydratase involved in biosynthesis of medium-chain-length polyhydroxyalkanoates in recombinant Escherichia coli, *J Bacteriol.* **185**, 5391-7.
112. Anderson, A. J., Haywood, G. W. & Dawes, E. A. (1990) Biosynthesis and composition of bacterial poly(hydroxyalkanoates), *Int J Biol Macromol.* **12**, 102-5.
113. Penning, T. M., Bennett, M. J., Smith-Hoog, S., Schlegel, B. P., Jez, J. M. & Lewis, M. (1997) Structure and function of 3 alpha-hydroxysteroid dehydrogenase, *Steroids.* **62**, 101-11.
114. Leenders, F., Dolez, V., Begue, A., Moller, G., Gloeckner, J. C., de Launoit, Y. & Adamski, J. (1998) Structure of the gene for the human 17beta-hydroxysteroid dehydrogenase type IV, *Mamm Genome.* **9**, 1036-41.
115. Qin, Y. M., Haapalainen, A. M., Kilpelainen, S. H., Marttila, M. S., Koski, M. K., Glumoff, T., Novikov, D. K. & Hiltunen, J. K. (2000) Human peroxisomal multifunctional enzyme type 2. Site-directed mutagenesis studies show the importance of two protic residues for 2-enoyl-CoA hydratase 2 activity, *J Biol Chem.* **275**, 4965-72.
116. Hiltunen, J. K., Wenzel, B., Beyer, A., Erdmann, R., Fossa, A. & Kunau, W. H. (1992) Peroxisomal multifunctional beta-oxidation protein of Saccharomyces cerevisiae. Molecular analysis of the fox2 gene and gene product, *J Biol Chem.* **267**, 6646-53.
117. Kiema, T. R., Engel, C. K., Schmitz, W., Filppula, S. A., Wierenga, R. K. & Hiltunen, J. K. (1999) Mutagenic and enzymological studies of the hydratase and isomerase activities of 2-enoyl-CoA hydratase-1, *Biochemistry.* **38**, 2991-9.
118. Osumi, T. & Hashimoto, T. (1979) Subcellular distribution of the enzymes of the fatty acyl-CoA beta-oxidation system and their induction by di(2-ethylhexyl)phthalate in rat liver, *J Biochem.* **85**, 131-9.

119. Koski, M. K., Haapalainen, A. M., Hiltunen, J. K. & Glumoff, T. (2004) A two-domain structure of one subunit explains unique features of eukaryotic hydratase 2, *J Biol Chem.* **279**, 24666-72.
120. Koski, K. M., Haapalainen, A. M., Hiltunen, J. K. & Glumoff, T. (2005) Crystal structure of 2-enoyl-CoA hydratase 2 from human peroxisomal multifunctional enzyme type 2, *J Mol Biol.* **345**, 1157-69.
121. Tsuge, T., Hisano, T., Taguchi, S. & Doi, Y. (2003) Alteration of chain length substrate specificity of Aeromonas caviae R-enantiomer-specific enoyl-coenzyme A hydratase through site-directed mutagenesis, *Appl Environ Microbiol.* **69**, 4830-6.
122. Leibundgut, M., Jenni, S., Frick, C. & Ban, N. (2007) Structural basis for substrate delivery by acyl carrier protein in the yeast fatty acid synthase, *Science.* **316**, 288-90.
123. Jenni, S., Leibundgut, M., Boehringer, D., Frick, C., Mikolasek, B. & Ban, N. (2007) Structure of fungal fatty acid synthase and implications for iterative substrate shuttling, *Science.* **316**, 254-61.
124. Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A. & Brinkman, F. S. (2016) Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database, *Nucleic Acids Res.* **44**, D646-53.
125. Pompilio, A., Crocetta, V., Scocchi, M., Pomponio, S., Di Vincenzo, V., Mardirossian, M., Gherardi, G., Fiscarelli, E., Dicuonzo, G., Gennaro, R. & Di Bonaventura, G. (2012) Potential novel therapeutic strategies in cystic fibrosis: antimicrobial and anti-biofilm activity of natural and designed alpha-helical peptides against Staphylococcus aureus, Pseudomonas aeruginosa, and Stenotrophomonas maltophilia, *BMC Microbiol.* **12**, 145.
126. Emerson, J., Rosenfeld, M., McNamara, S., Ramsey, B. & Gibson, R. L. (2002) Pseudomonas aeruginosa and other predictors of mortality and morbidity in young children with cystic fibrosis, *Pediatr Pulmonol.* **34**, 91-100.
127. Pesci, E. C. & Iglewski, B. H. (1997) The chain of command in Pseudomonas quorum sensing, *Trends Microbiol.* **5**, 132-4; discussion 134-5.
128. Bjarnsholt, T., Jensen, P. O., Jakobsen, T. H., Phipps, R., Nielsen, A. K., Rybtke, M. T., Tolker-Nielsen, T., Givskov, M., Hoiby, N., Ciofu, O. & Scandinavian Cystic Fibrosis Study, C. (2010) Quorum sensing and virulence of Pseudomonas aeruginosa during lung infection of cystic fibrosis patients, *PLoS One.* **5**, e10115.
129. Smith, R. S. & Iglewski, B. H. (2003) P. aeruginosa quorum-sensing systems and virulence, *Curr Opin Microbiol.* **6**, 56-60.
130. Singh, P. K., Schaefer, A. L., Parsek, M. R., Moninger, T. O., Welsh, M. J. & Greenberg, E. P. (2000) Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms, *Nature.* **407**, 762-4.
131. Palmer, K. L., Mashburn, L. M., Singh, P. K. & Whiteley, M. (2005) Cystic fibrosis sputum supports growth and cues key aspects of Pseudomonas aeruginosa physiology, *J Bacteriol.* **187**, 5267-77.
132. Kang, Y., Zarzycki-Siek, J., Walton, C. B., Norris, M. H. & Hoang, T. T. (2010) Multiple FadD acyl-CoA synthetases contribute to differential fatty acid degradation and virulence in Pseudomonas aeruginosa, *PLoS One.* **5**, e13557.
133. Zarzycki-Siek, J., Norris, M. H., Kang, Y., Sun, Z., Bluhm, A. P., McMillan, I. A. & Hoang, T. T. (2013) Elucidating the Pseudomonas aeruginosa fatty acid degradation pathway: identification of additional fatty acyl-CoA synthetase homologues, *PLoS One.* **8**, e64554.
134. Miller, R. M., Tomaras, A. P., Barker, A. P., Voelker, D. R., Chan, E. D., Vasil, A. I. & Vasil, M. L. (2008) Pseudomonas aeruginosa twitching motility-mediated chemotaxis towards phospholipids and fatty acids: specificity and metabolic requirements, *J Bacteriol.* **190**, 4038-49.
135. Bernhard, W., Wang, J. Y., Tschernig, T., Tummler, B., Hedrich, H. J. & von der Hardt, H. (1997) Lung surfactant in a cystic fibrosis animal model: increased alveolar phospholipid pool size

without altered composition and surface tension function in cftrm1HGU/m1HGU mice, *Thorax.* **52**, 723-30.
136. Bernhard, W., Hoffmann, S., Dombrowsky, H., Rau, G. A., Kamlage, A., Kappler, M., Haitsma, J. J., Freihorst, J., von der Hardt, H. & Poets, C. F. (2001) Phosphatidylcholine molecular species in lung surfactant: composition in relation to respiratory rate and lung development, *Am J Respir Cell Mol Biol.* **25**, 725-31.
137. Velasco-Garcia, R., Mujica-Jimenez, C., Mendoza-Hernandez, G. & Munoz-Clares, R. A. (1999) Rapid purification and properties of betaine aldehyde dehydrogenase from Pseudomonas aeruginosa, *J Bacteriol.* **181**, 1292-300.
138. Wargo, M. J., Szwergold, B. S. & Hogan, D. A. (2008) Identification of two gene clusters and a transcriptional regulator required for Pseudomonas aeruginosa glycine betaine catabolism, *J Bacteriol.* **190**, 2690-9.
139. Fujita, Y., Matsuoka, H. & Hirooka, K. (2007) Regulation of fatty acid metabolism in bacteria, *Mol Microbiol.* **66**, 829-39.
140. Magnuson, K., Jackowski, S., Rock, C. O. & Cronan, J. E., Jr. (1993) Regulation of fatty acid biosynthesis in Escherichia coli, *Microbiol Rev.* **57**, 522-42.
141. Cronan, J. E., Jr. (1997) In vivo evidence that acyl coenzyme A regulates DNA binding by the Escherichia coli FadR global transcription factor, *J Bacteriol.* **179**, 1819-23.
142. Raman, N., Black, P. N. & DiRusso, C. C. (1997) Characterization of the fatty acid-responsive transcription factor FadR. Biochemical and genetic analyses of the native conformation and functional domains, *J Biol Chem.* **272**, 30645-50.
143. Lowden, M. J., Skorupski, K., Pellegrini, M., Chiorazzo, M. G., Taylor, R. K. & Kull, F. J. (2010) Structure of Vibrio cholerae ToxT reveals a mechanism for fatty acid regulation of virulence genes, *Proc Natl Acad Sci U S A.* **107**, 2860-5.
144. Spector, M. P., DiRusso, C. C., Pallen, M. J., Del Portillo, F. G., Dougan, G. & Finlay, B. B. (1999) The medium-/long-chain fatty acyl-CoA dehydrogenase (fadF) gene of Salmonella typhimurium is a phase 1 starvation-stress response (SSR) locus, *Microbiology (Reading).* **145 (Pt 1)**, 15-31.
145. Ulrich, L. E., Koonin, E. V. & Zhulin, I. B. (2005) One-component systems dominate signal transduction in prokaryotes, *Trends Microbiol.* **13**, 52-6.
146. Yang, H. L., Zubay, G. & Levy, S. B. (1976) Synthesis of an R plasmid protein associated with tetracycline resistance is negatively regulated, *Proc Natl Acad Sci U S A.* **73**, 1509-12.
147. Beck, C. F., Mutzel, R., Barbe, J. & Muller, W. (1982) A multifunctional gene (tetR) controls Tn10-encoded tetracycline resistance, *J Bacteriol.* **150**, 633-42.
148. Bertrand, K. P., Postle, K., Wray, L. V., Jr. & Reznikoff, W. S. (1983) Overlapping divergent promoters control expression of Tn10 tetracycline resistance, *Gene.* **23**, 149-56.
149. Postle, K., Nguyen, T. T. & Bertrand, K. P. (1984) Nucleotide sequence of the repressor gene of the TN10 tetracycline resistance determinant, *Nucleic Acids Res.* **12**, 4849-63.
150. Wray, L. V., Jr., Jorgensen, R. A. & Reznikoff, W. S. (1981) Identification of the tetracycline resistance promoter and repressor in transposon Tn10, *J Bacteriol.* **147**, 297-304.
151. Kisker, C., Hinrichs, W., Tovar, K., Hillen, W. & Saenger, W. (1995) The complex formed between Tet repressor and tetracycline-Mg2+ reveals mechanism of antibiotic resistance, *J Mol Biol.* **247**, 260-80.
152. Ahn, S. K., Cuthbertson, L. & Nodwell, J. R. (2012) Genome context as a predictive tool for identifying regulatory targets of the TetR family transcriptional regulators, *PLoS One.* **7**, e50562.
153. Yamamoto, M., Ueda, A., Kudo, M., Matsuo, Y., Fukushima, J., Nakae, T., Kaneko, T. & Ishigatsubo, Y. (2009) Role of MexZ and PA5471 in transcriptional regulation of mexXY in Pseudomonas aeruginosa, *Microbiology (Reading).* **155**, 3312-3321.
154. Chen, C. H., Cheng, J. C., Cho, Y. C. & Hsu, W. H. (2005) A gene cluster for the fatty acid catabolism from Pseudonocardia autotrophica BCRC12444, *Biochem Biophys Res Commun.* **329**, 863-8.

155. Matsuoka, H., Hirooka, K. & Fujita, Y. (2007) Organization and function of the YsiA regulon of Bacillus subtilis involved in fatty acid degradation, *J Biol Chem.* **282**, 5180-94.
156. Anand, S., Singh, V., Singh, A. K., Mittal, M., Datt, M., Subramani, B. & Kumaran, S. (2012) Equilibrium binding and kinetic characterization of putative tetracycline repressor family transcription regulator Fad35R from Mycobacterium tuberculosis, *FEBS J.* **279**, 3214-28.
157. Zhang, Y. M., Marrakchi, H. & Rock, C. O. (2002) The FabR (YijC) transcription factor regulates unsaturated fatty acid biosynthesis in Escherichia coli, *J Biol Chem.* **277**, 15558-65.
158. Feng, Y. & Cronan, J. E. (2011) Complex binding of the FabR repressor of bacterial unsaturated fatty acid biosynthesis to its cognate promoters, *Mol Microbiol.* **80**, 195-218.
159. Zhang, Y. M., Zhu, K., Frank, M. W. & Rock, C. O. (2007) A Pseudomonas aeruginosa transcription factor that senses fatty acid structure, *Mol Microbiol.* **66**, 622-32.
160. Pan, X., Liang, H., Zhao, X., Zhang, Q., Chen, L., Yue, Z., Yin, L., Jin, Y., Bai, F., Cheng, Z., Bartlam, M. & Wu, W. (2023) Regulatory and structural mechanisms of PvrA-mediated regulation of the PQS quorum-sensing system and PHA biosynthesis in Pseudomonas aeruginosa, *Nucleic Acids Res.* **51**, 2691-2708.
161. Pan, X., Fan, Z., Chen, L., Liu, C., Bai, F., Wei, Y., Tian, Z., Dong, Y., Shi, J., Chen, H., Jin, Y., Cheng, Z., Jin, S., Lin, J. & Wu, W. (2020) PvrA is a novel regulator that contributes to Pseudomonas aeruginosa pathogenesis by controlling bacterial utilization of long chain fatty acids, *Nucleic Acids Res.* **48**, 5967-5985.
162. Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature.* **227**, 680-5.
163. Waterborg, J. H. & Matthews, H. R. (1994) The Lowry method for protein quantitation, *Methods Mol Biol.* **32**, 1-4.
164. Qin, S., Xiao, W., Zhou, C., Pu, Q., Deng, X., Lan, L., Liang, H., Song, X. & Wu, M. (2022) Pseudomonas aeruginosa: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics, *Signal Transduct Target Ther.* **7**, 199.
165. Pang, Z., Raudonis, R., Glick, B. R., Lin, T. J. & Cheng, Z. (2019) Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies, *Biotechnol Adv.* **37**, 177-192.
166. Filloux, A. (2011) Protein Secretion Systems in Pseudomonas aeruginosa: An Essay on Diversity, Evolution, and Function, *Front Microbiol.* **2**, 155.
167. Gellatly, S. L. & Hancock, R. E. (2013) Pseudomonas aeruginosa: new insights into pathogenesis and host defenses, *Pathog Dis.* **67**, 159-73.
168. Liao, C., Huang, X., Wang, Q., Yao, D. & Lu, W. (2022) Virulence Factors of Pseudomonas Aeruginosa and Antivirulence Strategies to Combat Its Drug Resistance, *Front Cell Infect Microbiol.* **12**, 926758.
169. Li, R., Zhang, P., Wang, Y. & Tao, K. (2020) Itaconate: A Metabolite Regulates Inflammation Response and Oxidative Stress, *Oxid Med Cell Longev.* **2020**, 5404780.
170. Zrieq, R., Sana, T. G., Vergin, S., Garvis, S., Volfson, I., Bleves, S., Voulhoux, R. & Hegemann, J. H. (2015) Genome-wide Screen of Pseudomonas aeruginosa in Saccharomyces cerevisiae Identifies New Virulence Factors, *Front Cell Infect Microbiol.* **5**, 81.
171. Waterson, R. M. & Hill, R. L. (1972) Enoyl coenzyme A hydratase (crotonase). Catalytic properties of crotonase and its possible regulatory role in fatty acid oxidation, *J Biol Chem.* **247**, 5258-65.
172. Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K. & Wierenga, R. K. (1996) Crystal structure of enoyl-coenzyme A (CoA) hydratase at 2.5 angstroms resolution: a spiral fold defines the CoA-binding pocket, *EMBO J.* **15**, 5135-45.
173. Grosse-Kunstleve, R. W. & Adams, P. D. (2003) Substructure search procedures for macromolecular structures, *Acta Crystallogr D Biol Crystallogr.* **59**, 1966-73.

174. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix, *Acta Crystallogr D Struct Biol.* **75**, 861-877.
175. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L. W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard, *Acta Crystallogr D Biol Crystallogr.* **65**, 582-601.
176. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L. W., Read, R. J. & Adams, P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, *Acta Crystallogr D Biol Crystallogr.* **64**, 61-9.
177. Terwilliger, T. C. (2003) Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement, *Acta Crystallogr D Biol Crystallogr.* **59**, 1174-82.
178. Emsley, P. & Cowtan, K. (2004) Coot: model-building tools for molecular graphics, *Acta Crystallogr D Biol Crystallogr.* **60**, 2126-32.
179. Holm, L., Laiho, A., Toronen, P. & Salgado, M. (2023) DALI shines a light on remote homologs: One hundred discoveries, *Protein Sci.* **32**, e4519.
180. Seifert, E. (2014) OriginPro 9.1: scientific data analysis and graphing software-software review, *J Chem Inf Model.* **54**, 1552.
181. Benning, M. M., Wesenberg, G., Liu, R., Taylor, K. L., Dunaway-Mariano, D. & Holden, H. M. (1998) The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from Pseudomonas sp. Strain CBS-3, *J Biol Chem.* **273**, 33572-9.
182. Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D. & Holden, H. M. (2003) The structure of 4-hydroxybenzoyl-CoA thioesterase from arthrobacter sp. strain SU, *J Biol Chem.* **278**, 43709-16.
183. Li, J., Derewenda, U., Dauter, Z., Smith, S. & Derewenda, Z. S. (2000) Crystal structure of the Escherichia coli thioesterase II, a homolog of the human Nef binding enzyme, *Nat Struct Biol.* **7**, 555-9.
184. Dillon, S. C. & Bateman, A. (2004) The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases, *BMC Bioinformatics.* **5**, 109.
185. Tsuge, T., Sato, S., Hiroe, A., Ishizuka, K., Kanazawa, H., Shiro, Y. & Hisano, T. (2015) Contribution of the distal pocket residue to the acyl-chain-length specificity of (R)-specific enoyl-coenzyme A hydratases from Pseudomonas spp, *Appl Environ Microbiol.* **81**, 8076-83.
186. Johansson, P., Castell, A., Jones, T. A. & Backbro, K. (2006) Structure and function of Rv0130, a conserved hypothetical protein from Mycobacterium tuberculosis, *Protein Sci.* **15**, 2300-9.
187. Halloum, I., Carrere-Kremer, S., Blaise, M., Viljoen, A., Bernut, A., Le Moigne, V., Vilcheze, C., Guerardel, Y., Lutfalla, G., Herrmann, J. L., Jacobs, W. R., Jr. & Kremer, L. (2016) Deletion of a dehydratase important for intracellular growth and cording renders rough Mycobacterium abscessus avirulent, *Proc Natl Acad Sci U S A.* **113**, E4228-37.
188. Agnihotri, G. & Liu, H. W. (2003) Enoyl-CoA hydratase. reaction, mechanism, and inhibition, *Bioorg Med Chem.* **11**, 9-20.
189. Hubbard, P. A., Yu, W., Schulz, H. & Kim, J. J. (2005) Domain swapping in the low-similarity isomerase/hydratase superfamily: the crystal structure of rat mitochondrial Delta3, Delta2-enoyl-CoA isomerase, *Protein Sci.* **14**, 1545-55.
190. Liberati, N. T., Urbach, J. M., Miyata, S., Lee, D. G., Drenkard, E., Wu, G., Villanueva, J., Wei, T. & Ausubel, F. M. (2006) An ordered, nonredundant library of Pseudomonas aeruginosa strain PA14 transposon insertion mutants, *Proc Natl Acad Sci U S A.* **103**, 2833-8.
191. Sun, Z., Kang, Y., Norris, M. H., Troyer, R. M., Son, M. S., Schweizer, H. P., Dow, S. W. & Hoang, T. T. (2014) Blocking phosphatidylcholine utilization in Pseudomonas aeruginosa, via mutagenesis

of fatty acid, glycerol and choline degradation pathways, confirms the importance of this nutrient source in vivo, *PLoS One.* **9**, e103778.
192. Wells, G., Palethorpe, S. & Pesci, E. C. (2017) PsrA controls the synthesis of the Pseudomonas aeruginosa quinolone signal via repression of the FadE homolog, PA0506, *PLoS One.* **12**, e0189331.
193. Chang, G. G. & Tong, L. (2003) Structure and function of malic enzymes, a new class of oxidative decarboxylases, *Biochemistry.* **42**, 12721-33.
194. Crousilles, A., Dolan, S. K., Brear, P., Chirgadze, D. Y. & Welch, M. (2018) Gluconeogenic precursor availability regulates flux through the glyoxylate shunt in Pseudomonas aeruginosa, *J Biol Chem.* **293**, 14260-14269.
195. Laarman, A. J., Bardoel, B. W., Ruyken, M., Fernie, J., Milder, F. J., van Strijp, J. A. & Rooijakkers, S. H. (2012) Pseudomonas aeruginosa alkaline protease blocks complement activation via the classical and lectin pathways, *J Immunol.* **188**, 386-93.
196. Casilag, F., Lorenz, A., Krueger, J., Klawonn, F., Weiss, S. & Haussler, S. (2016) The LasB Elastase of Pseudomonas aeruginosa Acts in Concert with Alkaline Protease AprA To Prevent Flagellin-Mediated Immune Recognition, *Infect Immun.* **84**, 162-71.
197. van Aalten, D. M., DiRusso, C. C., Knudsen, J. & Wierenga, R. K. (2000) Crystal structure of FadR, a fatty acid-responsive transcription factor with a novel acyl coenzyme A-binding fold, *EMBO J.* **19**, 5167-77.
198. Gan, J. H., Liu, J. X., Liu, Y., Chen, S. W., Dai, W. T., Xiao, Z. X. & Cao, Y. (2023) DrugRep: an automatic virtual screening server for drug repurposing, *Acta Pharmacol Sin.* **44**, 888-896.

Chapter 8

Publication

FEBS Letters

FEBS PRESS science publishing by scientists

RESEARCH ARTICLE

Structural and functional insights of itaconyl-CoA hydratase from *Pseudomonas aeruginosa* highlight a novel N-terminal hotdog fold

Atanu Pramanik[1] and Saumen Datta[1,2]

1 Department of Structural Biology and Bio-informatics, CSIR-Indian Institute of Chemical Biology (CSIR-IICB), Kolkata, India
2 Biological Sciences, Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India

Correspondence
S. Datta, Department of Structural Biology and Bio-informatics, CSIR-Indian Institute of Chemical Biology (CSIR-IICB), Kolkata 700032, India
Tel: +91 33 24995 896
E-mail: saumen_datta@iicb.res.in

(Received 21 July 2023, revised 29 February 2024, accepted 5 March 2024)

doi:10.1002/1873-3468.14867

Edited by Seema Mattoo

Itaconyl-CoA hydratase in *Pseudomonas aeruginosa* (*Pa*Ich) converts itaconyl-CoA to (*S*)-citramalyl-CoA upon addition of a water molecule, a part of an itaconate catabolic pathway in virulent organisms required for their survival in humans host cells. Crystal structure analysis of *Pa*Ich showed that a unique N-terminal hotdog fold containing a 4-residue short helical segment α3-, named as an "eaten sausage", followed by a flexible loop region slipped away from the conserved β-sheet scaffold, whereas the C-terminal hotdog fold is similar to all MaoC. A conserved hydratase motif with catalytic residues provides mechanistic insights into catalysis, and existence of a longer substrate binding tunnel may suggest the binding of longer CoA derivatives.

Keywords: asymmetric unit; DHF; itaconyl-CoA hydratase; *Pa*Ich; substrate binding tunnel

Itaconic acid (methylene succinic acid) is an unsaturated 1,4 dicarboxylic acid, the most promising organic acid found in soil. This compound is mainly produced by fungi like *Aspergillus terreus* as well as some other organisms such as *Escherichia coli*, and *Ustillago* spp. [1–4]. It is also produced by mammalian macrophages during active infection [5,6]. Itaconate (methylenesuccinate) is also known as a potent inhibitor of isocitrate lyase of bacterial glyoxylate cycle which is required for assimilation of acetyl-CoA as a carbon source for bacteria upon degradation from fatty acids [7–9]. Many pathogenic bacteria like *Pseudomonas* spp., *Yersinia* spp., *Micrococcus* spp., *Salmonella* spp. used itaconate as a sole carbon source and catabolized into acetyl-CoA and pyruvate for their survival in human host cells [4,10–12]. Some non-pathogenic bacteria like *Burkholderia* spp. also metabolize itaconate [13]. Itaconate utilization gene cluster was previously identified in those organisms. In this itaconate catabolic pathway, itaconate is first activated to its corresponding CoA-ester by succinyl-CoA: itaconate CoA transferase (Ict), then itaconyl-CoA is hydrated to (*S*)-citramalyl-CoA by (*R*)-specific itaconyl-CoA hydratase (Ich) and at last (*S*)-citramalyl-CoA is cleaved by (*S*)-citramalyl-CoA lyase (Ccl) to acetyl-CoA and pyruvate [14]. In *Pseudomonas aeruginosa* PAO1 strain, these three enzymes are encoded by genes named PA0882 (*Paict*), PA0878 (*Paich*), and PA0883 (*Paccl*) respectively. Similarly, the gene for

Abbreviations
ADP, adenosine di-phosphate; ASU, asymmetric unit; Blastp, basic local alignment search tool for protein; Ccl, (*S*)-citramalyl-CoA lyase; CoA, coenzyme A; DHF, double hotdog fold; HDC, hydroxydecanoyl-CoA; Hyss, hybrid substructure search; Ich, itaconyl-CoA hydratase; Ict, succinyl-CoA:itaconate CoA transferase; IPTG, isopropyl β-D-1-thiogalactopyranoside; ITC, isothermal titration calorimetry; LB, Luria broth; MIRAS, multiple isomorphous replacement with anomalous scattering; NTA, nitrilotriacetic acid; PCR, polymerase chain reaction; Pec1, *Pseudomonas aeruginosa* effector candidate 1; PEG, polyethylene glycol; PMSF, phenyl methyl sulfonyl fluoride; RMSD, root mean square deviation; SDS/PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SHF, single hotdog fold.

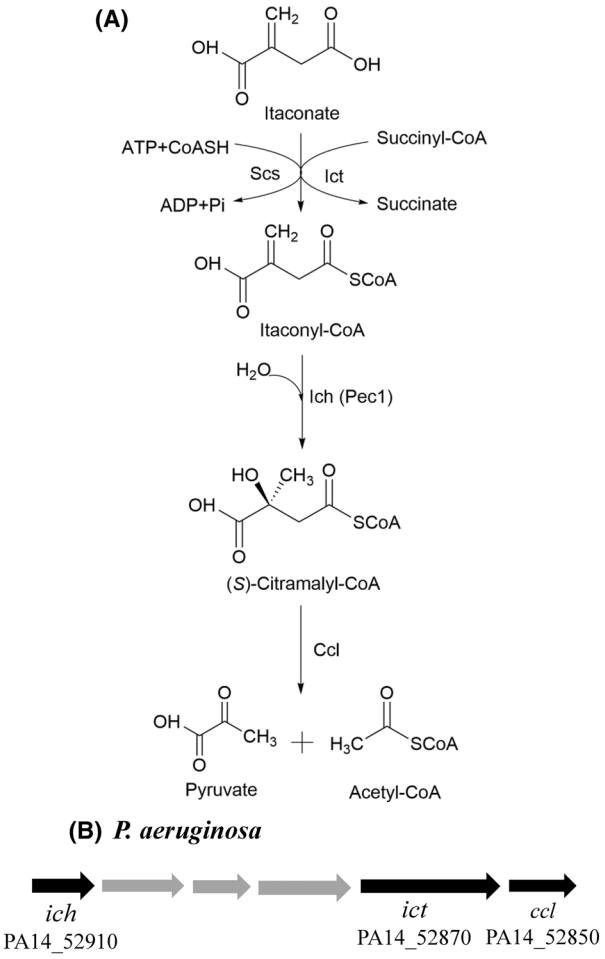


Fig. 1. Schematic representation of itaconate catabolic pathway and corresponding genetic constituents of *Pseudomonas aeruginosa*. (A) Itaconate degradation pathway showing conversion of itaconate to pyruvate and acetyl-CoA. Itaconyl-CoA hydratase (Ich) from *P. aeruginosa* also known as Pec1 catalyzes the conversion of itaconyl-CoA to citramalyl-CoA by addition of H_2O. (B) Corresponding gene clusters of itaconate degradation pathway.

*Pa*Ich product in PA14 strain was named PA14_52910 (Fig. 1). Later, Zrieq *et al.* [15] identified some virulence factors in PA14 strain by genome-wide screening of *P. aeruginosa* expression library based on yeast growth phenotype. Of them, PA14_52910 gene product was shown to have increased cellular toxicity toward macrophages and the *Caenorhabditis elegans* model. So, they called this gene product PA14_52910 (*Pa*Ich) as *P. aeruginosa* effector candidate 1 (Pec1) based on showing virulence in eukaryotic host cells. Here, in this present study, we have found that sequence of Pec1 is identical to previously described itaconyl-CoA hydratase from *P. aeruginosa* (*Pa*Ich) of itaconate catabolic pathway by blastp analysis. Thus, a crucial role of *Pa*Ich in itaconate degradation pathway as part of bacterial defense weaponry from macrophages was established in previous work. Despite this significance, structural and mechanistic insights regarding *Pa*Ich and hydration of itaconyl-CoA remain undetermined. So, In this work, we report the first three-dimensional structure of *Pa*Ich at 1.98 Å resolution. The crystal structure of *Pa*Ich shows that it is dimeric in nature confirmed by also in-solution by mass spectrometry analysis. The structure resembles MaoC-like hydratases/dehydratases consisting of N- and C-terminal domains connected by a long stretch of flexible loop. Herein, each domain represents a "hotdog fold" found previously in other (*R*)-specific enoyl-CoA hydratases [16–18]. Although C-terminal hotdog fold of *Pa*Ich was found similar to other hydratases, the N-terminal hotdog fold carries unique characteristic features which were not observed previously in any other hydratases. Unlike other hydratases, structural dynamicity is not restricted to helix/loop region but is also present in main frame β-sheet of *Pa*Ich. Based on multiple sequence alignment and structural superposition of C-terminal domain of *Pa*Ich with other DHF-containing hydratases delineate the conserved hydratase motif with catalytic residues followed by mechanistic point of view in catalysis. A remarkably long substrate binding tunnel predicts the binding of C_5 or more acyl chains of CoA derivatives. We found that it shows a significant binding affinity towards acetoacetyl-CoA like in crotonase and mitochondrial enoyl-CoA hydratase of bovine and rat liver respectively [19,20]. So, this study provides the first atomic-level insights of *Pa*Ich which fulfills the missing nexus of itaconate degradation pathway in *P. aeruginosa*.

Materials and methods

Materials

The primers used in this study were procured from (Integrated DNA Technologies, Coralville, IA, USA). *E. coli* DH5α, BL21(DE3) cells and all enzymes were purchased from Thermo Fisher Scientific, USA. All buffers and chemicals were purchased from (Sigma-Aldrich, St. Louis, MO, USA). Crystallization screens, reagents and other accessories were purchased from (Hampton Research, Aliso Viejo, CA, USA) and (Rigaku Corp., Tokyo, Japan).

Methods

Cloning, expression and purification

*Pa*Ich was cloned, heterologously produced and purified to extreme homogeneity prior to crystallization. Full-length gene was amplified using Nde1 and BamH1 as restriction sites and the following primers used in this study are

5′-TGCCATATGAGTGAGTCCGCTTTCGCC-3′ as forward primer and 5′-ATTGGATCCTTAGTCGAATTCC ACGTCGCC-3′ as reverse primer. The PCR product was cloned into N-terminal His_6-tagged pET28a vector (Novagen, St. Louis, MO, USA) using the same restriction site. This recombinant plasmid was then transformed into chemically competent BL21(DE3) cells (Thermo Fisher Scientific, Waltham, MA, USA). Culture was grown in 500 mL of LB supplemented with 100 mg·mL^{-1} (final conc.) ampicillin at 37 °C in a continuous shaking condition until the O.D. reached ~ 0.6–0.8. Then the culture was induced with 0.5 mM IPTG at 37 °C for overnight in a shaking condition. Cells were harvested at 5322 rcf for 10 min and pellets were dissolved in re-suspension buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 5 mM Imidazole and 5% glycerol) with 1 mM PMSF prior to sonication. Lysed cells were then centrifuged at 19 837 rcf for 1 h and the supernatant was then loaded into Ni^{2+}-NTA column which was pre-equilibrated with equilibration buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 10 mM Imidazole and 5% Glycerol). 1 column volume of wash 1 buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 35 mM Imidazole and 5% Glycerol) followed by 3 mL of wash 2 buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 50 mM Imidazole and 5% Glycerol) was added. After washing, protein was eluted with elution buffer (50 mM Tris; pH 8.0, 150 mM NaCl, 250 mM Imidazole and 5% Glycerol). Eluted protein was dialyzed against dialysis 1A buffer (50 mM Tris; pH 8.0, 150 mM NaCl, and 5% Glycerol) for 3–4 h. and then it was shifted to dialysis 1B buffer (50 mM Tris; pH 8.0, 150 mM NaCl, and 5% Glycerol) for another 9–10 h. The protein was then loaded to Hiload 16/60 Superdex 75 prep grade column pre-equilibrated with gel filtration buffer (50 mM Tris; pH 8.0, 150 mM NaCl, and 2% Glycerol). Eluted protein was collected and concentrated in 30 kDa cut-off spin concentrator until the concentration reached ~ 13 mg·mL^{-1}. The purity of protein was then checked by SDS/PAGE analysis.

Crystallization, data collection and processing

Purified protein was crystallized in sitting drop vapor-diffusion method. Initial crystal hit was obtained in 20% PEG 3350, 200 mM KNO_3 of Wizard 3 screen and crystals appeared in a thin plate-like morphology stacked by one after another. The whole setup was incubated at 20 °C. Single and large crystals were grown in 3–4 weeks and obtained in 22% PEG 3350, 220 mM KNO_3. Datasets for both native and heavy atom derivatives were collected in Bruker D8 Venture using Cu K_α source (~ 1.54 Å) and Photon III CCD detector was used here for collecting frames (Table 1). The data were integrated and scaled in Proteum 3 software and were found to belong to space group P1 with unit cell parameter a = 65.312 Å, b = 66.197 Å, c = 77.462 Å, α = 102.45°, β = 95.07° and γ = 101.06°. Matthew's coefficient of 2.57 Å^3·Da^{-1} assumed that there are four copies present in each asymmetric unit and the solvent content is 52.2%. Initially, the coordinate of heavy atom position and occupancy was calculated in Phenix by Hyss method [21,22]. Estimation of phasing power was calculated by Autosol and initially, the model was built by Autobuild in Phenix respectively [22–25]. Successive model building and refinement were done in Coot and Phenix respectively [26]. Structure-based sequence alignment was done by DALI analysis [27]. All structural representations were made through PYMOL [28]. The atomic coordinates for *Pa*Ich have been deposited in the Protein Data Bank (PDB ID: 8HUC).

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) was performed using Microcal VP-ITC 200 (Malvern, Enigma Business Park, Malvern, UK) with a cell volume of 350 µL. Acetoacetyl-CoA (syringe) at 45 µM was titrated against 1.75 µM of *Pa*Ich (cell) with a stirring speed of 300 rpm using titration buffer (25 mM Tris; pH 8.0, 150 mM NaCl) at 25 °C. A total of 28, 10 µL injections were made every 180 s. The raw data was analyzed in ORIGIN software [29].

Results

Overall structure of *Pa*lch

Full-length N-terminal His_6-tagged *Pa*Ich (~ 32 kDa) was heterologously produced in *E. coli* and purified to apparent homogeneity using affinity chromatography method which was further analyzed by SDS/PAGE. In size exclusion chromatography method, the protein eluted as a dimer (~ 64 kDa) which was further confirmed by mass spectrometry analysis (Fig. S1). Here, we report the first crystal structure of itaconyl-CoA hydratase from *P. aeruginosa* (*Pa*Ich) at 1.98 Å resolution. An initial attempt to get phase by molecular replacement method was not successful probably due to very low sequence similarity (< 20%) with other hydratases/dehydratases family of enzymes. Eventually, the structure was solved by multiple isomorphous replacement with anomalous scattering (MIRAS) method using heavy atom derivatives (Au and Hg). The more detailed structural statistics of native and heavy atom derivatized crystals are summarized in Table 1. The crystallographic asymmetric unit (ASU) of *Pa*Ich consists of four copies (chain A, B, C and D) of a monomeric unit forming two dimers (A–C and B–D) where each dimer represents a biologically active form (Fig. 2A,B). In ASU, two dimers are oriented at 90° to each other and each monomer of this dimer was related to each other by two-fold symmetry.

Table 1. Data collection and refinement statistics.

Dataset	Native	KAuCN	$HgCl_2$	C_2H_5HgCl
Data collection and phasing statistics				
Wavelength (Å)	1.54	1.54	1.54	1.54
Resolution range (Å)[a]	31.76–1.98 (2.05–1.98)	27.60–1.66 (1.71–1.66)	27.89–3.27 (3.39–3.27)	27.60–3.01 (3.12–3.01)
Spacegroup	P1	P1	P1	P1
Unit cell dimensions (a, b, c = Å; α, β, γ = °)	a = 65.31, b = 66.19, c = 77.46, α = 102.45, β = 95.07, and γ = 101.06	a = 65.43, b = 66.18, c = 77.60, α = 102.46, β = 95.04, and γ = 101.08	a = 65.39, b = 66.40, c = 77.43, α = 102.49, β = 95.13, and γ = 101.29	a = 65.38, b = 66.08, c = 77.34, α = 102.78, β = 95.09, and γ = 101.20
Total reflections[a]	1 608 435 (54 330)	1 279 481 (9533)	550 217 (45 323)	528 740 (47 828)
Unique reflections[a]	83 791 (7025)	137 054 (7165)	18 987 (1830)	24 258 (2396)
Redundancy	19.2	9.3	29.0	21.8
Completeness (%)[a]	98.2 (82.6)	93.0 (48.6)	99.6 (96.0)	99.8 (98.2)
Overall I/σ[a]	18 (1.5)	9.8 (2.6)	20 (8.7)	28 (12)
R_{merge} (%)[a]	12.7 (12.3)	17.5 (69.5)	18.7 (40.6)	10.4 (20.5)
Refinement statistics				
Resolution range (Å)	31.76–1.98			
Number of used reflections	77 191			
R_{work} (%)	18.22			
R_{free} (%)	22.06			
Total number of atoms	17 191			
Protein	16 186			
Water/glycerol/KNO_3	982/7/14			
Average B-factors ($Å^2$)	34.0			
Root mean square deviations				
Bonds (Å)	0.012			
Angles (°)	1.168			
Ramachandran plot				
Most favored region (%)	97.74			
Ramachandran outliers (%)	0.20			

[a]Values in parentheses refer to the highest resolution shell.

The overall structure of *Pa*Ich resembles the structure of MaoC-like hydratases/dehydratases (PDB ID: 1PN2, 3KH8, and 5I7N) composed of two domains, the N-terminal half and the C-terminal half connected by an intervening bridge (Fig. 2C). Furthermore, C-terminal domain of *Pa*Ich was found to be similar to other (*R*)-specific enoyl-CoA hydratases/dehydratases whereas N-terminal domain was shown to be different. The R.M.S.D. values of structural superposition of N-terminal domain of *Pa*Ich with 1PN2, 3KH8, and 5I7N are 2.03, 0.97 and 1.11 Å respectively. Each domain of *Pa*Ich consists of 2–4 α-helices and 5–6 continuous parallel β-strand comprising a β-sheet scaffold. This β-sheet scaffold of each domain provides the main structural framework of each monomer. The N-terminal end of *Pa*Ich consists of an overhanging segment of 3–7 residues followed by first strand of β-sheet scaffold. Each β-sheet scaffold of N- and C-terminal domains represented as a bun of the so-called "hotdog fold" which was seen in other hotdog folds containing (*R*)-specific enoyl-CoA hydratases/dehydratases (PDB ID: 1PN2, 3KH8, and 5I7N) [16,17,30–33]. A long central helix α13 (Gly_{207}-Ala_{221}) represented as a sausage was properly wrapped up by the C-terminal β-sheet scaffold constituting the C-terminal hotdog fold. On the contrary, a very short helical segment α3 (Trp_{47}-Ala_{50}) which we termed as an eaten sausage is not properly swathed by a β-sheet scaffold comprising the N-terminal hotdog fold. The architecture of C-terminal hotdog fold is very much similar to other DHF (PDB ID: 1PN2, 3KH8, and 4E3E) and also SHF-containing (*R*)-specific enoyl-CoA hydratases/dehydratases (PDB ID: 1IQ6, and 5CPG) whereas N-terminal hotdog fold has distinguished characteristic

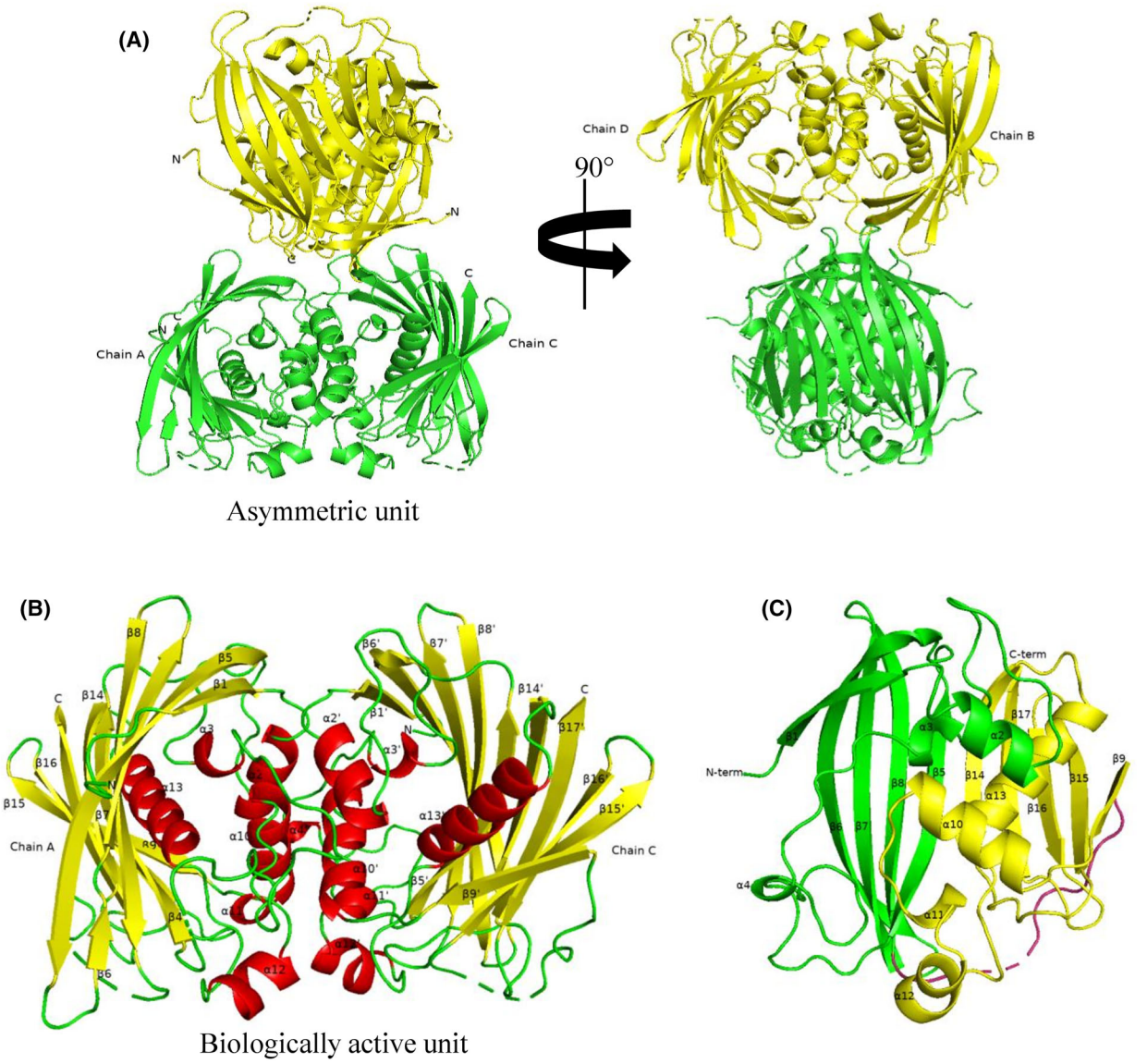


Fig. 2. Overall structure of itaconyl-CoA hydratase from *Pseudomonas aeruginosa* (*Pa*lch). (A) The whole asymmetric unit of *Pa*lch consists of a tetrameric assembly in which two dimers are positioned around 90° from each other. Chain A and chain C formed one dimer (green) whereas chain B and chain D made another dimer (yellow). A 90° rotation about *y*-axis of whole unit is shown here. (B) Overall structural view of a dimeric assembly (chain A–chain C) is represented here as a functional unit of *Pa*lch. Each α-helix and parallel β-pleated sheet are colored red and yellow respectively and they are labeled sequentially in order of their arrangement. (C) Cartoon representation of monomeric unit (chain C) showing the N-terminal and C-terminal domain comprised of a double hotdog fold. The N-terminal and C-terminal domains are colored green and yellow respectively and the intervening highly flexible segment is colored warm pink. All structural representations were made through PYMOL v2.5.4.

features [33–36]. Helix α3 (eaten sausage) and the succeeding β5 strand of N-terminal domain were connected by a long flexible loop (Phe_{51}-Arg_{79}; average b-factor 45.08 $Å^2$). An extra short helix α4 (Ala_{58}-Gly_{60}) is introduced in this loop which is found in all monomers (chain B, C and D) in the asymmetric unit except in chain A which means that helix–loop or loop–helix transition occurred in that region. Moreover, this particular Arg_{69}-Arg_{79} region of that extended loop (Phe_{51}-Arg_{79}) shows a high b-factor of ~ 81.62 $Å^2$ comprising 163 atoms. Similar to other DHF-containing hydratases, α2 (Val_{25}-Phe_{32}) helix of the N-terminal domain is stacked on α10 helix of the C-terminal domain in *Pa*Ich.

Dimeric unit

As mentioned earlier, *Pa*Ich has two dimers in the crystallographic asymmetric unit where chains A-C make one dimer and chains B-D form another dimer. The whole *Pa*Ich dimeric unit shows like a crab's shell with a dimension of ~ 77.5 × 60.8 × 54.2 Å. The most noticeable feature observed in this dimeric assembly was that each monomer consists of two β-sheets with variable number of strands as well as their arrangement (Fig. 2B). However, no significant R.M.S.D. value was observed between the monomers. Each monomer contains a total of 275 amino acid residues consisting of 6 α-helices and 11 strands of parallel β-sheet in chain A while 7 helices and 10 strands of β-sheet in its complementary chain C (Fig. 2B). Each helix and β-strand of a monomer was sequentially numbered on the basis of their arrangement (Fig. 2B, C). These strands of β-sheet range from 3 to 13 residues in length. The order of their arrangement is 1-4-5-6-7-8-9-14-15-16-17 found in chain A while β4 strand is not present in chain C (R.M.S.D value over C_α atoms between chain A and C is 0.207; Fig. 2B). However, consistency of β-sheet arrangement of A-C was not found in another dimer B-D. The arrangement of β-strands in chain B is 1-5-6-7-8-9-10-15-16-17-18-19 order while in subunit D it is found in 1-5-6-7-8-9-14-15-16-17 order. So, it concludes that each dimer of an ASU constitutes a different pattern of β-sheet arrangement (R.M.S.D value over C_α atoms between A–C and B–D is 0.271) while chain C and chain D consists of similar arrangement from dimer A–C and B–D respectively. An extra β6 strand (Ala_{118}-Leu_{120}) is formed in chain A by splitting up β7 strand of chain C (Gly_{117}-Gln_{130}). The C-terminal end of β5 followed by the N-terminal end of successive β6 strands of chain C make a bent in the opposite direction from the usual hotdog fold while it was not observed in counterpart chain A. In contrast to dimer A-C, β5 and β16 strands of chain D split to make two new strands β5-β6 and β16-β17 respectively in counter chain B. Similar to other DHF-containing hydratases, two domains of *Pa*Ich are also connected by a highly flexible solvent-exposed loop. Due to high B-factor of that region (average b-factor ~ 90 $Å^2$), the electron density of some residues (Thr_{148}-Glu_{156}) was not found.

In dimeric form, two monomers are highly attached by most of the hydrophobic residues in the interface regions and ~ 252 $Å^2$ surface area is buried by each other. The total surface area of *Pa*Ich is about 25 978 $Å^2$. Helices α2, α10, α12 and the C-terminal half of the first strand of β-sheet from both monomers participate in the dimeric stability. It is noticed that α2 (Val_{25}-Phe_{32}) helix of chain A interacts with the loop region (Phe_{51}-Glu_{57}) of its counterpart chain C whereas α2 from chain C makes contact with part of a large disordered loop region (Gln_{53}-Ala_{58}) of chain A to provide some extent of conformational stability to that region. ε-amide of Lys_{26} of chain A participates in H-bonding interaction with γ-carbonyl oxygen of Asp_{54} from chain C whilst side chain of Lys_{26} of chain C makes bifurcated H-bonding interactions with Gln_{53} and Asp_{54} of chain A respectively. The side chain of another residue Arg_{27} of chain C also makes H-bonds to Ala_{50} and Val_{180} present in its dimeric counterpart. α10 helix of both monomers forming dimers are arranged at an angle of 90° to each other. These centrally located α10 helices (Pro_{171}-Thr_{181}) of hotdog fold from both subunits are buried deep inside this dimeric interface and around 75% of residues of those helices are hydrophobic in nature. Each α10 helix interacts with a short loop ($Phe_{1}82$-Gly_{184}) from its dimeric counterpart to provide overall stability. Being exposed to solvent, both α12 helices in dimer are situated parallelly to each other. N-terminal part of small helix-12 (Tyr_{189}-Pro_{192}) from both subunits interacts hydrophobically with each other. Along with this, the C-terminal end of β1 (His_{17}-Gln_{19}) of both chains confers structural stability to this dimer. It was observed that δ-O carbonyl amide of Gln_{19} of chain A interacts with δ-NH_2 of Gln_{19} from chain C by making an H-bond between them. In another, the γ-carboxylic group of Asp_{18} and carbonyl oxygen of Gln_{19} of counterpart chain A also participates in H-bonding interactions to γ-carbonyl amide of Asn_{23} from its counterpart subunit C and vice versa. Two pi-pi interactions have also been observed between aromatic side chains of both Phe_{175} and Tyr_{189} from both chains.

Comparisons between N-terminal domain of *Pa*lch and other MaoC family hydratases

It has been stated earlier that C-terminal domain was found to be similar with other (*R*)-specific enoyl-CoA hydratases but N-terminal domain showed distinguished characteristic features which were not present in other hydratases. A common characteristic feature of all hotdog fold-containing hydratases exemplified a 9–17 residues long central helix represented as sausage properly wrapped by β-sheet as a bun [33–35,37]. This characteristic feature is distributed through all domains of life until the revelation of a new kind of hotdog fold as observed in *Pa*Ich. It is first noticed in the N-terminal domain of *Pa*Ich that a very short 4-residues helical segment α3 (Trp_{47}-Ala_{50}) which we named as an eaten sausage is not properly placed

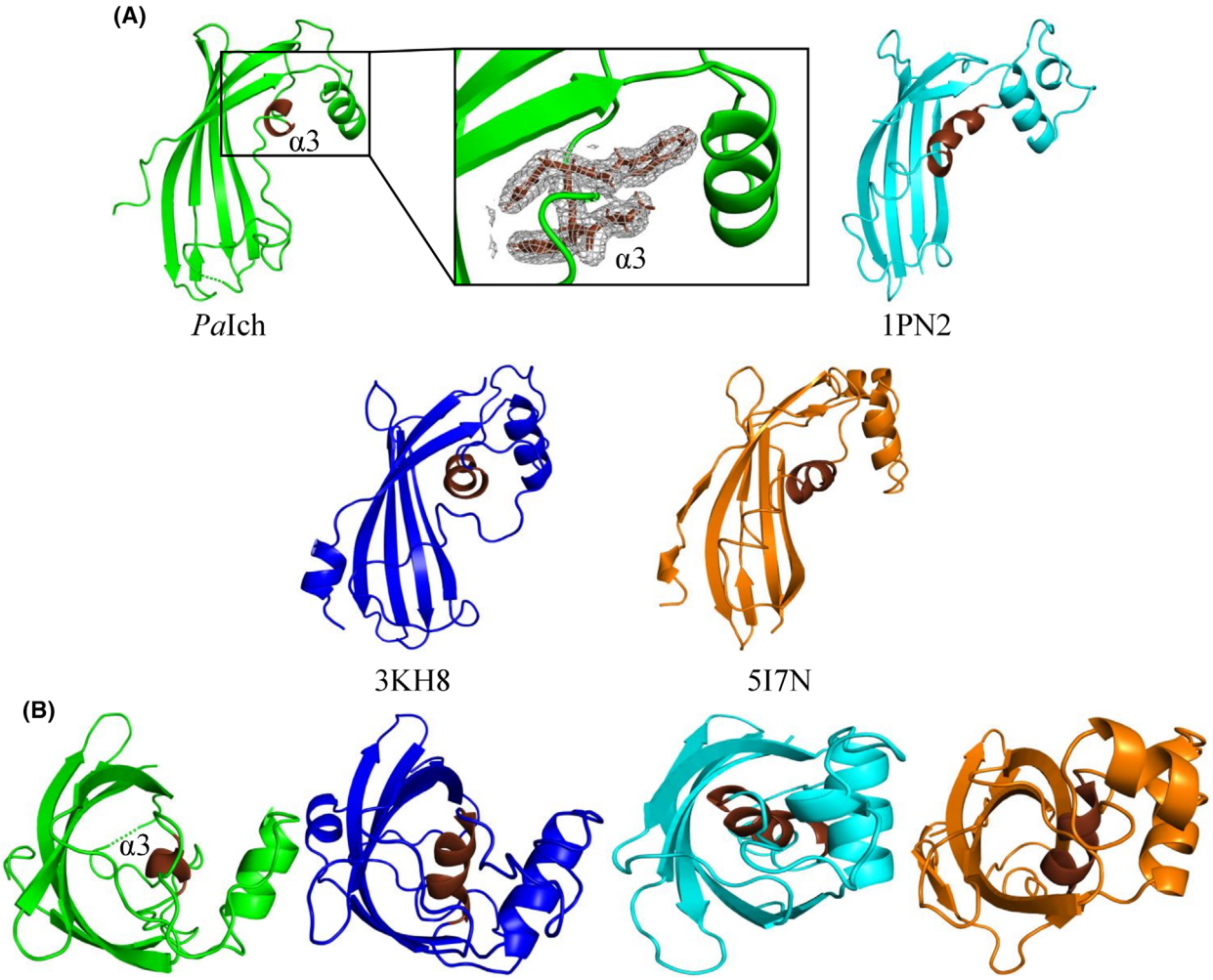


Fig. 3. A unique N-terminal hotdog fold of *Pa*lch and model structure of itaconyl-CoA hydratase from *Yersinia pestis* (*Yp*lch). (A) Cartoon representation of side view of N-terminal hotdog fold. The β-sheets represented here as buns of *Pa*lch from *Pseudomonas aeruginosa*, eukaryotic hydratase 2 from *Candida tropicalis* (PDB ID: 1PN2) and MaoC-like dehydratases from *Phytophthora capsici* (PDB ID: 3KH8) and *Mycobacteroides abscessus* (PDB ID: 5I7N) are shown in green, cyan, blue and orange color respectively whereas all central helices of those β-sheets of so-called "hotdog fold" represented as sausage are shown in chocolate color. A 9–17 residues long central α-helix (chocolate) represents the sausage of hotdog fold as seen in 1PN2, 3KH8, 5I7N and other DHF-containing hydratases where a 4-residues short helical segment α3 (Trp_{47}-Ala_{50}) named as an eaten sausage of *Pa*lch resemble the sausage counterpart shown in other hydratases. Inset: Final $2F_o$-F_c map showing that central helix "α3" (chocolate) of *Pa*lch is in right fit (contoured at 1.0 σ). (B) Top view of N-terminal hotdog fold of *Pa*lch (green) showing that central helix (α3; chocolate) is slipped away from the β-sheet scaffold (bun) while central helices (sausage) of other (*R*)-specific enoyl-CoA hydratases (PDB ID: 1PN2, 3KH8, and 5I7N) are properly wrapped by their respective hotdog folds. All structural representations were made through PYMOL v2.5.4.

inside the core of its hotdog fold (Fig. 3A,B). Top view representation of *Pa*Ich showed that α3 characterized as an eaten sausage has been slipped away from the β-sheet scaffold (Fig. 3B). The extended loop region (~ 7 residues) succeeding helix α3 is the missing region of sausage seen in other hydratases. The reason for appearance of very short helical segment α3 is the presence of α2 and α10 helices from its dimeric counterpart (Figs 2B and 4A). α2 and α10 helices of each monomer interact with the extended loop region (Phe_{51}-Gly_{65}) of the other monomer (Fig. 4A). Mainly α2 rather than α10 helix takes part in this interaction due to α2 being situated in close proximity to extended loop region (Phe_{51}-Gly_{65}). Around 69% of residues of this extended loop region (Phe_{51}-Gly_{65}) are of hydrophobic in nature. So, a strong hydrophobic interaction is seen between α2/α10 and the extended loop region (Phe_{51}-Gly_{65}) between each dimer which ultimately

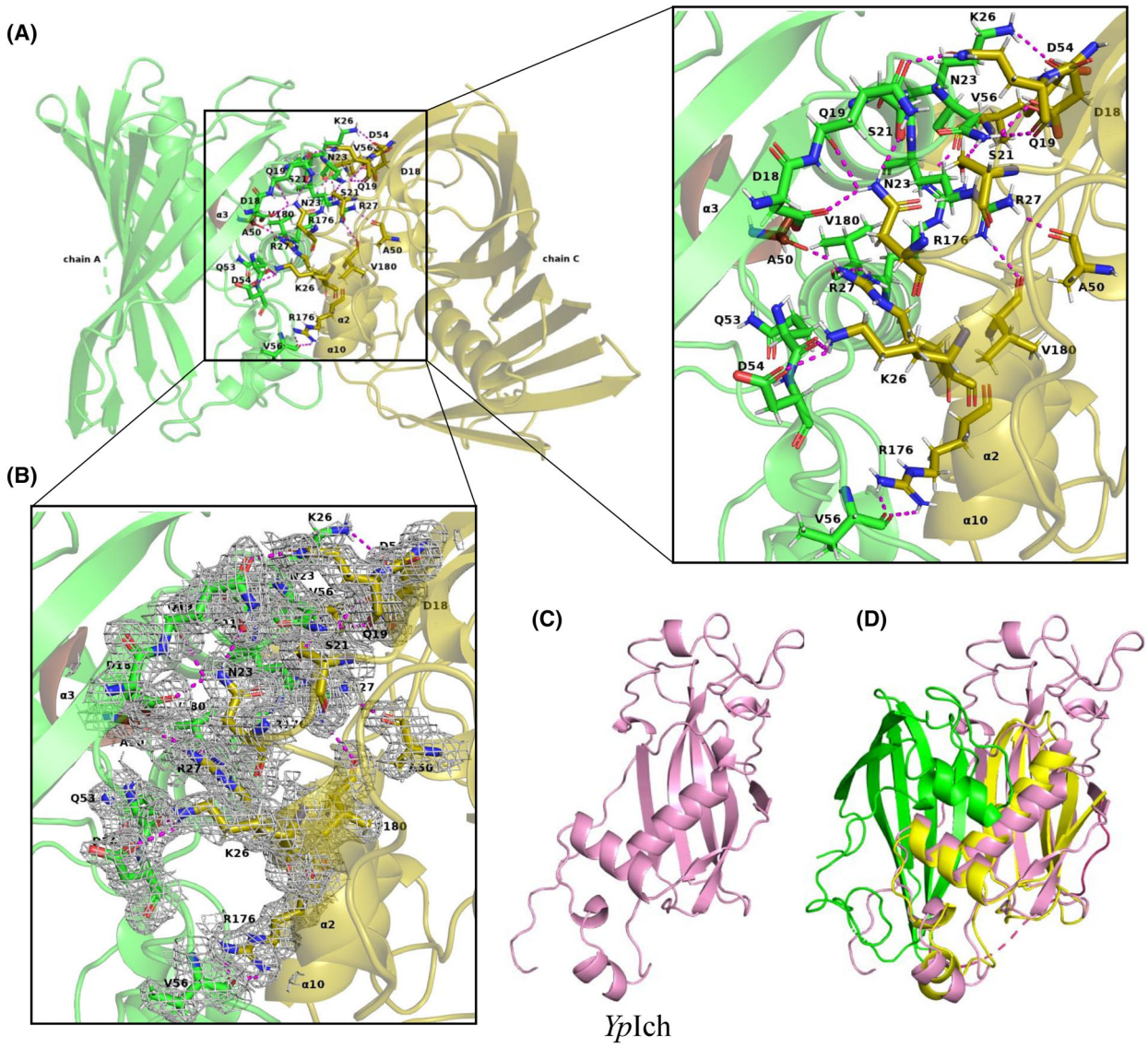


Fig. 4. Dimeric interfaces between two monomers (chainA-chainC). (A) α3 represented as an eaten sausage of monomer A (green) was pulled out from hotdog fold by α2/α10 helices of dimeric counterpart C (olive). Inset: Residues involved in H-bonding interactions between both monomers. Only polar contacts (magenta) are shown here. (B) Inset: Final $2F_o$-F_c map around residues involved in dimeric interfaces showed that they were in right fit (contoured at 1.0 σ). (C) Side view representation of model structure of Ich (pink) from *Yersinia pestis* (*Yp*Ich) consists of a single hotdog fold. (D) Structural superposition between *Yp*Ich and *Pal*ch showed that SHF of *Yp*Ich showed maximum similarity with C-terminal domain (yellow) of *Pal*ch with DHF. All structural representations were made through PYMOL v2.5.4.

gives overall dimeric stability. Conversely, very low polar contacts were made between α3 and the surrounding β-sheet scaffold which ultimately caused relaxed grasping in N-terminal hotdog fold. His_{48} and Trp_{49} are making an H-bonding interaction with the side chain carboxylic group of Glu_{138}. So, due to poor grasping of β-sheet scaffold, α2/α10 helices pulled out the α3 helix (eaten sausage) towards them (Fig. 4A). Another remarkable feature showed the most disordered region (His_{66}-Arg_{79}) of the extended loop (Phe_{51}-Arg_{79}) was positioned far away from the β-sheet main frame which is not similar in other hydratases. Due to high b-factor (~ 82 $Å^2$ over 163 atoms) of this disordered region (His_{66}-Arg_{79}), electron density of ~ 5 residues were not found. However, this region in other hydratases is located in close proximity to the interior wall of the β-sheet scaffold.

Structural comparisons between ich from *P. aeruginosa* (*Pa*lch) and ich (RipB) from *Yersinia pestis*

Here, we have generated a structure of Ich (RipB) of *Y. pestis* using MaoC domain containing dehydratase from *Chloroflexus aurantiacus* as a model (PDB ID: 4E3E) by SWISS-PDB (Fig. 4C). In *Pa*Ich two hotdog folds connected by a long linker region seen in the structure whereas only a single hotdog fold is present in *Yp*Ich homolog. The C-terminal hotdog fold of *Pa*Ich is found to be similar to hotdog fold of *Yp*Ich like in other SHF-containing hydratases (Fig. 4D). In contrast, a five-stranded β-sheet encompassing 13 residues long central α-helix (Ser_{119}-Val_{131}) forms hotdog fold in *Yp*Ich. A 70-residues long-disordered N-terminal stretch followed by the first strand of β-sheet forms the N-terminal region of *Yp*Ich which is not similar in case of *Pa*Ich. Instead, a 13-residues long overhanging segment followed by first strand of β-sheet scaffold made the N-terminal region of *Pa*Ich.

Substrate binding tunnel and active site

Unlike many SHF-containing (*R*)-specific enoyl-CoA hydratases consisting of two active sites located at both top and bottom of the dimeric interfaces, it was observed that only one active site region is present in DHF-containing (*R*)-specific enoyl-CoA hydratases [30,31,33,34,36]. Superposition of *Pa*Ich with its most closely related homolog of *Ct* hydratase 2 (PDB ID: 1PN2) showed a similar conserved hydratase motif present in the structure (Fig. 5A). This conserved hydratase motif is not only present in *Ct* hydratase 2 but it was consistently found in other DHF-containing dehydratases from *Phytophthora capsici*, *Mycobacteroides abscessus* and *Chloroflexus aurantiacus* (PDB ID: 3KH8, 5I7N, and 4E3E) [37]. Based on structural superposition, this hydratase motif is present at the interface of double hotdog fold of *Pa*Ich. This conserved hydratase motif in *Pa*Ich denoted as α11 responsible for hydration of itaconyl-CoA. Due to higher compactness of the core region of C-terminal domain, no active site cleft was seen in that region of all SHF- and DHF-containing hydratases (PDB ID: 3KH8, 5I7N, and 4E3E).

Multiple sequence alignment by clustalW and structure-based sequence alignment by DALI analysis of *Pa*Ich with other DHF-containing (*R*)-specific enoyl-CoA hydratases/dehydratases also identified the conserved active site residues in this hydratase motif responsible for catalysis (Fig. 6) [27]. Active site architecture of (3*R*)-hydroxydecanoyl-CoA (HDC) bound complex of *Ct* hydratase 2 (PDB ID: 1PN4) also confirmed the catalytic residues responsible for this hydration. Asp_{182}, Asn_{184}, and Gln_{187} are three catalytic residues in *Ct* hydratase 2 responsible for hydration of 3(*R*)-HDC [33]. Based on analysis of sequence alignment, Asn_{183}, His_{185}, and His_{188} are three putative catalytic residues in *Pa*Ich responsible for hydration to the C3 position of itaconyl-CoA (Fig. 6). Due to similarity of 3(*R*)-HDC with itaconyl-CoA, a complex of *Pa*Ich with itaconyl-CoA was modeled based on *Ct* hydratase 2 bound 3(*R*)-HDC complex (PDB ID:1PN4; Fig. 5B). Upon addition of a water molecule to itaconyl-CoA by *Pa*Ich, the substrate is converted to citramalyl-CoA which is further hydrolyzed by (*S*)-citramalyl-CoA lyase (Ccl), the last enzyme of the itaconate degradation pathway (Fig. 1A). The position of side chain of putative catalytic residues was also found to be similar with side chain of active site residues in *Ct* hydratase 2. The amide group of Gly_{207} makes an H-bonding interaction with carbonyl oxygen of acyl chain of itaconyl-CoA by creating an oxyanion hole which initiates the process of catalysis. Similar to 3(*R*)-HDC bound *Ct* hydratase 2 complex, The phosphorylated ADP-moiety of itaconyl-CoA bound *Pa*Ich is situated outside of the substrate binding tunnel in a bent conformation whereas the acyl chain is placed inside the central cavity of the double hotdog fold. An H-bonding interaction is observed between the -NH_2 group of adenine moiety and the carbonyl oxygen of Gly_{234} as well as side chain of Lys_{112} makes another polar contact with this ADP-moiety. Another notable difference is also observed between α11 and α13 hydratase motifs of *Pa*Ich and *Ct* hydratase 2 (PDB ID: 1PN2) respectively. It was found that the position of α11 motif of *Pa*Ich is slightly shifted away (~ 2.83 Å) from α13 motif of *Ct* hydratase 2.

Although hydratase motif with active site residues is conserved in all DHF-containing hydratases/dehydratases, the size of active site pocket showed different in *Pa*Ich. It is first noticed that a deep slender substrate binding tunnel greater than its expected length of C_5 acyl chain of itaconyl-CoA seen in *Pa*Ich. However, in *Ct* hydratase 2 (PDB ID: 1PN2) like other DHF-containing hydratases/dehydratases, the range of substrate binding tunnel is limited to the acyl chain length of its corresponding substrate 3(*R*)-HDC. The most remarkable feature found that the tunnel consisting of C_{10} acyl chain of HDC in *Ct* hydratase 2 is smaller than the tunnel consisting of C_5 acyl chain of itaconyl-CoA in *Pa*Ich. No difference in the length of the tunnel was observed between apo form (PDB ID: 1PN2) and 3(*R*)-HDC bound complex of *Ct* hydratase 2 (PDB ID: 1PN4). Similar to other hydratases/dehydratases, the

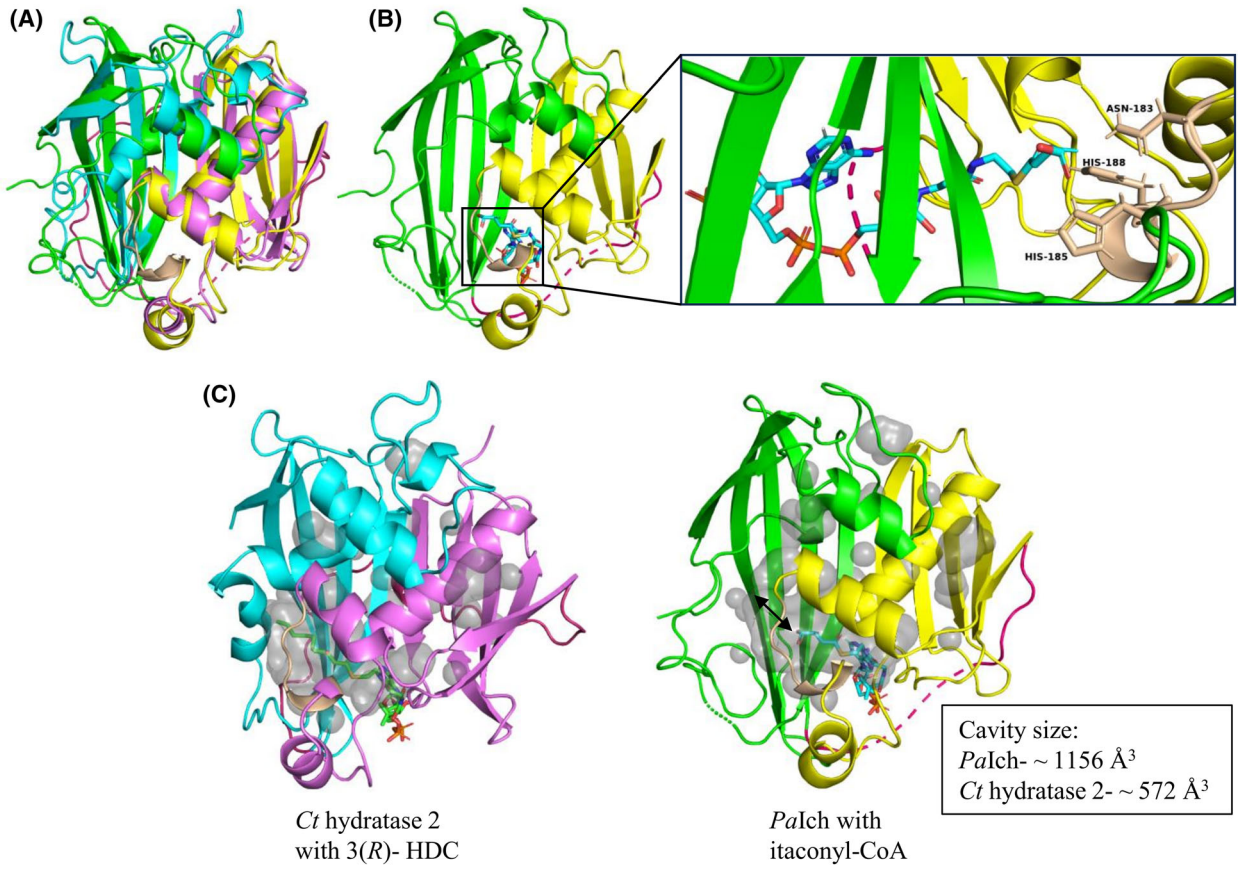


Fig. 5. The enzymatic active site and substrate binding tunnel of *Pa*lch. (A) Superposition of *Pa*lch (N-terminal domain: green and C-terminal domain: yellow) with *Ct* hydratase 2 (N-terminal domain: cyan and C-terminal domain: violet) showed the similar conserved hydratase motif (wheat) responsible for catalysis. (B) A complex of *Pa*lch with itaconyl-CoA was modeled from its most closely related homolog *Ct* hydratase 2 bound 3(*R*)-HDC complex due to structural similarity of 3(*R*)-HDC with itaconyl-CoA. Inset: The three putative catalytic residues are Asn_{183}, His_{185}, and His_{188} (yellow) taking part in hydratase reaction. ADP-moiety of itaconyl-CoA was placed outside of the active site cavity in a bent conformation while the acyl chain is situated into the pocket at dimeric interface. (C) Surface representation of active site pocket (gray) of *Pa*lch consists of an extra region marked as an arrow beyond its C_5 acyl chain of substrate (itaconyl-CoA). It predicts that *Pa*lch can bind longer acyl chain derivatives of CoA while in *Ct* hydratase 2 complex, length of the tunnel (gray) is limited to its substrate's acyl chain length. The volume of active site pockets of *Pa*lch and *Ct* hydratase 2 obtained by PyVol v1.7.8 are ~ 1156 and ~ 572 Å^3 respectively. All structural representations were made through PYMOL v2.5.4.

path of this tunnel was restricted by the central helix (sausage) followed by a long-disordered loop region of N-terminal hotdog fold whereas no such restriction was observed in *Pa*Ich. The presence of central helical segment α3 (Trp_{47}-Ala_{50}) represented as eaten sausage followed by subsequent loop region (Phe_{51}-Gly_{65}) of chain A pulled out by α2 and α10 helices from counter chain C cleared the path of the tunnel in chain A and vice versa. It may access the longer acyl chain derivatives of CoA in the substrate binding tunnel. So, the active site pocket of *Pa*Ich may propose the binding of longer acyl chain derivatives of CoA than C_5 acyl chain of itaconyl-CoA. Despite different tunnel lengths in *Pa*Ich with other hydratases, the entry of the tunnel is similar to other DHF-containing hydratases/dehydratases. Electrostatic potential map of *Pa*Ich and the *Ct* hydratase 2 also showed the positively charged residues lining the active site pocket region (Fig. S2). The mouth of the tunnel of *Pa*Ich is a little wider than *Ct* hydratase 2. His_{66}, His_{185}, Arg_{186}, His_{188} and His_{206} of *Pa*Ich contribute the positively charged surface to the inner lining of the tunnel whereas Arg_{145} contribute to the mouth position. The residues positioned at the entry of tunnel are not so well conserved between them and other (*R*)-specific enoyl-CoA hydratases. Leu_{235}, Arg_{115}, Arg_{145}, and Leu_{119} are situated at the mouth of this tunnel whereas Ile_{233}, Leu_{72}, Phe_{132}, and Phe_{198} are found to be positioned at that region of *Ct* hydratase 2. Likely to other hydratases/dehydratases, β4 and β14 strands of chain A in *Pa*Ich are parallelly situated at the two-

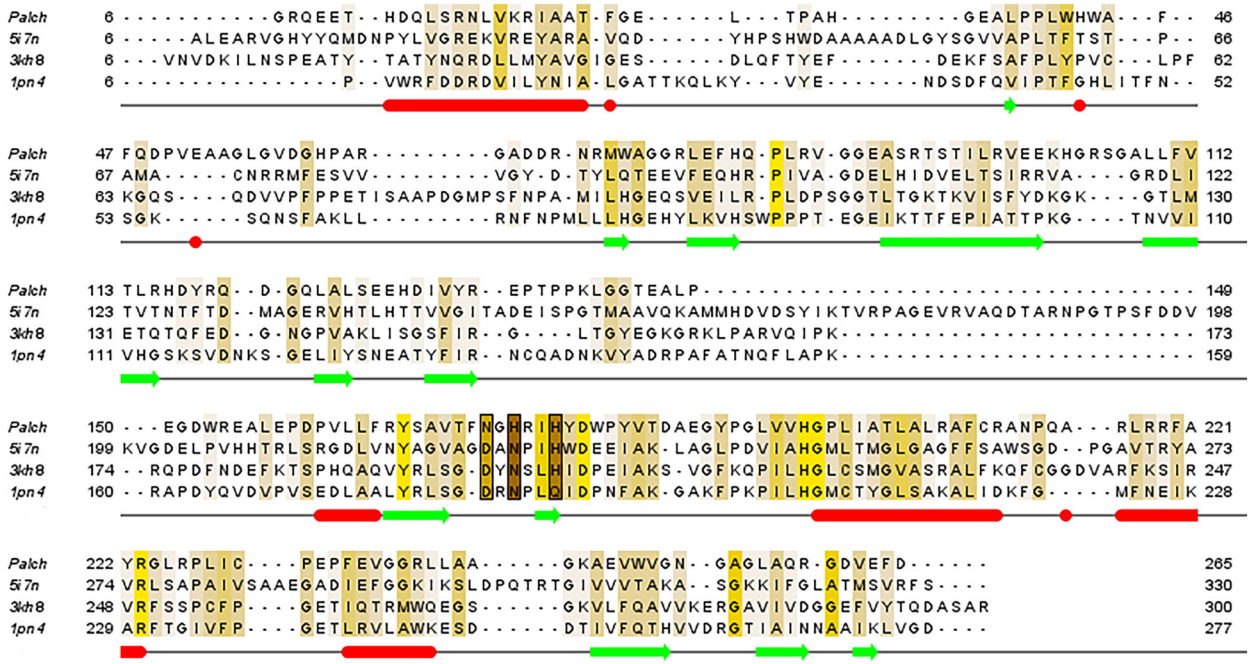


Fig. 6. Structure-based sequence alignment of *Pa*lch with other DHF-containing (*R*)-specific enoyl-CoA hydratases by DALI analysis (PDB ID: 5I7N, 3KH8, 1PN4) and pictorial representation was made through Jalview. Residues of these alignment are colored based on their conservation. Secondary structure prediction analysis showed cylinder (red) and arrow (green) as α-helix and β-sheet respectively. Similar hydratase motif consisting of conserved active site residues marked under square box (black).

domain interface. The C-terminal end of β4 splayed apart from N-terminal end of β14 strand creating the mouth of the tunnel. Another notable difference was observed between the entrance of the tunnel of *Pa*Ich and other DHF-containing hydratases. The N-terminal end of β6 strand of *Ct* hydratase 2 (PDB ID: 1PN2) was more bent towards the core region of N-terminal domain compared to the corresponding β4 strand of *Pa*Ich. At the opening of mouth in *Pa*Ich, Asn_{80}-Trp_{83} of β4 strand produced a slight kink interiorly to make a path for entry of the itaconyl-CoA whereas in *Ct* hydratase 2, Leu_{71}-His_{73} made such a kink. So, a greater arc was required for the entry of longer acyl chain substrate like C_{10}-acyl chain of 3(*R*)-HDC in *Ct* hydratase 2 (PDB ID: 1PN4) rather than shorter arc for C_5-acyl chain of itaconyl-CoA in *Pa*Ich.

Affinity towards acetoacetyl-CoA

In previous studies, it was reported that certain CoA derivatives function as inhibitors for (*R*)-specific enoyl-CoA hydratases [38]. Among these derivatives, acetoacetyl-CoA has demonstrated inhibitory effects on crotonase and mitochondrial enoyl-CoA hydratase (PDB ID: 1DUB) in bovine and rat liver, respectively [19,20]. *Pa*Ich belongs to the MaoC family of (*R*)-specific enoyl-CoA hydratases, and its substrate, itaconyl-CoA, shares structural similarities with acetoacetyl-CoA. Therefore, we attempted to determine the affinity and inhibitory effects through competitive binding assays between itaconyl-CoA and acetoacetyl-CoA. Our findings revealed a significant binding affinity of acetoacetyl-CoA with *Pa*Ich. However, due to the poor stability of itaconyl-CoA, we were unable to perform a competitive binding assay between itaconyl-CoA and acetoacetyl-CoA. Thus, in this study, we exclusively present the binding affinity of acetoacetyl-CoA with *Pa*Ich. The dissociation constant (K_d) value of *Pa*Ich with acetoacetyl-CoA was determined to be 0.214 μM using isothermal titration calorimetry (ITC) experiments (Fig. 7). A control experiment was carried out where *Pa*Ich and buffer (25 mM Tris; pH 8.0 and 150 mM Nacl) were given in cell and syringe respectively.

Discussion

The importance of itaconate degradation pathway in pathogenic bacteria has become increasingly evident over the past few years. It was first identified the genes encoding enzymes responsible for terminating the antimicrobial effect of itaconate by converting it to

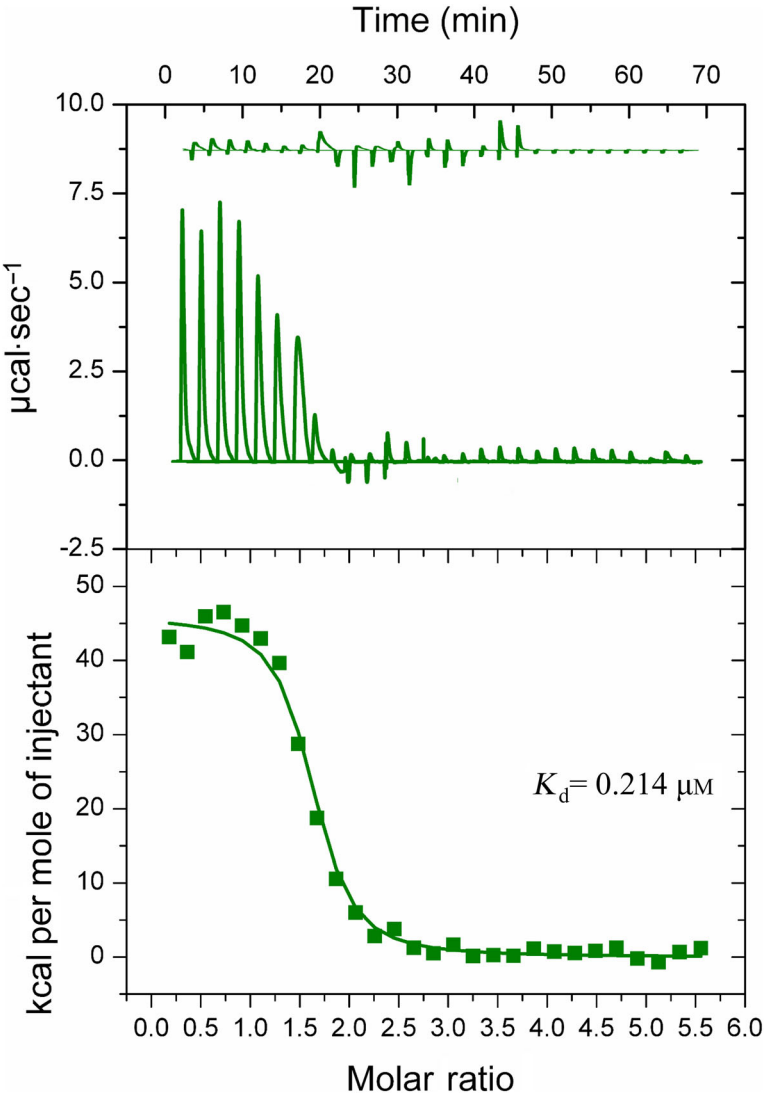


Fig. 7. Isothermal titration calorimetry (ITC) analysis between *Pa*Ich and acetoacetyl-CoA. Acetoacetyl-CoA showed a significant binding affinity of K_d = 0.214 µM. The ITC experiments were carried out at a constant temperature of 25 °C. The binding isotherm for control experiment was shown above the experimental data. The raw data were analyzed through ORIGIN.

acetyl-CoA and pyruvate as nutrient one [14]. Detailed investigation of structural and functional insights of two enzymes, Ict and Ccl of this pathway have been performed so far but the information about the second enzyme Ich remained obscure. Sequence similarity searches through blastp analysis of *Pa*Ich showed maximum matches found between the same genus but different species of *Pseudomonas* whereas very low sequence similarity was observed in other genera. However, in another genus like *Yersinia*, a homolog *Yp*Ich showed ~ 31% sequence identity with *Pa*Ich by blastp analysis. It showed that both enzymes have genetic dissimilarities but carried similar functions. So, it concludes that occurrence of itaconyl-CoA hydratase between different genera likely arose through convergent evolution. In this work, we have determined the first crystal structure of *Pa*Ich which catalyzes conversion of itaconyl-CoA to citramalyl-CoA and the functional insights about active sites and respective substrate binding tunnel. The 1.9 Å crystal structure of *Pa*Ich from *P. aeruginosa* resembles MaoC-like hydratase/dehydratase subfamily proteins consisting of double hotdog folds connected by a long flexible loop. Homology modeling of *Yp*Ich sheds light on its basic structure and showed that it consists of only one domain which is similar to the C-terminal domain of *Pa*Ich. It means structure of *Yersinia* homolog *Yp*Ich share 50% similarities with *Pa*Ich from *Pseudomonas* but both have similar functions. So, the N-terminal domain of *Pa*Ich is supposed to have evolved via gene duplication from single hotdog fold of *Yersinia* or any other hydratases/dehydratases carrying similar functions. Although the number and arrangement of strands in each monomer of a dimeric unit showed

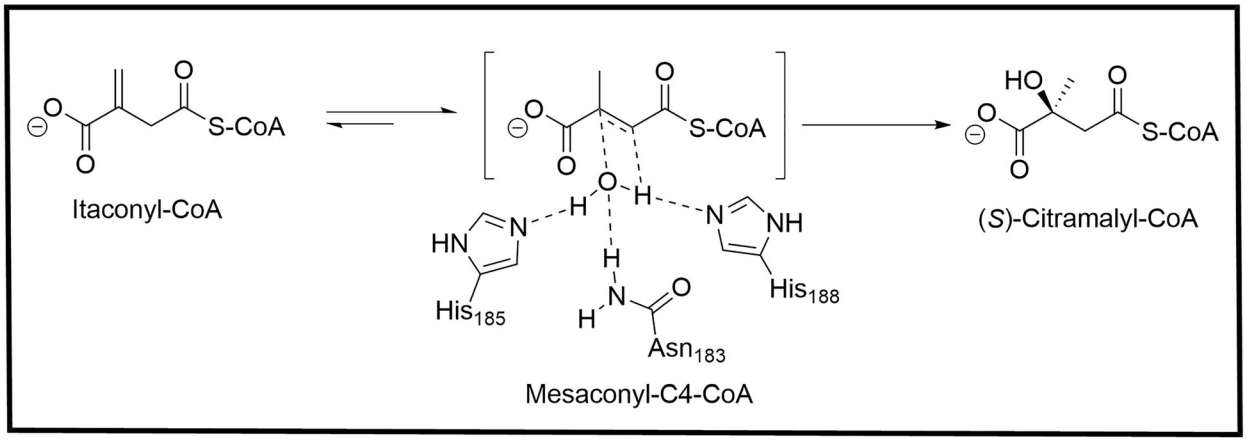


Fig. 8. Proposed catalytic mechanism of *Pa*lch. It converts itaconyl-CoA to citramalyl-CoA through a transient intermediate mesaconyl-C4-CoA. Asn_{183}, His_{185} and His_{188} are three catalytic residues making an H-bonding network with a water molecule which further adds OH^- ion at C3 and a proton at C2 to produce the final product (*S*)-citramalyl-CoA.

differences whereas chain C and chain D from two dimers (A–C and B–D) of an ASU consists of a similar arrangement. It confers pliability is not restricted within helix/loop region but is also present in main frame β-sheet. Herein, we assume that each monomer modulates the main frame architecture of hotdog fold of other monomer while keeping itself fixed in a particular state.

The new N-terminal hotdog fold consisting of a short helical segment α3 pulled out by α2/α10 helices of dimeric counterpart creates a long substrate binding tunnel. So, active participation of two protomer may influence the binding of longer acyl chain coA substrate which is not seen in other hydratases. The sizes of active site pocket of *Ct* hydratase 2 and *Pa*Ich measured by PyVol v1.7.8 analysis are ~ 572 and ~ 1156 $Å^3$ respectively [39]. In contrast to N-terminal hotdog fold of *Pa*Ich, a long 10–17 residues helix represented as full sausage followed by an extended loop region was found in close proximity to the interior wall of N-terminal β-sheet frame in other hydratases/dehydratases (PDB ID: 1PN2, 3KH8, and 5I7N; Fig. 3A). This kind of short helix represented as eaten sausage in *Pa*Ich was most likely arose due to helix–loop transition of 6–10 residues which might be facilitated by pulling effect of α2/α10 helices of counterpart chain (Fig. 4A). Despite slippage of α3 helix from the N-terminal hotdog fold, the main frame β-sheet still holds the similar characteristic features of bun like other hotdog folds. It may be concluded that all MaoC-like hydratases/dehydratases consist of a similar β-sheet scaffold (bun) which is conserved through all domains of life while central helix α3 in *Pa*Ich is found differently as eaten sausage.

It has been found previously that *Pa*Ich of itaconate degradation pathway catalyzes the conversion of itaconyl-CoA to (*S*)-citramalyl-CoA through an unstable intermediate mesaconyl-C4-CoA [14]. It was reported that it carries out two partial activities as itaconyl-CoA-isomerase and mesaconyl-CoA-hydratase where itaconyl-CoA-isomerase first transforms 3-*cis* itaconyl-CoA to 2-*trans* mesaconyl-C4-CoA and later mesaconyl-CoA-hydratase converts this transient intermediate to (*S*)-citramalyl-CoA upon addition of a water [14,40]. At first, a proton is abstracted from C2 of itaconyl-CoA and re-added to C5 by carbonyl amide of Gly_{207} of itaconyl-CoA-isomerase/*Pa*Ich to produce mesaconyl-C4-CoA. Later, Asn_{183}, His_{185} and His_{188} of mesaconyl-CoA-hydratase/*Pa*Ich made an H-bonding network with a catalytic water molecule which further catalyzes the reaction from mesaconyl-C4-CoA to (*S*)-citramalyl-CoA by a nucleophilic attack of OH^- at C3 and addition of a proton to C2 (Fig. 8). So, this study on structural as well as mechanistic insights of *Pa*Ich which connects the missing link of itaconate degradation pathway in virulent organisms like *P. aeruginosa*. So, the structural information about active site pocket with putative catalytic residues will help to find more potent drug candidates against *Pa*Ich in future. More importantly, future drug discovery based on structural insights of *Pa*Ich will ultimately replenish the defensive role of macrophages and prevent the abusive usage of antibiotics.

Acknowledgement

This work was funded by a grant (P07: IP1-SD/503) from the Council of Scientific and Industrial Research,

India. A.P. was financially supported by a UGC scholarship. The authors thank Mr. Jishu Mondal (Technical officer, CSIR-IICB) for assistance in ITC experiments.

Author contributions

AP and SD designed the research; AP performed the research and produced data; and AP and SD analyzed the data; AP and SD wrote and edited the paper.

Data accessibility

The atomic coordinates and structure factors for *Pa*Ich have been deposited in the PDB as entry ID: 8HUC. Additional data supporting the results of this study can be found in the Supporting Information.

References

1 Chen M, Huang X, Zhong C, Li J and Lu X (2016) Identification of an itaconic acid degrading pathway in itaconic acid producing *Aspergillus terreus*. *Appl Microbiol Biotechnol* **100**, 7541–7548.
2 Okabe M, Lies D, Kanamasa S and Park EY (2009) Biotechnological production of itaconic acid and its biosynthesis in *Aspergillus terreus*. *Appl Microbiol Biotechnol* **84**, 597–606.
3 Willke T and Vorlop KD (2001) Biotechnological production of itaconic acid. *Appl Microbiol Biotechnol* **56**, 289–295.
4 Yamamoto K, Nagata K, Ohara H and Aso Y (2015) Challenges in the production of itaconic acid by metabolically engineered *Escherichia coli*. *Bioengineered* **6**, 303–306.
5 Strelko CL, Lu W, Dufort FJ, Seyfried TN, Chiles TC, Rabinowitz JD and Roberts MF (2011) Itaconic acid is a mammalian metabolite induced during macrophage activation. *J Am Chem Soc* **133**, 16386–16389.
6 Michelucci A, Cordes T, Ghelfi J, Pailot A, Reiling N, Goldmann O, Binz T, Wegner A, Tallam A, Rausell A *et al*. (2013) Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production. *Proc Natl Acad Sci USA* **110**, 7820–7825.
7 McFadden BA and Purohit S (1977) Itaconate, an isocitrate lyase-directed inhibitor in pseudomonas indigofera. *J Bacteriol* **131**, 136–144.
8 Li R, Zhang P, Wang Y and Tao K (2020) Itaconate: a metabolite regulates inflammation response and oxidative stress. *Oxidative Med Cell Longev* **2020**, 5404780.
9 Hillier S and Charnetzky WT (1981) Glyoxylate bypass enzymes in yersinia species and multiple forms of isocitrate lyase in *Yersinia pestis*. *J Bacteriol* **145**, 452–458.
10 Martin WR, Frigan F and Bergman EH (1961) Noninductive metabolism of itaconic acid by *Pseudomonas* and *Salmonella* species. *J Bacteriol* **82**, 905–908.
11 Cooper RA, Itiaba K and Kornberg HL (1965) The utilization of aconate and itaconate by *Micrococcus* sp. *Biochem J* **94**, 25–31.
12 Cooper RA and Kornberg HL (1964) The utilization of itaconate by *Pseudomonas* sp. *Biochem J* **91**, 82–91.
13 Kronen M, Sasikaran J and Berg IA (2015) Mesaconase activity of class I fumarase contributes to mesaconate utilization by *Burkholderia xenovorans*. *Appl Environ Microbiol* **81**, 5632–5638.
14 Sasikaran J, Ziemski M, Zadora PK, Fleig A and Berg IA (2014) Bacterial itaconate degradation promotes pathogenicity. *Nat Chem Biol* **10**, 371–377.
15 Zrieq R, Sana TG, Vergin S, Garvis S, Volfson I, Bleves S, Voulhoux R and Hegemann JH (2015) Genome-wide screen of *Pseudomonas aeruginosa* in *Saccharomyces cerevisiae* identifies new virulence factors. *Front Cell Infect Microbiol* **5**, 81.
16 Leesong M, Henderson BS, Gillig JR, Schwab JM and Smith JL (1996) Structure of a dehydratase-isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* **4**, 253–264.
17 Benning MM, Wesenberg G, Liu R, Taylor KL, Dunaway-Mariano D and Holden HM (1998) The three-dimensional structure of 4-hydroxybenzoyl-CoA thioesterase from *Pseudomonas* sp. strain CBS-3. *J Biol Chem* **273**, 33572–33579.
18 Zhuang Z, Song F, Takami H and Dunaway-Mariano D (2004) The BH1999 protein of *Bacillus halodurans* C-125 is gentisyl-coenzyme a thioesterase. *J Bacteriol* **186**, 393–399.
19 Waterson RM and Hill RL (1972) Enoyl coenzyme a hydratase (crotonase). Catalytic properties of crotonase and its possible regulatory role in fatty acid oxidation. *J Biol Chem* **247**, 5258–5265.
20 Engel CK, Mathieu M, Zeelen JP, Hiltunen JK and Wierenga RK (1996) Crystal structure of enoyl-coenzyme a (CoA) hydratase at 2.5 angstroms resolution: a spiral fold defines the CoA-binding pocket. *EMBO J* **15**, 5135–5145.
21 Grosse-Kunstleve RW and Adams PD (2003) Substructure search procedures for macromolecular structures. *Acta Crystallogr D Biol Crystallogr* **59**, 1966–1973.
22 Liebschner D, Afonine PV, Baker ML, Bunkoczi G, Chen VB, Croll TI, Hintze B, Hung LW, Jain S, McCoy AJ *et al*. (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr D Struct Biol* **75**, 861–877.
23 Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV,

Zwart PH and Hung LW (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**, 582–601.
24 Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ and Adams PD (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**, 61–69.
25 Terwilliger TC (2003) Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement. *Acta Crystallogr D Biol Crystallogr* **59**, 1174–1182.
26 Emsley P and Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
27 Holm L, Laiho A, Toronen P and Salgado M (2023) DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci* **32**, e4519.
28 Rigsby RE and Parker AB (2016) Using the PyMOL application to reinforce visual understanding of protein structure. *Biochem Mol Biol Educ* **44**, 433–437.
29 Seifert E (2014) OriginPro 9.1: scientific data analysis and graphing software-software review. *J Chem Inf Model* **54**, 1552.
30 Thoden JB, Zhuang Z, Dunaway-Mariano D and Holden HM (2003) The structure of 4-hydroxybenzoyl-CoA thioesterase from *Arthrobacter* sp. strain SU. *J Biol Chem* **278**, 43709–43716.
31 Li J, Derewenda U, Dauter Z, Smith S and Derewenda ZS (2000) Crystal structure of the *Escherichia coli* thioesterase II, a homolog of the human Nef binding enzyme. *Nat Struct Biol* **7**, 555–559.
32 Dillon SC and Bateman A (2004) The hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics* **5**, 109.
33 Koski MK, Haapalainen AM, Hiltunen JK and Glumoff T (2004) A two-domain structure of one subunit explains unique features of eukaryotic hydratase 2. *J Biol Chem* **279**, 24666–24672.
34 Hisano T, Tsuge T, Fukui T, Iwata T, Miki K and Doi Y (2003) Crystal structure of the (R)-specific enoyl-CoA hydratase from *Aeromonas caviae* involved in polyhydroxyalkanoate biosynthesis. *J Biol Chem* **278**, 617–624.
35 Tsuge T, Sato S, Hiroe A, Ishizuka K, Kanazawa H, Shiro Y and Hisano T (2015) Contribution of the distal pocket residue to the acyl-chain-length specificity of (R)-specific enoyl-coenzyme a hydratases from *Pseudomonas* spp. *Appl Environ Microbiol* **81**, 8076–8083.
36 Johansson P, Castell A, Jones TA and Backbro K (2006) Structure and function of Rv0130, a conserved hypothetical protein from *Mycobacterium tuberculosis*. *Protein Sci* **15**, 2300–2309.
37 Halloum I, Carrere-Kremer S, Blaise M, Viljoen A, Bernut A, Le Moigne V, Vilcheze C, Guerardel Y, Lutfalla G, Herrmann JL *et al.* (2016) Deletion of a dehydratase important for intracellular growth and cording renders rough *Mycobacterium abscessus* avirulent. *Proc Natl Acad Sci USA* **113**, E4228–E4237.
38 Agnihotri G and Liu HW (2003) Enoyl-CoA hydratase. Reaction, mechanism, and inhibition. *Bioorg Med Chem* **11**, 9–20.
39 Smith RH, Dar AC and Schlessinger AJB (2019) PyVOL: a PyMOL plugin for visualization, comparison, and volume calculation of drug-binding sites. *bioRxiv* 816702. 10.1101/816702 [PREPRINT]
40 Hubbard PA, Yu W, Schulz H and Kim JJ (2005) Domain swapping in the low-similarity isomerase/hydratase superfamily: the crystal structure of rat mitochondrial Delta3, Delta2-enoyl-CoA isomerase. *Protein Sci* **14**, 1545–1555.

Supporting information

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

Fig. S1. Mass spectrometry analysis of *Pa*Ich.

Fig. S2. Electrostatic surface potential map of *Pa*Ich and *Ct* hydratase 2.