Structural and functional insights into the initial steps of phenazine biosynthesis

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften (Dr. rer. nat.)
des Fachbereichs Chemie der Technischen Universität Dortmund

Angefertigt am Max-Planck-Institut für Molekulare Physiologie

vorgelegt von

Qi-Ang Li

Dortmund, Juli 2011

Erklärung/Declaration

Die vorliegende Arbeit wurde in der Zeit von März 2007 bis Mai 2011 am Max-Planck-Institut für Molekulare Physiologie in Dortmund unter der Anleitung von Prof. Dr. Wulf Blankenfeldt und Prof. Dr. Roger S. Goody durchgeführt. Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbständig und nur mit den angegebenen Hilfsmitteln angefertigt habe.

The present work was accomplished between March 2007 and May 2011 at Max-Planck-Institute for Molecular Physiology in Dortmund under the guidance of Prof. Dr. Wulf Blankenfeldt and Prof. Dr. Roger S. Goody. I hereby declare that I performed the work presented independently and did not use any other aids but the indicated.

Dortmund, May 2011

Qi-Ang Li

Dedicated to the loving memory of my grandparents.

献给我最怀念的爷爷奶奶

TABLE OF CONTENTS

1 INTRODUCTION

1.1 Natural products: secondary metabolites

Natural products are defined as a large group of organic compounds that are produced by living systems. It can be divided into three major categories: The first category are the primary metabolites, which play critical roles in primary metabolism and are essential for growing, development and reproduction of the producers. The primary metabolites include nucleic acids, amino acids and sugars etc. The second category contain high molecular weight bio-polymers such as lignen and cellulose, which are important for maintaining physical structures of the living cells. The third category are the secondary metabolites, which have attracted great research interests due to their diverse biological activities towards other organisms. Mostly, the term "natural products" is regarded to mean secondary metabolites (Hanson, 2003).

Unlike the primary metabolites, secondary metabolites are naturally synthesized organic compounds that are not directly involved in the growth and development of the producing organisms. Organisms impaired with secondary metabolite synthesis normally do not die immediately, but rather suffer from a long-term damage of their survivability and fecundity. Due to the immense diversity of their structures, functions and biosynthesis routes, it is therefore difficult to appoint natural products into just a few categories. However, in practice it is generally believed that there are five main classes of secondary metabolites: alkaloids, terpenoids and steroids, non-ribosomal polypeptides, fatty acids and polyketides, and enzyme cofactors.

Throughout the development of modern organic chemistry and medicinal chemistry, the natural products have been regarded as the largest pool for novel biological active compounds. For example in the development of cancer treatment, around 50% of the drugs approved since the 1940s are either natural products or their direct derivatives; and a significant number of those drugs/leads are actually produced by microbes (Newman & Cragg, 2007). Therefore, the study of natural products, especially those

with microbial origins, is today one of the most rapidly growing areas and has attracted great research interests.

1.2 Phenazines

1.2.1 Phenazines and phenazine producers

Phenazines are a class of nitrogen-containing heterocyclic compounds that was first discovered and isolated about 150 years ago (Fordos, 1859) and have been extensively studied ever since. To date, more than 6000 phenazine-containing compounds have been reported, and several hundred are known to possess biological activities of which only around 100 are naturally synthesized (Laursen & Nielsen 2004; Mavrodi et al., 2006). Most of these naturally synthesized phenazines are pigmented and due to the modification of the core phenazine ring structure, these compounds show typical colors at a spectra ranging from deep-red to light-blue (Britton, 1983; Price-Whelan et al., 2006) (Figure 1.1).
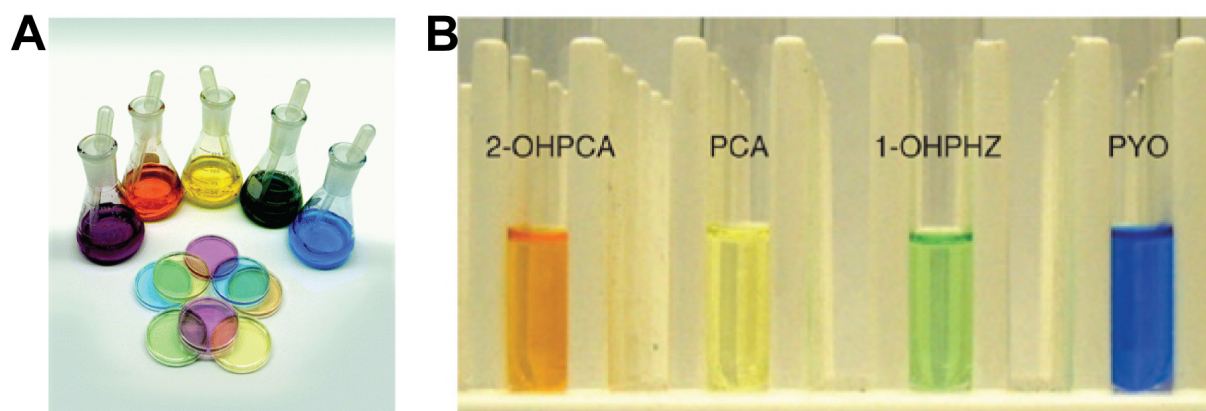
Figure 1.1: (A) Phenazine pigments produced by *Pseudomonas* species come in all the colors of the rainbow (Figure source: American Society for Microbiology). (B) Some of the phenazines produced by *Pseudomonas* strains in aqueous solution. 2-OHPCA: 2-hydroxyl-PCA, PCA: phenazine-1-carboxylic acid, 1-OHPHZ: 1-hydroxyphenazine, PYO: pyocyanine. (Price-Whelan et al., 2006)

Except for the archebacterium *Methanosarcina* (Abken et al., 1998), the natural origin of phenazine are almost exclusively limited to a number of bacterial genera including the Gram-negative fluorescent *Pseudomonas*, as well as *Burkholderia*, *Pantoea* and the Gram-positive *Streptomyces* etc. (Mavrodi et al., 2006). Of all natural

phenazine producers, the best studied are the fluorescent *Pseudomonas* spp., including *P. aeruginosa*, *P. fluorescens*, and *P. chlororaphis* (previously known as *P. aureofaciens*). Each of them is capable of synthesizing two or more phenazine compounds except for *P. fluorescens*, which produces only phenazine-1-carboxylic acid (PCA) (Figure 1.2).

Pyocyanin
(PYO)

Phenazine-1-carboxamide
(PCN/Chlororaphin)

4,9-Dimethoxyphenazine-1,6-dicarboxylic acid dimethylester

1-Hydroxyphenazine

Phenazine-1-carboxylic acid
(PCA)

Phenazine-1,6-dicarboxylic acid
(PDC)

Sendomycin A
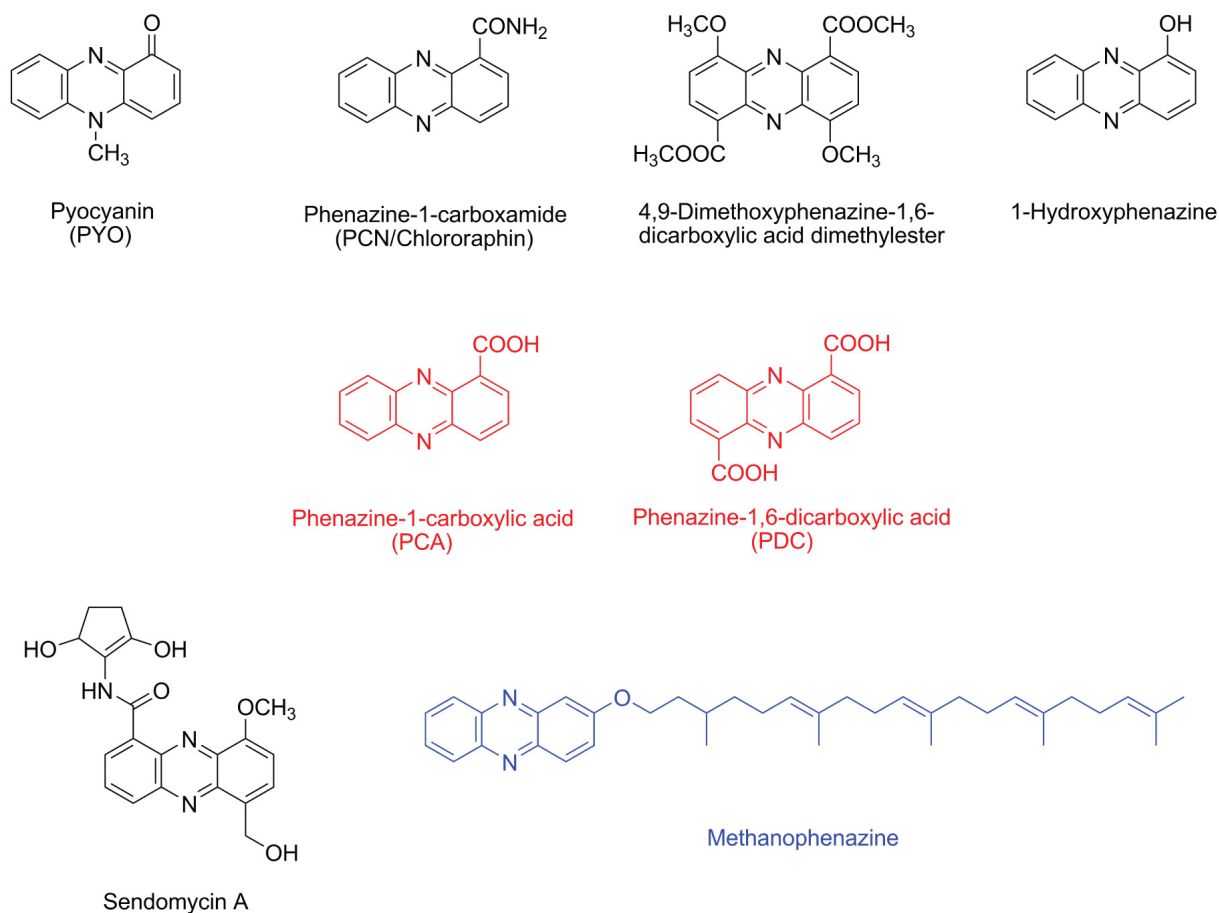
Methanophenazine

Figure 1.2: Some of the naturally produced phenazines. PCA and PDC (red) are precursors of other phenazines produced by bacteria (black). Methanophenazine (blue) is the only known archaeal phenazine that is produced through a different biosynthesis route.

Pseudomonas aeruginosa is a Gram-negative opportunistic pathogen of animals, insects, nematodes and plants, and can also infect immune-compromised individuals, causing both acute and chronic lung disease, that results in a high mortality rate among the patients (Villavicencio, 1998). Pyocyanin (PYO), which is produced by over 90% of *Pseudomonas aeruginosa* isolates, is believed to act as a virulence factor and

3

contribute to the pathogenesity of its producer in the pulmonary tissue damage observed with chronic lung infections of cystic fabrosis patients (Wilson et al., 1988) (Finnan et al., 2004). This is also supported by the observation that PYO deficient *P. aeruginosa* strains are attenuated in both acute and chronic lung infections in a mouse model (Lau et al., 2004). Some strains of another phenazine producer *Burkholderia cepacia*, which produces 4,9-dihydroxyphenazine-1,6-dicarboxylic acid dimethylester (Figure 2.2), is also frequently observed colonizing lung tissues of cystic fibrosis patients, and is reported in many cases as responsible or involved in the mortality of these patients after lung transplantation (Chaparro et al., 2001).

Other phenazine producers including non-human pathogenic pseudomonads such as *P. fluorescens* and *P. chlororaphis* have also been extensively studied. These bacteria are commonly isolated from the rhizosphere (the soil environment surrounding the roots of plants), and are believed to colonize the roots of the plants and contribute to microbial competitiveness. While *P. fluorescens* 2-79 was one of the first strains from which purified phenazine compounds were shown to have anti-fungal activity (Gurusiddaiah et al., 1986), phenazine-1-carboxamide (PCN) produced by *P. aeruginosa* as well as *P. chlororaphis* showed the hightest overall anti-fungal activity in vitro (Smirnov & Kiprianova, 1990). In vivo experiments were later performed which provide direct evidence in correlating phenazine products with the anti-microbial activity of its producers, where Thomashow & Weller showed that phenazine-1-carboxylic acid (PCA) is the major factor that determines the anti-fungal activity of *P.fluorescens* 2-79. They found out that the ability to inhibit pathogens was lost in the phenazine-deficient strains of *P.fluorescens* 2-79, but can be subsequently restored by completing PCA production with the wild-type DNA (Thomashow & Weller, 1988). The importance of phenazines to inhibit phytopathogens has also been documented for *P. chlororaphis* PCL1391, which produces PCA and PCN that contribute to the suppression of foot and root rot of tomato caused by its fungal pathogen *F. oxysporum* (Chin-A-Woeng et al., 2000).

1.2.2 **The redox-activity of phenazines**

Regardless of whether acting as virulence factor contributing to pathogenesis, or as antibiotics in the rhizosphere, studies on the biological function of phenazines have been focusing mainly on their redox properties, which are important characteristics of the phenazine derivatives. For example, PYO in its reduced form can be oxidized by molecular oxygen, resulting in the accumulation of toxic reactive oxygen species (ROS) such as superoxide or hydrogen peroxide (Hassan & Fridovich, 1980), and the oxidized phenazine can further be reduced by various reducing agent including NADH, NADPH or GSH to complete a redox-cycle (Figure 1.3). The antibiotic activity of phenazines is therefore mainly resulting from their ability to generate free radical species, which is potentially harmful to microbes competing with phenazine producers in the environment.
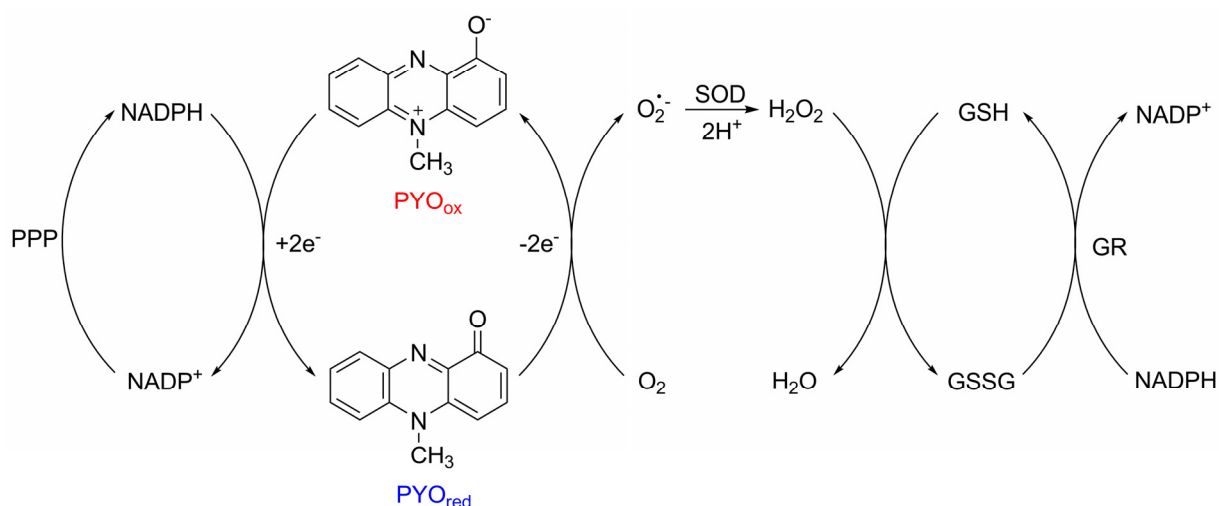
Figure 1.3: The redox-cycling of pyocyanin. PPP: pentose phosphate pathway. PYO$_{ox}$: oxidized form of pyocyanin. PYO$_{red}$: reduced form of pyocyanin. SOD: superoxide dismutase. GSH: glutathione. GSSG: glutathione disulfide. GR: glutathione reductase.

In human alveolar epithelial cells, the action mode of PYO is consisting of several simultaneous processes, including the inhibition of catalase (O'Malley et al. 2004), the reducing of cAMP and ATP (Kanthakumar et al., 1993), and the depletion of the major cellular antioxidant GSH (Muller, 2002). These processes together produce not only ROS, but also pyocyanin free radicals, which further contribute to the virulence effect of PYO. The consequences of the depleting cellular GSH may also cause the activation of redox-sensitive transcription factors that mediate pro-inflammatory processes. During

this PYO stimulated process, neutrophils are attracted into airways, causing neutrophil-mediated tissue damage and inflammation (Denning et al., 1998; Lauredo et al., 1998).

Previous research has also focused on the role of phenazines in assisting iron acquisition of the producing organism. For example, PYO is able to reduce the transferrin-bound Fe^{3+} from the human host, hence making it more available for the infectious *P. aeruginosa* (Cox, 1986). Another example is *P. chlororaphis*, a strain of phenazine-producing pseudomonads isolated from soil, which produces phenazine-1-carboxamide (PCN). *P. chlororaphis* is able to reductively dissolve insoluble iron and manganese oxides which could be further taken up by siderophores, whereas a strain carrying a mutation in one of the phenazine-biosynthetic genes (*phzB*) is not. In this case, PCN is believed to act as an electron shuttle, and the small amount of PCN produced relative to a larger amount of ferric iron reduced indicates that PCN is recycled several times (Hernandez et al., 2004). The role of phenazines as electron shuttle is also supported by the observation that the membrane bound methanophenazine produced by the methanogenic archaea *Methanosarcina mazei* Gö1 acts as an electron carrier that mediates electron transfer between membrance bound targets (Abken et al., 1998; Beifuss et al., 2000).

1.2.3 Physiological roles of phenazines

Interestingly, Lau and coworkers have documented that the growth of *P. aeruginosa* in mouse infection models benefits from the production of phenazines regardless of the presence of other competing microorganisms (Lau et al., 2004). In addition, Bankhead et al. have observed that the composition of the rhizobacterial community did not change after *P. aeruginosa* colonized the root (Bankhead et al., 2004). These evidences suggesting that apart from acting as antibiotics, virulence factors or electron carriers, phenazines may also play important physiological and ecological roles that contribute to the overall biological control of their producers.

Therefore, studies regarding the function of phenazines have been extended to discuss the relevance of phenazine metabolism in the producers themselves. Recently, novel results from Dietrich and coworkers suggest that in *Pseudomonas aeruginosa*, the

phenazine product PYO can function in redox homeostasis to re-oxidize NADH in order to support primary metabolitic pathways such as glycolysis under anaerobic conditions (Price-Whelan et al., 2007). It also has profound activities in controlling the structure and size of colony biofilms, where phenazine-deficient mutants showed over-growing of the colony size. They have also demonstrated that PYO directly activates the iron-containing oxidative stress response regulator SoxR, which subsequently regulates a number of genes involved in transformation/transport of small molecules and the superoxide stress response (Dietrich et al., 2008). This suggests that, independent of introducing oxdiative stress directly, phenazines, which have previously been regarded as secondary metabolites without a direct role in primary cellular processes, play important roles also in the control of gene expression and colony growth of the producers.

In addition, phenazines are also shown to possess abilities in polynucleotide intercalation and topoisomerase inhibition. A number of phenazine molecules produced by pseudomonads have been studied for their properties to bind double-stranded DNA/RNA, and a π–π interaction that leads to intercalation between the planar aromatic phenazine ring and the base pairs was observed (Hollstein & Van Gemert, 1971). Although none of the natural phenazines has been reported to inhibit topoisomerase, synthetic analogues of phenazines have been intensively studied for their capacity to act as topoisomerase inhibitors. For example, synthetic phenazine-1-carboxamide derivatives have been reported to have advantages in multi-receptor targeting, which addresses drug resistance issues in topoisomerase inhibition (Stewart et al., 2001).

1.3 Biosynthesis of phenazines

1.3.1 Precursors of phenazines

While early studies regarding the origin of phenazines have been mainly focused on the common nutrients fed to the producers, no direct evidence has been provided on the identity of the immediate precursors of phenazines (Turner & Messenger, 1986; Mentel et al., 2009). The mass production of penicillin facilitate the research of modern microbiology by means of selective culturing in the 1940s, which led to the discovery of

shikimic acid by selectively growing of *E.coli* mutants as a precursor for many microbial aromatic metabolites (Davis, 1951). By using radioactively labeled substrates, Millican showed that shikimic acid, not anthranilate, was incorporated into pyocyanin; and since shikimic acid is the precursor of anthranilate, it indicated that the biosynthesis of pyocyanine branches off from the shikimate pathway before the step in which anthranilate is formed (Millican, 1962). Similar results have been observed for other phenazine derivatives as well, indicating that the group of naturally occurring phenazines possibly shares a common precursor (Levitch & Stadtman, 1964; Levitch & Rietz, 1966; Podojil & Gerber, 1967; Chang & Blackwood, 1968). Ingledw and Campbell showed quatitative relationship between shikimic acid and pyocyanine in 1969, which further proved the role of shikimic acid as carbon source in phenazine biosynthesis (Ingledew & Campbell, 1969).

Since shikimic acid is a key intermediate in microbes and plants that is directed into a number of metabolite biosynthesis pathways, the exact branch point of phenazine biosynthesis from shikimate pathway remained veiled till chorismate was identified as a common precursor of pyocyanin and other phenazine derivatives (Calhoun et al., 1972; Longley et al., 1972). However, evidence provided by these researches was only indirect because radioactively labeled chorismic acid was not applicable in the experiments due to its poor cell permeability. Later in 1979, Byng and coworkers examined different mutants of *P. aeruginosa* for their capabilities of producing phenazines. They identified three classes of mutants that were pyocyanine deficient and proposed a biochemical scheme implying the precursor-product relationships that cover the terminal steps in pyocyanine biosynthesis, which also supported the argument that chorismic acid acts as phenazine precursor (Byng et al., 1979).

1.3.2 Phenazine biosynthesis genes

Although studies in the 1970s have already identified chorismic acid as the branch point to phenazine biosynthesis, very limited knowledge was provided towards understanding the key steps involved in the formation of the phenazine aromatic rings until genetic analysis of phenazine biosynthesis genes was performed. In the beginning

8

of 1990s, Essar and coworkers reported that the removal of a putative anthranilate synthase in the *P. aeruginosa* genome resulted in a dramatic decrease of pyocyanin production (Essar et al., 1990). Later in that decade, Pierson et al. cloned and sequenced part of the phenazine biosynthesis genes from *P.aureofaciens* 30-84 (Pierson & Thomashow, 1992; Pierson et al., 1995), which for the first time shed light on the previously unknown molecular mechanisms of this pathway. The full set of phenazine biosynthesis genes was then identified by Mavrodi et al. via a complete sequencing of the gene cluster, showing that seven genes *phzABCDEFG* form a defined operon in the genome of *P. fluorescens* (Mavrodi et al., 1998). In 2000, the complete genome sequence of *Pseudomonas aeruginosa* strain PAO1 was published (Stover et al., 2000) and two seven-gene phenazine biosynthesis loci *phzA1B1C1D1E1F1G1* and *phzA2B2C2D2E2F2G2* were cloned in 2001 (Mavrodi et al., 2001). Each of The two gene cluster copies is homologous to the previously sequenced phenazine biosynthesis operon from *P. fluorescens* and is capable of PCA production. The duplication of phenazine genes in *P. aeruginosa* could possibly explain why it is one of the most active phenazine producers.

In the meantime, studies of the *phz* operon have been extended to many other eubacterial phenazine producers (Mavrodi et al., 2001; Giddens et al., 2002; Haagen et al., 2006; Saleh et al., 2009). A collection of phenazine biosynthesis clusters are shown in Figure 1.4 (Mentel et al., 2009). By comparing the phenazine loci among its producers, it has been concluded that five proteins, encoded by five genes *phzB*, *phzD*, *phzE*, *phzF* and *phzG*, are absolutely required for the biosynthesis of phenazines (Mentel et al., 2009). Interestingly, all phenazine-making pseudomonads carry also an additional *phzA* gene which is approximately 80% identical to *phzB*, and a *phzC* gene which encodes a type-II 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) synthase is present in the *phz* operon of many phenazine producers, too. While DAHP synthases from bacteria normally belong to type-I subgroup, which catalyzes the first step in shikimate pathway and are feed-back inhibited by aromatic amino acids, PhzC probably bypasses the allosteric regulation due to the lack of a loop region (Webby et al., 2005). Therefore, it has been suggested that a non-regulated PhzC encoded in the

phz operon which targets at the upstream shikimate pathway is needed to ensure the sufficient production of phenazine precursors when other DAHP synthases are inhibited at the later stage of bacterial growth. Recent report showed that the *phz* operon has been distributed among their bacterial producers by either a conservative mechanism in *Pseudomonas spp.*, or horizontal gene transfer in *Burkholderia spp.* and *Pectobacterium spp.* (Mavrodi et al., 2010; Fitzpatrick, 2009), which is a strong indication that the phenazine biosynthesis in bacteria shares the same core pathway and the large variety of phenazine compounds produced in nature is due to specific modifications of a limited number of precursor molecules. This conclusion is further supported by the identification of genes encoding phenazine-modifying enzymes in the *phz* operon of most phenazine producing bacterial strains. In addition, several other genes that locate closely up- or down-stream to the phenazine biosynthesis core operon were also found encoding phenazine regulatory, resistance or transporter proteins (Figure 1.4).

Some early studies showed that phenazine-1-carboxylic acid (PCA) and phenazine-1,6-dicarboxylic acid (PDC) are highly incorporated into other strain-specific phenazines (Turner & Messenger, 1986). PCA is the only phenazine product of *P. fluorescens* 2-79, the strain with simplest *phz* operon (Figure 1.4), and is not converted from the symmetric phenazine PDC (McDonald et al., 2001). It has been suggested that both PCA and PDC are the core precursor molecules for other natural phenazine derivatives.

The phenazine production of *P. aeruginosa* has been shown controlled by a cell density-dependent, signal transduction mechanism called quorum sensing (QS) at the transcriptional level (Diggle et al., 2008; Girard & Bloemberg, 2008). QS relies on the production of small signaling molecules that can diffuse freely across bacterial cell-wall and activate downstream target proteins when certain concentration has been reached. For example, signaling system in *P. aeruginosa* that depends on 2-heptyl-3-hydroxy-4-quinolone, also named "*Pseudomonas* quinolone signal" (PQS), is known to control the biosynthesis of phenazine pyocyanin and several other virulence factors (Pesci et al., 1999; Calfee et al., 2001).

10
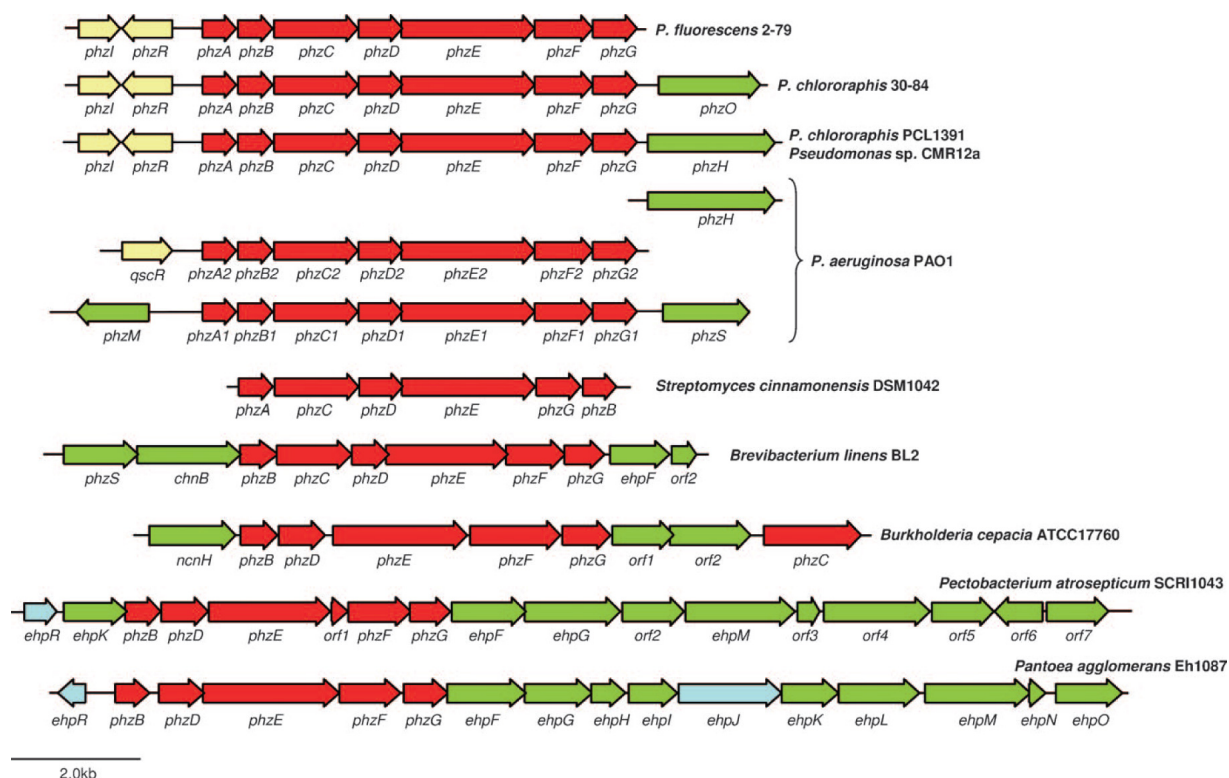
Figure 1.4: Phenazine biosynthesis gene clusters of bacteria. Colors indicate proposed function of the genes, including core phenazine biosynthesis (red), phenazine modification (green), phenazine regulation (yellow) and phenazine transporting/resistance (cyan). Note that although marked in red, PhzC is not distributed in all *phz* operons, and *phzA/phzB* gene duplication is only conserved in pseudomonads.

1.3.3 **Phenazine biosynthesis proteins**

The first systematic investigation of phenazine biosynthesis enzymes was performed in 2001, when McDonald et al. reexamined the point at which phenazine formation branches off from the shikimate pathway and used recombinant *E.coli* expressing all or different subsets of the *phzA–G* genes (McDonald et al., 2001). Their results showed that 2-amino-2-deoxyisochorismic acid (ADIC) is the phenazine precursor, which is converted from chorismate by an anthranilate synthase homologue PhzE. Nevertheless, unlike anthranilate synthases, PhzE lacks the lyase activity and releases the intermediate ADIC into the environment. PhzE catalyzes the first and critical step in phenazine biosynthesis and is responsible for incorporating nitrogen atoms from glutamine into the pathway. However, the direct evidence for ADIC

synthase activity of PhzE is missing at this point since chorismate is an important precursor for aromatic products in bacteria and the degradation of chorismate by other metabolitic enzymes from the cell extracts was almost inevitable.

In the pathway, ADIC is subsequently hydrolyzed to *trans*-2,3-dihydro-3-hydroxyanthranilic acid (DHHA) by PhzD, which is related to the isochorismate synthase family of EntB required for siderophore enterobactin biosynthesis. McDonald et al. proved that PhzD can also utilize isochorismate, 4-amino-4-deoxychorismate and chorismate as substrate (McDonald et al., 2001). The first crystal structure of PhzD from *P. aeruginosa* was determined by Parsons and coworkers in 2003, where they showed that PhzD is remarkably similar to enzymes from a family of α/β-hydrolases. Unlike most of the α/β-hydrolases, the catalytic mechanism of PhzD is distinct. While it lacks a catalytic cysteine that is always found important in other close structural relatives, vinyl ether hydrolysis is catalyzed by an aspartic acid residue (D38) in the active center (Parsons et al., 2003). However, a crystal structure reported recently indicated that a similar active site aspartic acid is conserved in EntB as well (Drake et al., 2006). Although intensive crystallographic studies have been done with respect to PhzD, neither the natural substrate ADIC nor the product DHHA has been observed in a complex crystal structure of this enzyme.

McDonald et al. further demonstrated in their report that, apart from PhzE and PhzD, PhzF and PhzG are also absolutely required for phenazine biosynthesis. However, the mechanism of the dimerization of DHHA to the phenazine ring system remained a myth until year 2004, when Blankenfeldt et al. successfully determined the crystal structure of PhzF in complex with a substrate analogue and proposed a catalytic mechanism for the multiple step condensation of DHHA to PCA. They have found out that DHHA is the substrate of PhzF, an isomerase that catalyzes a pericyclic reaction converting DHHA to 6-amino-5-oxocyclohex-2-ene-1-carboxylic acid (Blankenfeldt et al., 2004). Similar result was also presented by Parsons et al. independently in the same year (Parsons et al., 2004).

The ketone product (Figure 1.5, **1**) of PhzF further undergoes a possibly simultaneous condensation reaction with a second molecule of itself, generating a tricyclic phenazine precursor (Figure 1.5, **2**). Although this reaction does not absolutely require enzyme catalysis, recent studies indicated that the formation of the tricycle is catalyzed by PhzA/B heterodimer (Ahuja et al., 2008), which also explained the earlier observation by McDonald et al. that knock-out of *phzAB* genes decreases but not fully abolishes PCA production (McDonald et al., 2001). The PhzA/B reaction product analyzed by HPLC-coupled NMR spectroscopy indicated that the symmetrical tricyclic product was rearranged to contain four conjugated double bonds (Figure 1.5, **2a**). This molecule undergoes a possibly enzyme-independent oxidative decarboxylation to form an intermediate (Figure 1.5, **3**), that then needs to be oxidized to become fully aromatized. The terminal oxidation steps in phenazine biosynthesis do not require enzyme catalysis since evidence showed that PCA, PDC and also unsubstituted phenazine were produced by a reaction mixture containing only PhzF, PhzA/B and DHHA (Ahuja et al., 2008). Another principle towards the terminal aromatization of the tricyclic phenazine precursor involves PhzG, an FMN-dependent oxidase and the only oxidase in phenazine pathway (Pierson et al., 1995; Parsons et al., 2004). PhzG may be able to oxidize the intermediate **2a** to generate 5,10-dihydrophenazine-1-carboxylic acid, which is the reduced form of the phenazine end product PCA. This is conceivable since it has been shown that one phenazine modifying enzyme dihydrophenazine-1-carboxylate dimethylallyltransferase from *Streptomyces anulatus* possess higher activity towards reduced form of substrates (Saleh et al., 2009). moreover, PCA is further converted to pyocyanin in *P. aeruginosa* by the sequential actions of the putative *S*-adenosylmethionine-dependent *N*-methyltransferease PhzM and the putative flavin-dependent hydroxylase PhzS (Gohain et al., 2006b; Gohain et al., 2006a; Parsons et al., 2007). The existence of PhzG is important also due to the fact that it acts directly on the intermediate **2a** and prevents the oxidative decarboxylation of the substrate, which is suggested as the branch point between PDC and PCA biosynthesis, and also explains why PDC was not converted to PCA by the cell-free extract of *E.coli* expressing full *phz* operon (McDonald et al., 2001).
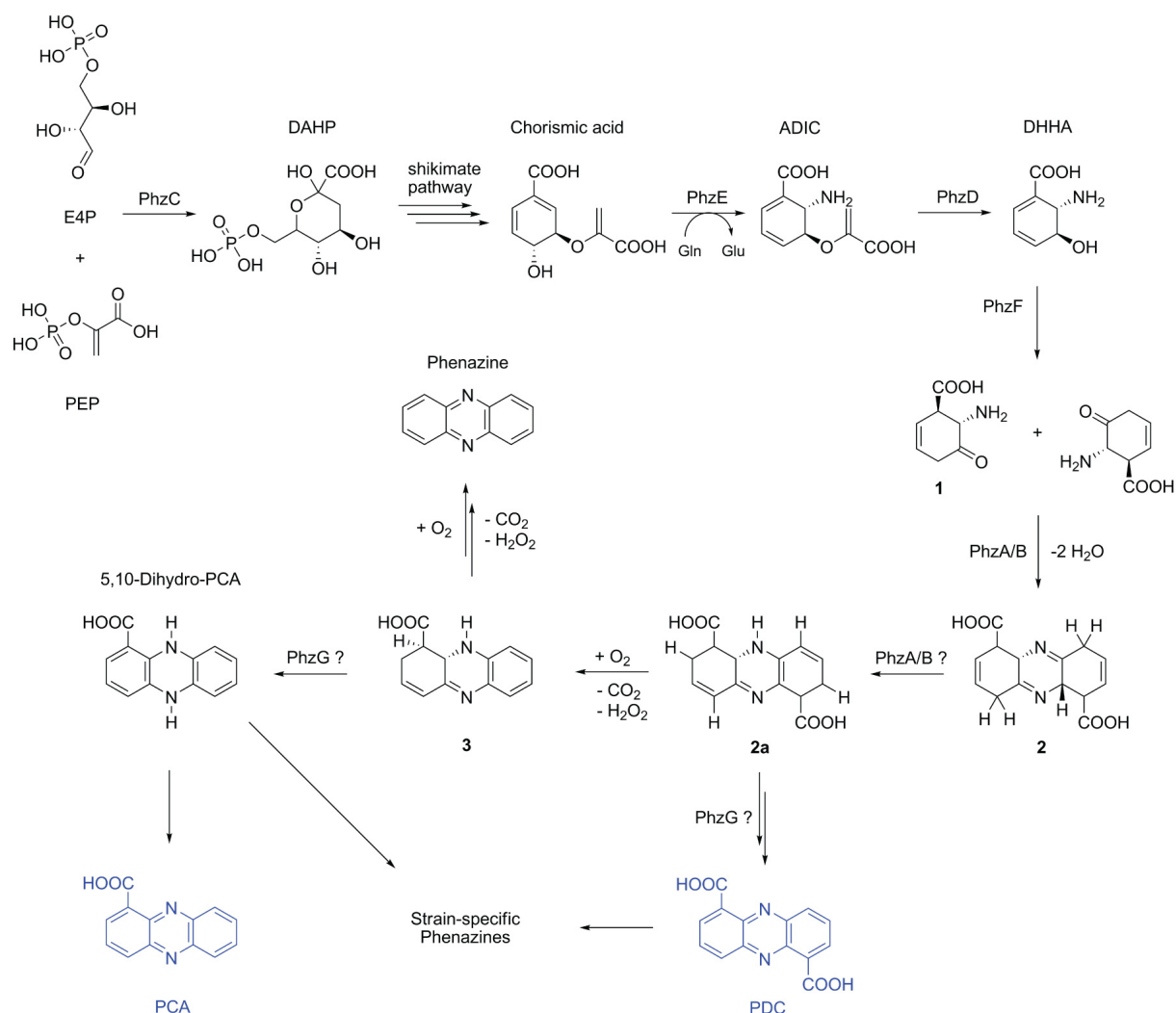
Figure 1.5: Curent understanding of phenazine biosynthesis pathway. Core phenazine products PCA and PDC are shown in blue. PEP: phosphoenol pyruvate, E4P: erythrose-4-phosphate, DAHP: 3-deoxy-D-arabino-heptulosonate-7-phosphate, ADIC: 2-amino-2-deoxyisochorismic acid, DHHA: *trans*-**2,3-dihydro-3-hydroxyanthranilic acid, PCA: phenazine-1-carboxylic acid, PDC: phenazine-1,6-dicarboxylic acid.**

Although the terminal steps of the phenazine biosynthesis pathway have not been fully understood, it is widely believed that PCA and PDC are the core phenazine products of bacteria. Because of the spontaneous decarboxylation of the intermediate **2a**, the PDC-only producers have not been documented. Therefore, the phenazine producing bacteria are currently grouped as PCA-only or PCA/PDC producers (Mavrodi et al., 2006).

14

1.4 **Chorismate utilizing enzymes**

In bacteria, fungi and plants, the shikimate pathway is a metabolic tree that branches carbohydrates into the biosynthesis of a broad range of products. It commits a seven-step synthesis starting with erythrose-4-phosphate (E4P) and phosphoenolpyruvate (PEP) to generate chorismate, which is the precursor of a broad range of primary and secondary metabolites including aromatic amino acids, ubiquinone, folate, vitamin K and the siderophores enterobactin and pyochelin, etc (Herrmann & Weaver, 1999; Dosselaere & Vanderleyden, 2001). The important role of chorismate as a branch point for various metabolic pathways has therefore attracted intensive studies on enzymes that are acting on chorismate, and since the production and utilization of chorismate is exclusively limited to prokaryotic microorganisms and plants, enzymes involved in chorismate metabolism are attractive targets for the development of anti-microbial drugs and herbicides. For example, chorismate mutase (CM) inhibitors have been studied for their roles in the fight against antibiotic-resistant Tuberculosis (TB) (Agrawal et al., 2007) and the more famous glyphosate, which inhibits 5-enolpyruvylshikimate-3-phosphate synthase (EPSPS) in the shikimate pathway, has been extensively used globally as a safe and effective herbicide (Alibhai & Stallings, 2001).

In the review article from Dosselaere & Vanderleyden, five classes of chorismate utilizing enzymes have been discussed regarding their structures and functions. They are chorismate mutase, anthranilate synthase, aminodeoxychorismate synthase, isochorismate synthase, and chorismate pyruvate-lyase (Dosselaere & Vanderleyden 2001b). Of these enzymes, isochorismate synthase and anthranilate synthase catalyze the initial step in menaquinone, siderophore and tryptophan biosynthesis, and are two members of the MST enzyme family (Kolappan et al., 2007). In 2008, the first naturally occurring 2-amino-2-deoxyisochorismate (ADIC) synthase SgcD has been isolated and studied, which confirmed the branching point of chorismate metabolism and added a new member to the MST enzyme family (Van Lanen et al., 2008). Very recently, Andexer and coworkers reported four enzymes comprising two new enzyme classes that are acting on chorismate and have not been reported before. The first class contains chorismate hydrolase FkbO and RapK, and the second class includes 3-

hydroxybenzoate (3HBA) synthase Hyg5 and Bra8 (Andexer et al., 2011). Taken together, there are eight classes of enzymes that have been identified to be chorismate utilizing (Figure 1.6). While focusing on the MST enzyme family, the current understanding of each enzyme class will be discussed separately in details.
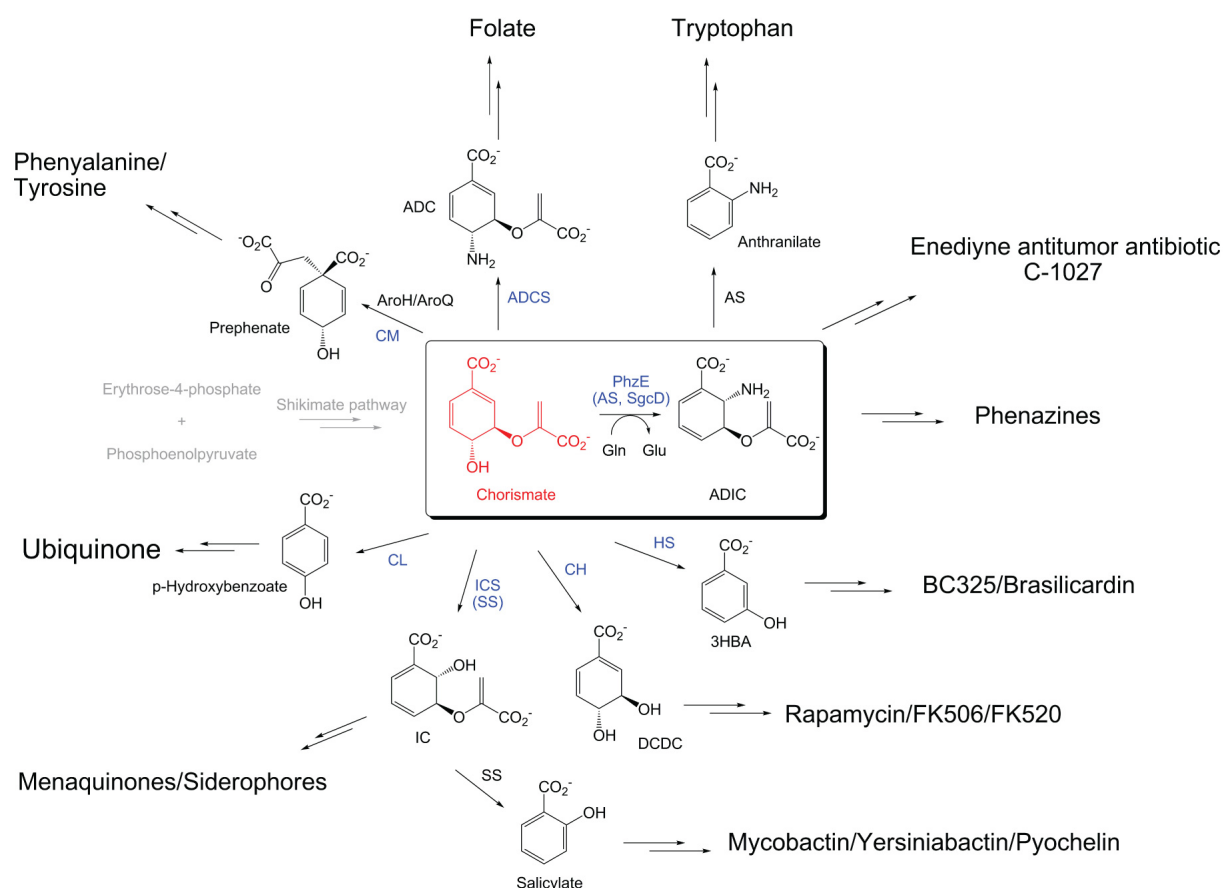
Figure 1.6: Biochemical conversion of chorismate. The PhzE reaction is highlighted with a black box. Enzymes that act directly on chorismate are marked in blue. CM: chorismate mutases of the AroH/AroQ type, CL: chorismate lyase, ICS: isochorismate synthase, SS: salicylate synthase, CH: chorismate hydrolase, HS: 3-hydroxybenzoate synthase, AS: anthranilate synthase, ADCS: 4-amino-4-deoxychorismate synthase, ADIC: 2-amino-2-deoxyisochorismate, IC: isochorismate, DCDC: 4,5-dihydroxycyclohexa-1,5-dienecarboxylic acid, 3HBA: 3-hydroxybenzoate.

1.4.1 **Some chorismate utilizing enzymes**

1.4.1.1 ADC synthase (ADCS)

4-Amino-4-deoxychorismate (ADC) is the precursor of *para*-aminobenzoate (PAB) and folate in microorganisms, and is produced from chorismate by ADC synthase (Figure 1.7). The production of ADC requires a nucleophilic addition of NH_3 to the chorismate ring. Therefore, very similar to anthranilate synthase, ADC synthase is encoded by two genes: *pabA* and *pabB*. While *pabA* encodes a glutamine amidotransferase (GATase), *pabB* encodes a chorismate binding ADC synthase (Dosselaere & Vanderleyden, 2001). Similar to ADIC synthase, ADC synthase also lacks the pyruvate-lyase activity. Therefore, unlike AS which has both a synthase and a lyase activity, the cleavage of pyruvate from ADC to form PAB is performed by an ADC pyruvate-lyase encoded by a separate gene *pabC* in the operon (Ye et al., 1990; Green & Nichols, 1991). The crystal structure of PabB has been determined in 2002 (Figure 1.10). As expected, it is shown that PabB has a complex α/β fold which is similar to the homologue TrpE subunit from AS. Surprisingly, a tryptophan ligand was oberserved in a binding pocket similar to the regulatory site of TrpE deeply buried in the structure of PabB, which cannot be dissociated without denaturing the enzyme (Parsons et al., 2002). It suggested that tryptophan is required for the structural integrity of PabB and may play a role as a positive regulator of folate biosynthesis pathway.
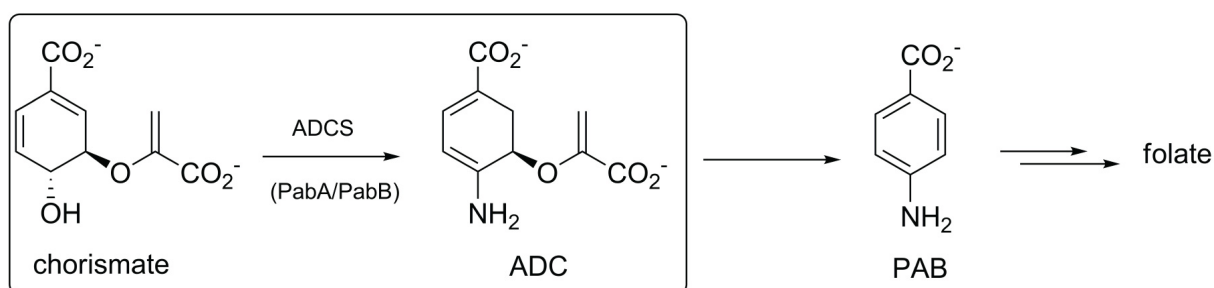
Figure 1.7: 4-Amino-4-deoxychorismate (ADC) synthase (ADCS) catalyzed the first reaction (black box) in folate biosynthesis pathway. PAB: *para*-aminobenzoate.

1.4.1.2 Chorismate mutase (CM)

Chorismate mutase is the enzyme acting on the branch-point reaction of phenylalanine and tyrosine biosynthesis, catalyzing an intramolecular 3,3-sigmatropic rearrangement of the enolpyruvyl moiety of chorismate to produce prephenate (Figure 1.8). Based on the structure, CMs can be divided into two groups: the AroH class and the AroQ class. The AroQ isoenzymes can be further divided into $AroQ_p$, $AroQ_t$, $AroQ_d$, $AroQ_f$, and $AroQ_r$ sub-classes, depending if they are fused to other enzymes or mono-functional and either allosterically regulated or unregulated, respectively. Although overall structures of CMs are fundamentally different, the two classes of CMs share similar catalytic setups at their active sites and possess comparable kinetic parameters (Mattei et al., 1999; Dosselaere & Vanderleyden, 2001). The first crystal structure determined for CM was the unregulated mono-functional CM from *Bacillus subtilis* (Chook et al., 1994, Chook et al., 1993). Interestingly, it was shown that the trimer structure displays a pseudo α/β barrel with the β-sheets from each monomer forming the core and the α-helices wrapping on the outside. The α/β structure of the trimer displays a certain level of similarity to that of the single chained MST (TrpE) domain (subunit) of chorismate utilizing enzymes from the MST family (Figure 1.10).
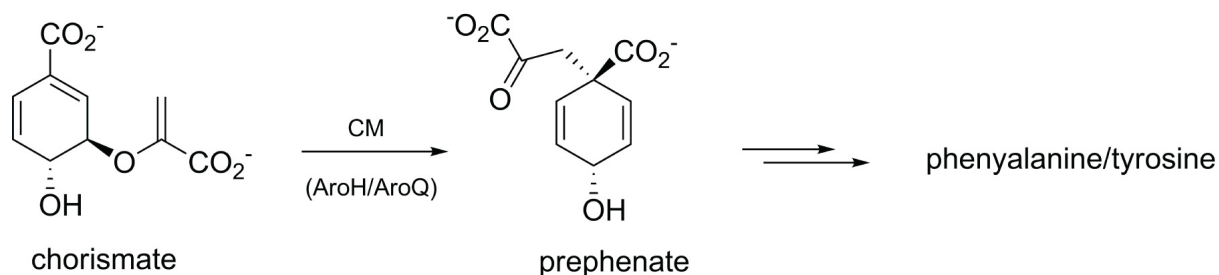
Figure 1.8: Reaction catalyzed by chorismate mutase (CM).

1.4.1.3 Chorismate lyase (CL)

Chorismate lyase catalyzes the conversion of chorismate to *para*-hydroxybenzoate (PHB), which is the first step in the ubiquinone biosynthesis pathway (Figure 1.9). Unlike other chorismate utilizing enzymes from the MST family and ADCS, which share significant sequence/fold similarities, the sequence and structure of CL and CM are

18

largely diversed (Dosselaere & Vanderleyden, 2001). The first crystal structure of CL from *E.coli* was determined in 2001 and a high resolution (1.0 Å) crystal structure of CL in complex with its product PHB was reported in 2006 (Figure 1.10) (Stover et al., 2000; Smith et al., 2006). The ligand-binding properties and mechanism of action have been well characterized and the inhibitor-mutant CL complexes revealed that vanillic acid is an inhibitor of the enzyme. In addition, it should be mentioned that alternative PHB biosynthesis pathways also exist in some bacteria and plants. In contrast to *E.coli*, in which the CL reaction is the only source of PHB (Siebert et al., 1994), *Corynebacterium cyclohexanicum* produces PHB from *para*-oxocyclohexane carboxylate, and in higher plants, PHB is produced from chorismate via the phenylpropanoid pathway in 10 successive enzymatic reaction steps (Kaneda et al., 1993; Loscher & Heide, 1994).
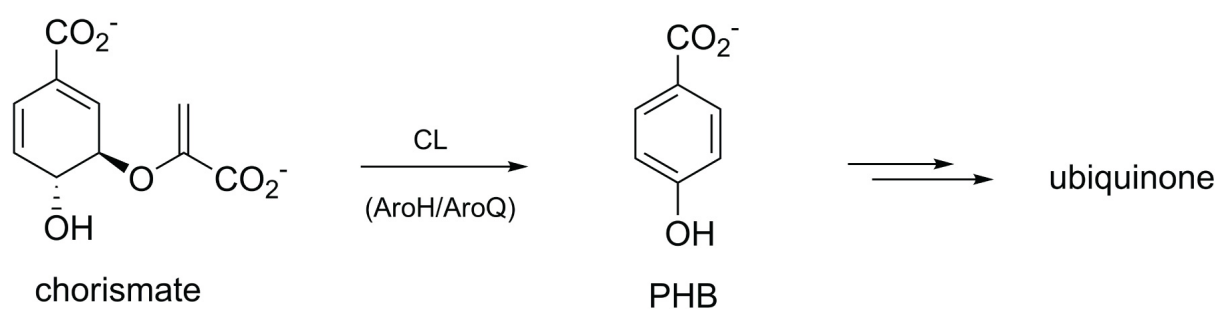
Figure 1.9: Reaction catalyzed by chorismate lyase (CL). PHB: *para*-hydroxybenzoate.

ADCS from *E.coli*   CM from *B. subtilis*   CL from *E. coli*
(PDB entry 1K0E)   (PDB entry 1COM)   (PDB entry 1TT8)
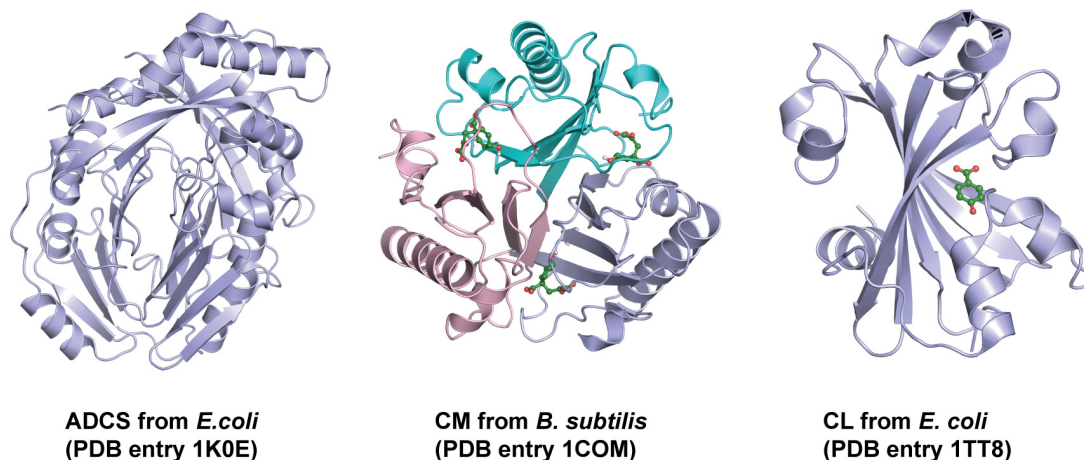
Figure 1.10: Structures of ADC synthase (ADCS) from *E.coli* (chain A), chorismate mutase (CM) from *B. subtilis* (chain A, B and C) and chorismate lyase (CL) from *E. coli*. Cartoon presentations are colored by chain.

1.4.1.4 Chorismate hydrolase (CH) and 3-hydroxybenzoate synthase (HS)

The very recent studies from Andexer et al. have added four enzymes into two new classes of the chorismate utilizing enzymes: the chorismate hydrolase (CH) FkbO and RapK, and 3-hydroxybenzoate (3HBA) synthase (HS) Hyg5 and Bra8 (Figure 1.11). Sequence analysis indicates that these four enzymes are closely related and that they belong to a large group of bacterial proteins with undefined function (Andexer et al., 2011). While FkbO and RapK are encoded respectively in the biosynthesis of macrocyclic polyketides FK506/FK520 and rapamycin, Bra8 was previously assigned an oxidative function and was encoded in the biosynthesis gene cluster for glycosylated diterpene natural product brasiliocardin (Hayashi et al., 2008), and Hyg5 is encoded in an uncharacterized biosynthesis gene cluster (*hyg*) of the rapamycin-producing strain (Ruan et al., 1997). The recognition of FkbO and RapK as CH strongly suggests that Bra8 and Hyg5 act directly on chorismate to generate 3HBA, which has indeed been confirmed by biochemical analysis (Andexer et al., 2011). Further studies are yet to be done regarding structure and function of members from these two enzyme classes.
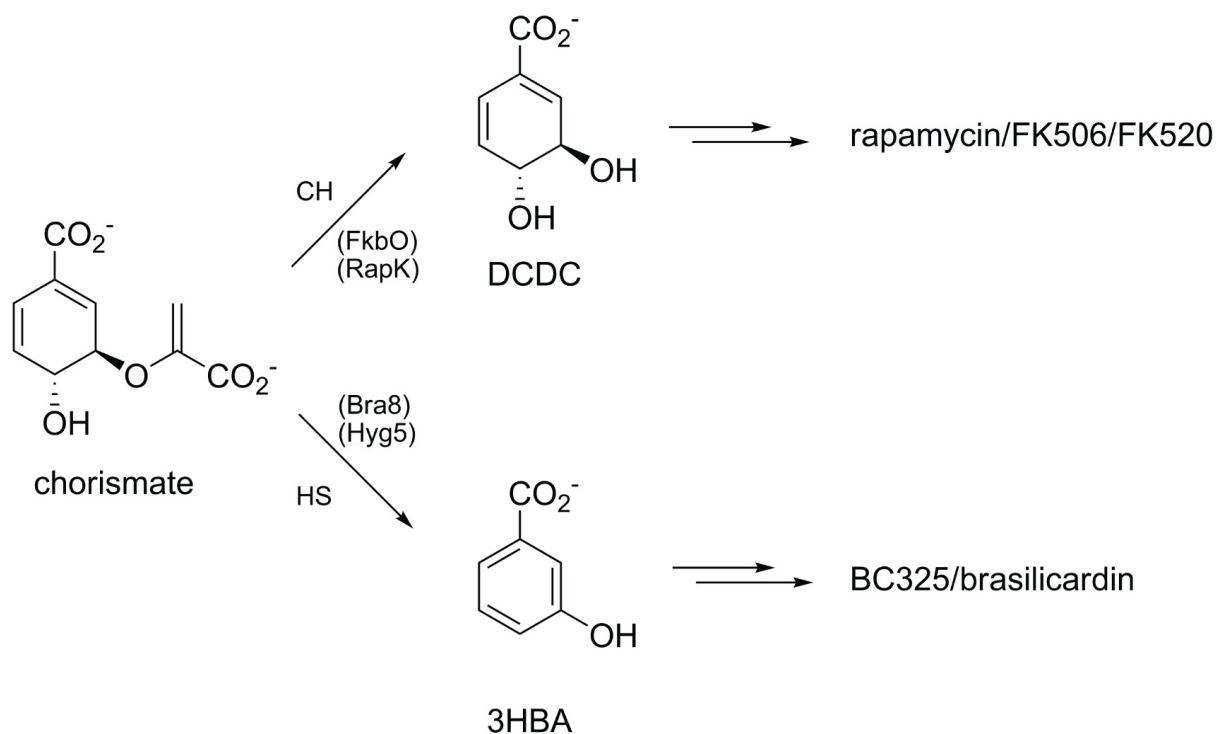
Figure 1.11: Reactions catalyzed by chorismate hydrolase (CH) and 3-hydroxybenzoate synthase (HS). DCDC: 4,5-dihydroxycyclohexa-1,5-dienecarboxylic acid. 3HBA: 3-hydroxybenzoate.

1.4.2 **The MST enzyme family**

The menaquinone, siderophore and tryptophan (MST) biosynthesis enzyme family is currently comprised of isochorismate synthase (ICS), salicylate synthase (SS), anthranilate synthase (AS) and 2-amino-2-deoxyisochorismate (ADIC) synthase (ADICS), which utilize ammonia (AS and ADICS) or water (ICS) to perform nucleophilic substitution at the C2 position of the chorismate ring with or without concomitant rearrangement of the double bond system in an Mg^{2+}-dependent reaction. Of these enzymes, ICS and ADICS release the isomerized product, whereas AS and SS initiate a subsequent sigmatropic rearrangement resulting in the elimination of pyvuvate and generate anthranilate and salicylate, respectively (Spraggon et al., 2001; Kerbarh et al., 2006; DeClue et al., 2005).

1.4.2.1 Isochorismate synthase (ICS) and salicylate synthase (SS)

Isochorismate is one of the precursors required for the biosynthesis of menaquinone/siderophores. The ICS PchA involved in pyochelin siderophore synthesis from *P. aeruginosa* was first reported in 2003 (Gaille et al., 2003). Later, crystal structures of two different ICS from *E. coli* have been reported in separate studies (Figure 1.13 **A** and **B**). Kolappan et al. determined the crystal structure of the first menaquinone-specific ICS MenF and extended the insight into its reaction mechanism, which had been proposed by He et al. in 2004. He and colleagues showed that MenF is a Mg^{2+} dependent chorismate binding enzyme which utilizes Lys190 as base to activate water for nucleophilic attack at the chorismate C2 carbon (He et al., 2004; Kolappan et al., 2007). More recently, the structure of MenF in complex with Mg^{2+} and sulfate bound in the active center was also reported (Parsons et al., 2008). Sridharan et al. published the structure of the enterobactin-specific ICS EntC in complex with Mg^{2+} and the product isochorismate, and performed mutagenesis studies in order to further understand the reaction mechanism of this enzyme (Sridharan et al., 2010).
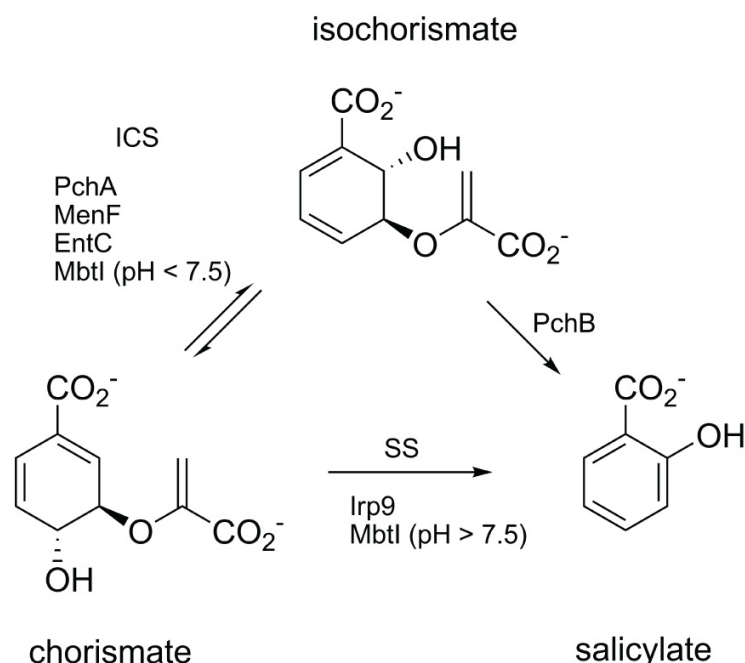
Figure 1.12: Reactions catalyzed by isochorismate synthase (ICS) and salicylate synthase (SS). Note that the function of MbtI is pH-dependent.

In some bacteria, salicylate is produced from chorismate either through a two-step process involving an isochorismate synthase and a pyruvate lyase as observed for *P. aeruginosa* (Gaille et al., 2003), or via a single-step reaction catalyzed by salicylate synthase as with *Y. enterocolitica* (Kerbarh et al., 2006) (Figure 1.12). Salicylate is utilized as one of the building blocks for the biosynthesis of siderophores, a group of low molecular mass iron-chelators, such as pyochelin in *P. aeruginosa*, mycobactin in *M. tuberculosis* and enterobactin in *E. coli*, just to name a few (Crosa & Walsh, 2002). Although isochorismate synthase and salicylate synthase share highly similar structures, SS, unlike ICS, catalyzes a sigmatropic elimination of pyruvate generating salicylate from isochorismate. The crystal structures of bacterial SS Irp9 from *Y. enterocolitic* and MbtI from *M. tuberculosis* have been reported (Figure 1.13 **C** and **D**). Irp9 is Mg^{2+}-dependent and has a complex α/β structure which is conserved in the MST enzyme family. The structure in complex with salicylate and pyruvate was obtained by soaking Irp9 with chorismate, indicating that the protein is still catalytically functional in the crystal (Kerbarh et al., 2006). The crystal structure of MbtI has been reported by two groups of researchers independently. Harrison et al. demonstrated that MbtI is a salicylate synthase which catalyzes the first reaction in the biosynthesis of the siderophore mycobactin (Harrison et al., 2006), while Zwahlen et al. showed that isochorismate is a kinetically competent intermediate in the conversion of salicylate from chorismate catalyzed by MbtI (Zwahlen et al., 2007). Interestingly, they showed that MbtI is a pH- and Mg^{2+}-dependent promiscuous enzyme. In the presence of Mg^{2+} and at pH below 7.5, isochorismate is the dominant product and at pH above 7.5, MbtI converts chorismate almost completely to salicylate (Figure 1.12). In the absence of Mg^{2+}, the protein possesses chorismate mutase activity similar to that of the isochorismate pyruvate lyase PchB from *P. aeruginosa* (Zwahlen et al., 2007).

Despite the high similarities between ICS and SS, differences in residues at the active center of ICS that confers pyruvate lyase acitivity as SS have not been identified (Sridharan et al., 2010). Therefore the exact mechanism that underlines the catalytic diversity of these two enzyme classes still remains a myth.
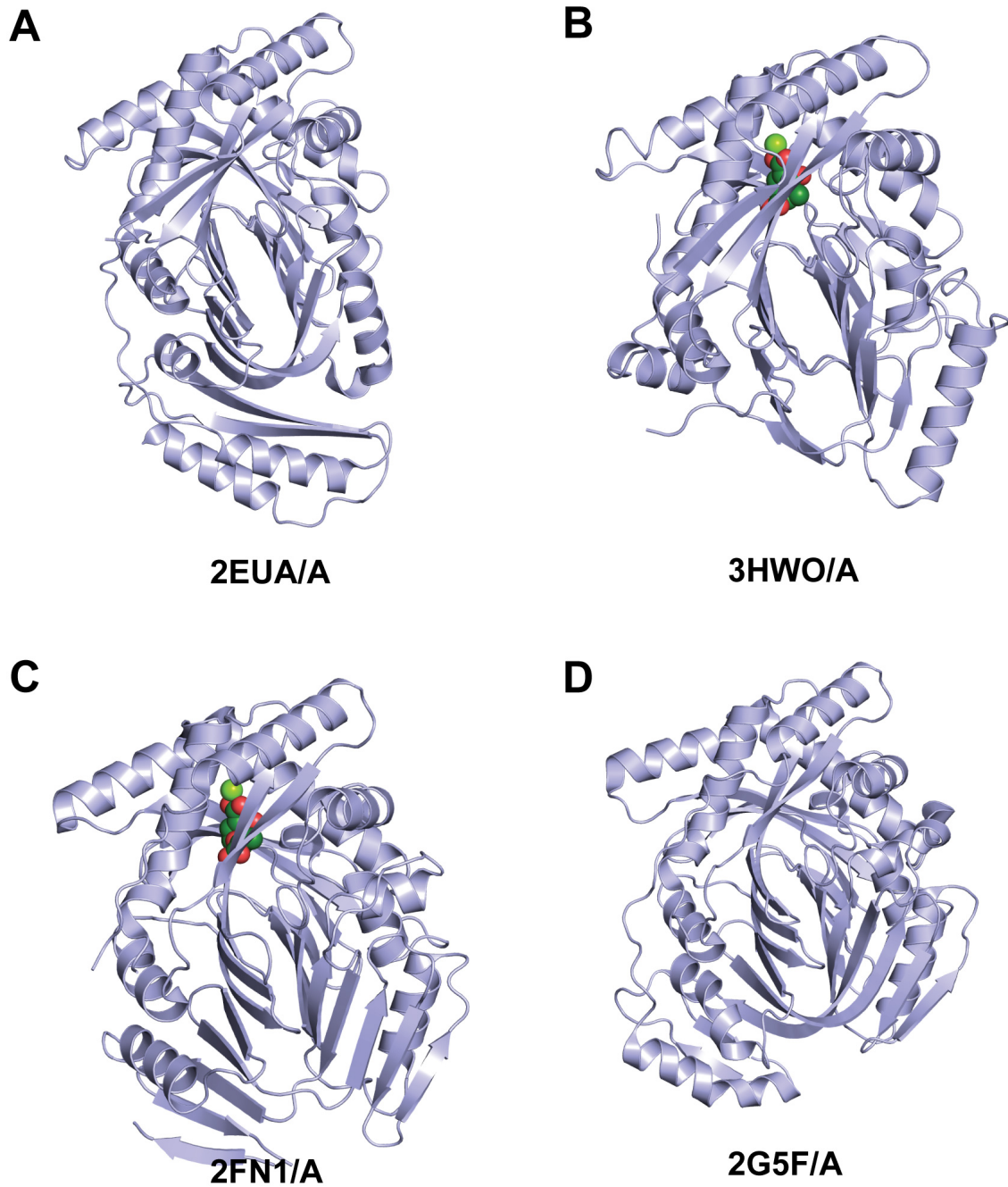
Figure 1.13: Crystal structures of bacterial isochorismate synthase (ICS) and salicylate synthase (SS). (A) ICS MenF from *E. coli*. (B) ICS EntC from *E.coli*. (C) SS Irp9 from *Y. enterocolitica*. (D) SS MbtI from *M. tuberculosis*. Only chain A from each structure is shown and ligands are indicated in stick and sphere presentation and colored in forest when available. Structures are aligned and figure is prepared with PyMOL (Schrödinger LLC).

1.4.2.2 ADIC synthase (ADICS) and anthranilate synthase (AS)

Anthranilate synthase (AS) from bacteria and yeast is a multi-functional enzyme catalyzing the initial reaction in tryptophan biosynthesis. It produces anthranilate through two steps: a reversible amination of chorismate at C2 position to 2-amino-2-deoxyisochorismate (ADIC) and an irreversible sigmatropic elimination of pyruvate from ADIC to anthranilate (Figure 1.14). Mg^{2+} is required for both reactions and ADIC is not released into the solvent during the whole reaction process (Morollo & Bauerle, 1993). The enzyme is composed of two functional polypeptide chains TrpE and TrpG (Zalkin, 1993). In some cases, for example the SvTrpEG from *Streptomyces venezuelae*, AS is translated as a single chained TrpE/TrpG fusion protein (Ashenafi et al., 2008). The TrpE subunit binds chorismate and catalyzes the formation of anthranilate from chorismate and ammonia and belongs to the MST enzyme family. Tryptophan feedback inhibits AS by binding to a distinct site of TrpE subunit, which triggers a conformational change in both subunits and stabilizes the enzyme in its inactive form (Caligiuri & Bauerle, 1991a, Caligiuri & Bauerle, 1991b; Knochel et al., 1999). The TrpG subunit belongs to the family of type 1 "triad" glutamine amidotransferases (GATase1), providing ammonia from *L*-glutamine for the amination at the TrpE active site. It is believed that the AS reaction is strictly ordered, while glutamine hydrolysis by GATase1 only initiates once chorismate has bound to TrpE, and that NH_3 is delievered through an intra-molecular path to the chorismate-binding site to avoid its loss to the solvent (Raushel et al., 2003; Huang et al., 2001).
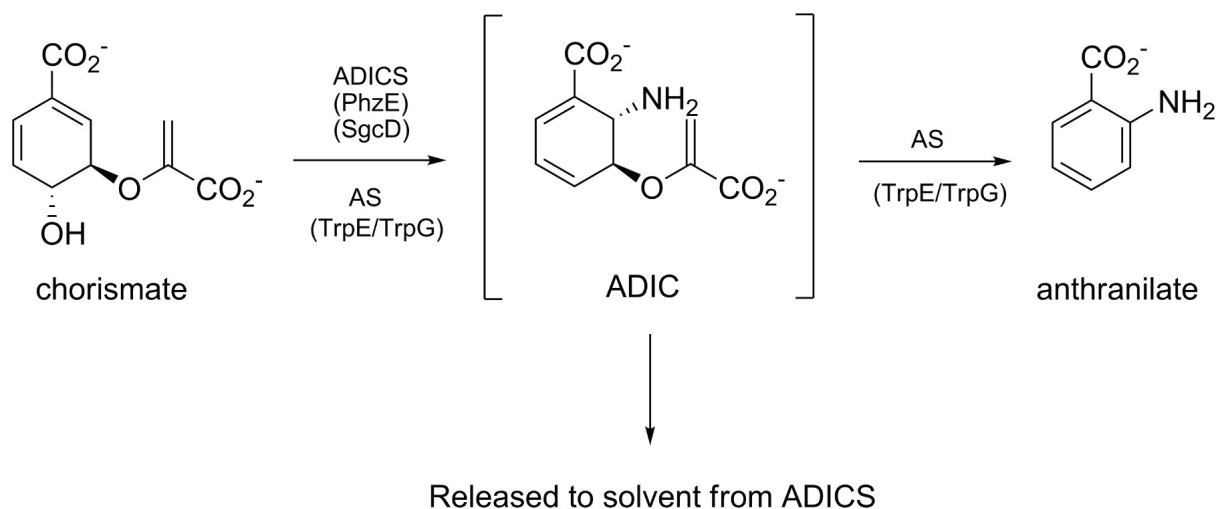
Figure 1.14: Reactions catalyzed by anthranliate synthase (AS) and ADIC synthase (ADICS). Instead of converting ADIC to anthranilate, ADICS releases it to the solvent.

The first crystal structure of anthranilate synthase was determined in its ligand-free form in 1999, when Knochel et al. showed that AS from *Sulfolobus solfataricus* is a TrpE$_2$/TrpG$_2$ heterotetramer, in which two functional TrpE/TrpG protomers associate mainly via the TrpG subunits (Knochel et al., 1999). In 2001, two AS structures in their ligand-bound from different bacterial origins have been reported. The structure of AS from *Salmonella typhimurium* in complex with its allosteric inhibitor *L*-tryptophan showed that binding of tryptophan stabilizes the inactive form of AS by restricting closure of the active site cleft of TrpE (Morollo & Eck, 2001); and the structure of AS from *Serratia marcescens* comfirmed a pyruvate and a putative anthranilate (with ambiguous amine group), as well as a covalently bound glutamyl thioester intermediate in the TrpG active site. It was also revealed that binding of tryptophan to only one of the two allosteric sites is sufficient to quench the catalytic acitivity of both TrpE subunits (Spraggon et al., 2001). Interestingly, although all the three AS structures share the same TrpE$_2$/TrpG$_2$ heterotetrameric setup and display similar TrpE/TrpG functional dimer pairs, they have completely different quaternary structures (Figure 1.15). In additions, neither the substrate chorismate nor the product anthranilate have ever been confirmed in these crystal structures, leaving the detailed reaction mechanism of TrpE still a question to answer.

Compared to the extensively studied anthranilate synthase, the current knowledge about ADIC synthase is rather limited. Unlike anthranilate synthase, which catalyzes a two-step converstion of chorismate to anthranilate, ADIC synthase releases the intermediate ADIC into the solvent without concomitant cleavage of the pyruvate group (Figure 1.14). As discussed above, previous studies on ICS, SS and AS have provided insights into the nucleophilic addition step of the reaction, however, the understanding of residues involved in the pyruvate elimination step is still lacking. Although Morollo et al. observed a transient accumulation of ADIC in the solvent by incubation chorismate with engineered AS mutant H398M from *Salmonella typhimirium*, the activity of the mutant was in fact very low (less than 1% of wild-type activity) and the efficiency of ADIC conversion was less than 15% of the substrate chorismate; and with an prolonged incubation, ADIC was completely converted to anthranilate in the same reaction mixture (Morollo & Bauerle, 1993). Later on, the first native ADIC synthase activity was proposed by genetic analysis for the enzyme PhzE, which catalyzes the initial step of phenazine biosynthesis (McDonald et al., 2001). However, biochemical and structural insight into this enzyme is still missing and contradictory evidence showing PhzE's ability to complement an AS mutant strain of *E.coli*, which in turns recover tryptophan biosynthesis, was also reported, indicating PhzE is an AS instead of ADIC synthase (Pierson et al., 1995). In 2008, Van Lanen et al. reported the first biochemical characterization of an ADIC synthase SgcD, which catalyzed the first reaction in the biosynthesis of C-1027, an enediyne antitumor antibiotic. They also questioned the role of PhzE as an ADIC synthase while PhzE and SgcD have onle less than 12% sequence identity (Van Lanen et al., 2008).
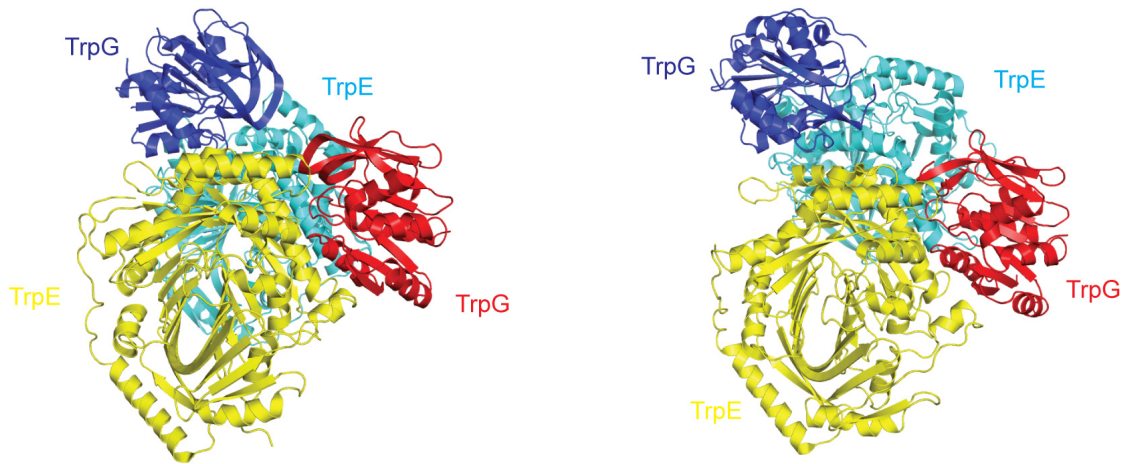
**AS from S. marcescens
(PDB entry 1I7Q)**

**AS from S. typhimurium
(PDB entry 1I1Q)**
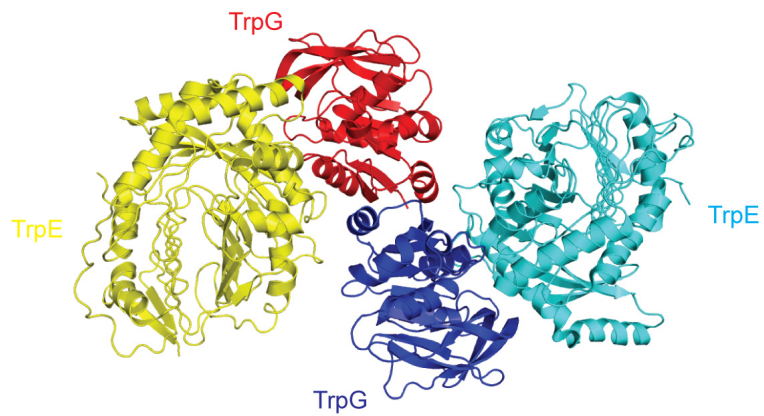
**AS from S. solfataricus
(PDB entry 1QDL)**

Figure 1.15: Cartoon presentation of quaternary structures of anthranilate synthase from different bacteria, colored by chains.

2 AIMS OF THIS STUDY

Phenazines are nitrogen-containing heterocyclic pigments produced by a number of bacterial genera, including fluorescent *Pseudomonas*, *Burkholderia*, *Brevibacterium* and *Streptomyces*. It is believed that phenazines are used as redox-active antibiotics in microbial competitiveness and may also have diverse physiological functions because they also act as signalling molecules and as respiratory pigments under anoxic conditions e.g. in the deeper anoxic layers of biofilm.

PhzE catalyzes the first reaction in the phenazine biosynthesis pathway, producing 2-amino-2-desoxyisochorimate (ADIC) from chorismate. The enzyme belongs to the menaquinone, siderophore, tryptophan biosynthesis (MST) family, using ammonia to substitute chorismic acid at C2 position without subsequent elimination of pyruvate. Since chorismate utilization is limited only bacteria and plants, PhzE may be an attractive target for pharmaceutical intervention.

Crystallographic and biochemical studies have lead to considerable structural and mechanistic insight into members of the MST family in recent years, yet structures of a MST/GATase1 fusion protein and of an ADIC synthase are lacking. In addition, the existence of an ammonia transporting channel has never been demonstrated in these enzymes. This study aimed at filling these gaps by determining the crystal structure of ADIC synthase PhzE from *Burkholderia lata* 383, an enzyme that consists of an MST/GATase1-fusion in a single chain. In addition, experiments were designed to confirm the functional assignment of PhzE as an ADIC synthase using a broad spectrum of biochemical, biophysical and analytical methods. It is also of great interest to investigate if the enzyme is subject to feedback inhibition similar to some of the related anthranilate synthases and the potential regulatory mechanism.

In order to assist the functional analysis of PhzE, crystallographic studies have been extended to PhzD from *Pseudomonas fluorescens* 2-79, which catalyzes the step following PhzE in phenazine biosynthesis.

3 MATERIALS AND METHODS

3.1 Materials

3.1.1 Chemicals

Chemicals used in this study were purchased from the following companies: Applichem (Darmstadt, DE), Boehringer (Mannheim, DE), Fluka (Neu-Ulm, DE), Gerbu (Gaiberg, DE), JT Baker (Deventer, NL), Merck (Darmstadt, DE), Roth (Karlsruhe, DE), Serva (Heidelberg, DE) and Sigma-Alderich (Deisenhofen, DE).

3.1.2 Kits, Markers and Enzymes

Kits	Supplier
QIAprep Spin Miniprep Kit	Qiagen (Hilden, DE)
QIAquick PCR Purification Kit	Qiagen (Hilden, DE)
QIAquick Gel Extraction Kit	Qiagen (Hilden, DE)
JETquick Plasmid Miniprep spin Kit	Genomed (St. Louis, USA)
BigDye Terminator Cycle Sequencing Kit	Fermentas (Langen, DE)
Markers	**Supplier**
GeneRuler 1kb DNA ladder	Fermentas (Langen, DE)
Unstained Protein Molecular Weight Marker	Fermentas (Langen, DE)
Gel Filtration Standard	Bio-Rad (München, DE)
Enzymes	Supplier (Source)
FastDigest Restriction Enzymes	Fermentas (St. Leon-Rot, DE)
Phusion DNA Polymerase	Finnzymes (Espoo, FI)
T4 DNA Ligase	Fermentas (St. Leon-Rot, DE)
Tobacco Etch Virus Protease	Lab prepared

3.1.3 Microorganisms

Strains	Genotype	Supplier (Source)
E. coli KA12	F2 l2 D(*pheA-tyrA*) *thi-1 endA1 hsdR17* D(*argF-*	Collaborator

(chorismate accumulating strain)	*lac*)*205*(*U169*) *supE44* D(*srlR-recA*)*306*::Tn*10*	
E. coli XL1-blue	*endA1 gyrA96*(nal[R]) *thi-1 recA1 relA1 lac glnV44* F'[::Tn10 *proAB[+] lacl[q]* Δ(lacZ)M15 Amy Cm[R]] *hsdR17(r_K[-] m_K[+])*	Stratagene (Santa Clara, USA)
E. coli Rosetta pLysS	F[-] *ompT hsdS*$_B$(r$_B^-$ m$_B^-$) *gal dcm* (DE3) pLysSRARE (Cam[R])	Novagen (Darmstadt, DE)
E. coli Rosetta2 pLysS	F[-] *ompT hsdS*$_B$(r$_B^-$ m$_B^-$) *gal dcm* (DE3) pLysSRARE2 (Cam[R])	Novagen (Darmstadt, DE)

3.1.4 **Culture Media and Antibiotics**

Medium	Composition
Luria-Bertani (LB)	10 g/L Bactotryptone, 10 g/L NaCl, 5 mM NaOH, 5 g/L yeast extract
Terrific Broth (TB)	12 g/L BactoTryptone, 24 g/L Bacto-yeast-extract, 4 g/L glycerol, 17 mM KH$_2$PO$_4$, 72 mM K$_2$HPO$_4$
GYT medium	10% glycerol, 0.125% (w/v) yeast extract, 0.25% (w/v) tryptone
TSS medium	LB medium containing 10% (w/v) PEG 8,000, 30 mM MgCl$_2$, 5% DMSO
M9	6 g/L Na$_2$HPO$_4$, 3 g/L KH$_2$PO$_4$, 0.5 g/L NaCl, 1 g/L NH$_4$Cl, 1 mg/L Thiamine, 1 mM MgSO$_4$, 0.1 mM CaCl$_2$, 0.2% (w/v) glucose

Antibiotics	Concentration applied	Supplier
Ampicillin	100 mg/L	Gerbu (Gaiberg, DE)
Chloramphinicol	34 mg/L	Gerbu (Gaiberg, DE)
Tetracycline	50 mg/L	Gerbu (Gaiberg, DE)

3.1.5 **Buffers and Solutions**

Buffers and Solutions	Composition
Protein Purification	
Ni-NTA Buffer A	50 mM Na_2HPO_4, pH 8.0, 500 mM NaCl, 5 mM 2-Mercaptoethanol.
Ni-NTA Buffer B	Ni-NTA Buffer A plus 500 mM Imidazol, pH 8.0.
Dialysis Buffer	50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM 2-Mercaptoethanol.
Gel Filtration Buffer	20 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM 2-Mercaptoethanol.
SDS-PAGE	
4× Stacking Gel Buffer	0.5 M Tris-HCl pH 6.8, 0.4% (w/v) SDS.
4× Separating Gel Buffer	1.5 M Tris-HCl pH 8.8, 0.4% (w/v) SDS.
4× SDS Sample Buffer	130 mM Tris-HCl pH 6.8, 200 mM DTT, 4% (w/v) SDS, 0.025% (w/v) Bromophenol blue, 20% Glycerine.
10× SDS Running Buffer	250 mM Tris, 2 M Glycerine, 1% (w/v) SDS.
SDS Staining Solution	0.15 % (w/v) Coomassie Brilliant Blue R 250, 44 % Ethanol, 12 % Acetic acid
SDS Destaining Solution	10% Acetic acid
Agarose Gel Electrophoresis	
TAE buffer	40 mM Tris pH 8.0, 0.1% (v/v) Acetic acid, 1 mM EDTA.
Loading Buffer Orange	6% Sucrose, 4% Glycerin, 0.04% Orange.
RP-HPLC	
Solvent A	0.1% Trifluoroacetic acid in water
Solvent B	0.1% Trifluoroacetic acid in acetonitrile
MALDI-TOF	
Matrix Solution	Saturated sinapinic acid solution in 0.3 % TFA/acetonitrile (2:1)

3.1.6 **Other Materials**

Materials	Supplier
Omnifix syringe (50 mL)	B.Braun (Melsungen, DE)
Filtropur S (0.2 µm)	SARSTEDT (Nümbrecht, DE)
Dialysis tubing (MWCO: 5-8 kDa)	Spectrum Lab Inc. (Rancho Dominguez, USA)
Electroporation cuvettes	Bio-Rad (München,DE)
HiTrap Ni-NTA column (1 mL, 5 mL)	Pharmacia Biotech (Uppsala, SE)
Superdex 75/200 Gel filtration columns	Pharmacia Biotech (Uppsala, SE)
ME 24 Membrane filter	Whatman (Dassel, DE)
Ultrafiltration Membranes (NMWL: 30,000)	Millipore (Billerica, USA)
Amicon Stirred Cells (50 mL, 100 mL)	Millipore (Billerica, USA)
Amicon Ultra-4 and -15 Centrifugal Filter Units	Millipore (Billerica, USA)
Eppendorf tubes (0.5 mL, 1.5 mL, 2.0 mL)	Eppendorf (Hamburg,DE)
Falcon Tubes (15 mL, 50 mL)	Falcon GmbH (Gräfeling-Locham, DE)
Illustra NAP-5 Columns	GE Healthcare (Buckinghamshire, UK)
Quartz cuvette (10 mm)	Hellma Optik GmbH (Jena, DE)
PRONTOSIL 120-5-C18-AQ 5 µm reverse phase column	Bischoff-Chrom (Leor, DE)
CrystalQuick 96-well low-profile plates	Greiner Bio-One GmbH (Frickenhausen, DE)
V-shaped 96-well plates	Greiner Bio-One GmbH (Frickenhausen, DE)
96-well COC protein crystallization micro plates (Product #3553)	Corning GmbH (Kaiserslautern, DE)
Linbro Plate	Hampton Research (Aliso Viejo, USA)

3.1.7 **Instruments**

Instrument	Supplier
ÄKTAprime Automated Liquid Chromatography System	Pharmacia Biotech (Uppsala, SE)
Microfluidizer 110S	Microfluidics (Newton, MA, USA)
Excella® Benchtop Incubator Shakers	New Brunswick Scientific (Edison, USA)

Minitron incubator	Infors (Bottmingen, CH)
Varioklave Steam Sterilizer	H+P Labortechnik GmbH (Oberschleißheim, DE)
PCR Sprint Temperature Cycling System	Thermo Scientific (Waltham, USA)
Eppendorf benchtop Centrifuge 5415D/5804R	Eppendorf (Hamburg, DE)
Avanti J20-XP Centrifuge	Beckman Coulter (Palo Alto, USA)
Optima L-70K Ultracentrifuge	Beckman Coulter (Palo Alto, USA)
Gel Doc XR System	Bio-Rad (München,DE)
Thermomixer comfort	Eppendorf (Hamburg, DE)
NanoDrop ND-1000 Spectrophotometer	PEQLAB (Erlangen, DE)
650E Advanced Protein Purification System	Waters (Eschborn, DE)
High pressure liquid chromatography (HPLC)	Waters (Eschborn, DE)
HPLC-ESI-MS	Agilent and Finnigan
MALDI-TOF-MS	Applied Biosystem (Darmstadt, DE)
E. coli Pulser Electroporation device	Bio-Rad (München,DE)
SDS-PAGE Chamber	Bio-Rad (München,DE)
Agarose Gel Chamber Horizon 58	Biometra (Göttingen, DE)
UV/Visible Spectrometer DU 640	Beckman Coulter (Palo Alto, USA)
Milli-Q Water System	Millipore (Eschborn, DE)
Vacuum Pump	Ilmvac (Ilmenau, DE)
Stirring Device MR-3000	Heidolph Instruments (Schwabach, DE)
pH-Meter 766 Calimatic	Knick (Berlin, DE)
Mosquito Crystallization Robot	TTP LabTech (Melbourn, UK)
Rock Imager 1000	Formulatrix (Waltham, USA)
Nonius / Bruker AXS MICRO Star	Bruker AXS (Karlsruhe, DE)
Rigaku MicroMax-007 HF	Rigaku Europe (Kent, UK)

3.2 **Methods**

3.2.1 **PCR and Plasmid Construction**

3.2.1.1 *PCR amplification of phzE and phzD*

*phz*E and *phz*D genes were amplified by Polymerase Chain Reaction (Saiki et al., 1985) using plasmids containing phenazine operons as templates (Table 3.1). Primers were designed manually and ordered from MWG biotech (Table 3.2). The reaction mixture was prepared (Table 3.3) and PCR program was carried out by a PCR Sprint Temperature Cycling System (Thermo Scientific) (Table 3.4).

Target gene	Name of Plasmid	Description
phzE B.lata 383	pKSII-phzAll-B.lata383	A pKSII plasmid containing phz operon from Burkholderia lata 383
phzD P.fl	pT7-6-AG-P.fl.	A pT7-6 plasmid containing phzA-G genes from Pseudomonas fluorescens

Table 3.1: Target genes and template plasmids.

Name	Sequence (5'-3')	Restriction site	Tm (ºC)
phzE_Blata_for	AGGTGCTCATATGAATGCCGCTCC	NdeI	66.3
phzE_Blata_rev	CGTGAAGGATCCTTAGGCGGTCAACG	BamHI	68.0
PhzD_NdeI_for	GCAGCCATATGACCGGCATTCCATCGATCGT CC	NdeI	73.2
PhzD_XhoI_rev	CAGCCGGATCCTCGAGTCATAGCACCTCATC GGT	XhoI	74.3

Table 3.2: List of PCR primers.

Composite (concentration)	Volume (µL)
Template DNA (50-100 ng/µL)	5
Forward primer (25 pmol/µL)	1
Reverse primer (25 pmol/µL)	1
Phusion HF buffer (5x)	10
dNTPs mix (10 mM)	1
Phusion DNA Polymerase	0.5
Milli-Q water	31.5

Table 3.3: Composition of PCR reaction mixture.

Step	Cycle	Temperature (ºC)	Time (s)
Denaturation	1	98	30
Denaturation		98	10
Annealing	30	60	20
Extension		72	60
Final step elongation	1	72	300
Hold	1	4	Hold

Table 3.4. PCR program for amplifying *phzE* and *phzD*.

3.2.1.2 Purification of PCR products

The PCR products of target genes were purified by agarose electrophoresis as described below:

1) 2 g agarose was dissolved in 200 ml TAE buffer.
2) 10 µL RedSafe Nucleic acid staining solution (20,000x) was then added.
3) The gel solution was casted into an agarose gel chamber and allowed for polymerization.
4) 5 µL of the PCR product was mixed with appropriate amount of loading buffer and the sample was loaded into the well. 10 µL of 1 kb DNA ladder was loaded as a molecular weight marker.

5) The gel was run in TAE buffer at 10 V/cm till band separation is complete.

6) The result was checked under UV light and the band of interest was excised using a scalpel and transferred into a 2 mL eppendorf tube.

7) The DNA fragment was then extracted from the gel by a gel extraction kit.

3.2.1.3 Restriction digest and ligation

The PCR products and vectors were multiply digested with FastDigest restriction enzymes at 37 ºC for 1 h and the reaction mixtures were prepared as described (Table 3.5). The reactions were stopped by incubating at 80 ºC for 5 min. The resulting DNA products were identified and purified on agarose gel.

	Vector (100 ng/µL) Digest	PCR product (≈15 ng/µL) Digest
DNA	10 µL	15 µL
FastDigest buffer (10x)	2 µL	3 µL
NdeI	1 µL	1 µL
BamHI (XhoI)	1 µL	1 µL
Milli-Q water	6 µL	10 µL
Total	20 µL	30 µL

Table 3.5: Reaction mixture for restriction digests.

To construct the plasmid, the DNA insert was mixed with the vector at 5:1 ratio. 5 unit of T4 DNA ligase was applied and the total volume of ligation mixture was adjusted to 20 µL (Table 3.6). The ligation mixture was left at room temperature overnight and was subsequently incubated at 65 ºC for 10 min in order to deactivate the ligase.

Composite	Volume (µL)
Insert (≈ 5 ng/µL)	10
Vector (≈ 50 ng/µL)	1
Ligation buffer (5x)	4
T4 DNA ligase	1
Milli-Q water	4
Total	20

Table 3.6: Reaction mixture for ligation.

3.2.1.4 Preparation of competent cells

100 mL LB culture were inoculated with overnight culture of the desired bacterial strain at a starting $OD_{600} = 0.2$ and then incubated at 37 ℃ in a shaker till $OD_{600} = 0.6$. To stop cell growing, the culture was left on ice for 20 min. Ice-chilled cell culture was then transferred to sterile centrifuge tubes and was centrifuged for 15 min at 3000×g at 4 ℃. The supernatant was discarded.

Electro-competent cells (*E. coli* XL1-Blue)

1) The pellet was resuspended with 500 mL 10% glycerol (sterile) and centrifuged for 15 min at 3000*g* at 4 ℃. Supernatant was dicarded.
2) Repeat step 1.
3) The pellet was resuspended with 30 mL 10% glycerol (sterile) and centrifuged for 15 min at 3000*g* at 4 ℃. Supernatant was dicarded.
4) The cells were resuspended with 2.5 mL pre-chilled GYT medium, shock frozen with liquid nitrogen in 75 µL aliquots and stored at -80 ℃.

Heat-shock competent cells (Rosetta pLysS and Rosetta2 pLysS)

The cells were resuspended with 10 mL pre-chilled TSS medium, shock frozen with liquid nitrogen in 100 µL aliquots and stored at -80 ℃.

3.2.1.5 Plasmid transformation

Electroporation of *E. coli* XL1-blue

Plasmids were transformed into *E. coli* XL1-blue cells by electroporation as described below:

1) Approximatly 1 ng plasmid DNA was incubated with 75 µL of competent cells on ice in a pre-chilled electroporation cuvette.
2) The electroporation (25 µF, 200 Ω, 2.5 kV) was carried out using an *E.coli* Pulser from Biorad.
3) After electroporation the cells were diluted immediately with 1 mL LB medium and incubated at 37 ℃ for 45 min in a shaker.

4) 200 µL of the culture was spread on an LB agar plate containing the appropriate antibiotics.

5) The plate was incubated at 37 ºC for 16 hours.

Heat-shock transformation of *E. coli* Rosetta pLysS and *E. coli* Rosetta2 pLysS

The transformation of Plasmid into *E. coli* Rosetta pLysS and *E. coli* Rosetta2 pLysS competent cells was carried out as described below:

1) Approximatly 1 ng plasmid DNA was incubated with 100 µL of competent cells on ice for 30 min.

2) The cells were heat-shocked at 42 ºC for 90 seconds and then immediately transferred on ice and let stand for 2 min.

3) 1 mL LB medium was added to the cells and incubated at 37 ºC for 45 min in a shaker.

4) 200 µL of the culture was spread on an LB agar plate containing the appropriate antibiotics.

5) The plate was incubated at 37 ºC for 16 hours.

3.2.1.6 Site-directed Mutagenesis

Site-directed mutagenesis was performed according to the manual of the QuikChange II XL system (Stratagene). A plasmid containing wild-type *phz*E (*phz*D) was used as PCR template. All primers used were designed manually and ordered from MWG biotech (Table 3.7). The reaction mixture was prepared (Table 3.8) and PCR program was carried out by a PCR Sprint Temperature Cycling System as described (Table 3.9).

Name	Sequence (5'-3')	Tm (ºC)
phzE variations		
E241A	GCG ATCGCAAGGCATCCGACGAG	67.8
E244A	AAGGAATCCGACGCGCTGTACATGGTG	68.0
E379A	CGCGGTATCGGCAGT CATGGAGACG	69.5

E382A	CATGCG TCGCCATGACTTCCGATACCG	66.3
S217A	GACGATGAACCCGATCGCAGGG ACTTATCGGTATC	73.0
T369G	GTCGGCTCAGGGCTCGTCCGGCATT	71.2
S368AT369G	CATCGGCGTCGGCGCAGGGCTCGTCC	75.0
E201Q	GCCACGCCGCAGCGCCACCTGACG	74.7
W184G	GAGGTCGGCGCGTACGGGATCTTCGTGATTC	73.5
E251A	GTACATGGTGCTCGACGCAGAACTCAAGATGATGGCG	73.9
E251Q	GTACATGGTGCTCGATCAAGAACTCAAGATGATGGCGCG	73.7
phzD vatiations		
D38A	CGTACTGTTGGTACATGCCATGCAGCGCTACTTC	70.2
D38N	CGTACTGTTGGTACATAACATGCAGCGCTACTTC	65.4

Table 3.7: List of primers for site-directed mutagenesis (For each set of primers, only the primer for the sense strand is listed, the primer for the anti-sense strand is reverse-complementary to the one for the sense strand).

Composite	Volume (µL)
Template Plasmid (50-150 ng/µL)	1
Forward primer (25 pmol/µL)	2
Reverse primer (25 pmol/µL)	2
dNTPs mix (10 mM)	2
DMSO	1
MgCl2 (50 mM)	1
Phusion HF buffer (5x)	10
Milli-Q water	30.5
Phusion DNA Polymerase	0.5
Total	50

Table 3.8. PCR mixture for site-directed mutagenesis.

Step	Cycle	Temperature (ºC)	Time (s)
Denaturation	1	98	45
Denaturation	16	98	15
Annealing		60*	20

40

Extension		72	30/kb
Final step elongation	1	72	600
Hold	1	4	Hold

Table 3.9: Program setup for site-directed mutagenesis. (* The annealing temperature differs accordingly to primers used, it is generally 5 ºC lower than the Tm of the primer)

After the program was finished, 1 µL DpnI enzyme was added into the 50 µL reaction mixture and incubated for 3 h at 37 ºC (To digest the methylated template plasmid). 1 µL PCR product was then transformed into *E. coli* XL1-blue by electroporation (see chapter 3.2.1.5). Single colony was picked from the plate to inoculate 5 mL LB culture and the Plasmid DNA was prepared from 4 mL of the culture using a Miniprep kit. Desired mutation was then confirmed by in house DNA sequencing (see chapter 3.2.1.7).

3.2.1.7 DNA sequencing

The BigDyeDesoxy terminator cycle sequencing kit was used to check the DNA sequence. The reaction mixture for sequencing was prepared as shown in Table 3.10. The PCR program was designed as shown in Table 3.11 and was carried out by a PCR Sprint Temperature Cycling System.

composite	Volume (µL)
Plasmid DNA(100-200 ng/µL)	3
Sequencing primer (10 pmol/µL)	1
BigDye mix	4
Milli-Q water	2
Total	10

Table 3.10: PCR Reaction mixture for sequencing.

Step	Cycle	Temperature (ºC)	Time (s)
Denaturation	1	96	30
Denaturation	25	96	10
Annealing		50	5

Extension		60	240
Hold	1	4	Hold

Table 3.11: PCR program for sequencing.

After PCR, the reaction mixture was transferred to a 0.5 mL Eppendorf tube and was processed by ethanol precipitation as described:

1) 10 µL Milli-Q water, 2 µL EDTA (125 mM), 2 µL sodium acetate (3 M) and 50 µL ethanol were added to the reaction solution.
2) The tube was gently inverted 4 times and incubated at room temperature for 15 min.
3) The tube was then centrifuged at 13,000g for 20 min.
4) Supernatant was discarded and 200 µL cold 70% ethanol was added to wash the pellet.
5) The tube was centrifuged again at 13,000g for 10 min.
6) Supernatant was discarded and the tube was kept to open-air for 30 min.

The sequencing samples were analyzed by an in house sequencing facility.

3.2.1.8 Plasmids constructed

Plasmids constructed in this study are listed in Table 3.12.

Name of Plasmids	Insert*	Antibiotic resistance
PhzE_pET19mod	phzE	Ampicillin
E241A_pET19mod	phzE_E241A	Ampicillin
E244A_pET19mod	phzE_E244A	Ampicillin
E379A_pET19mod	phzE_E379A	Ampicillin
E382A_pET19mod	phzE_E382A	Ampicillin
S217A_pET19mod	phzE_S217A	Ampicillin
T369G_pET19mod	phzE_T369G	Ampicillin
S368AT369G_pET19mod	phzE_S368AT369G	Ampicillin
S217A S368AT369G _pET19mod	phzE_S217AS368AT369G	Ampicillin
E201Q_pET19mod	phzE_E201Q	Ampicillin
W184G_pET19mod	phzE_W184G	Ampicillin

E251A_pET19mod	phzE_E251A	Ampicillin
PhzD_pET19mod	phzD	Ampicillin
D38A_pET19mod	phzD_D38A	Ampicillin
D38N_pET19mod	phzD_D38N	Ampicillin

Table 3.12: List of plasmids (* unless otherwise indicated, *phzE* genes are originated from *Burkholderia lata* 383 and *phzD* genes from *Pseudomonas fluorescens* 2-73)

3.3 Gene over-expression and protein purification

3.3.1 Gene over-expression

Escherichia coli strain Rosetta 2 pLysS transformed with recombinant plasmid was grown at 37 ºC in Terrific Broth media containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol with vigorous shaking until OD_{600} of 0.7 was reached and then induced by adding 0.5 mM isopropyl-ß-D-thiogalactopyranoside (IPTG). Cells were further incubated at 20 ºC for 16 hours then harvested by centrifugation (20 min, 6000×g, 4 ºC). The cell pellets were stored at -20 ºC if not immediately used.

3.3.2 Protein purification

The pellet was re-suspended in a lysis buffer (Buffer A + 2 mM PMSF) and then lysed by passing three times through a microfluidizer. The lysate was then ultracentrifuged (150000×g, 45 min, 4 ºC) and the resulting supernatant was filtered through a 0.2 µm filter to further remove precipitates, and then passed through a pre-packed 5 ml Hitrap Chelating column that had been charged with 100 mM nickel chloride and pre-equilibrated with lysis buffer on an Äkta Prime FPLC system. The column was washed with Buffer A containing 2% Buffer B. The bound protein was then eluted with a gradient of 2% to 25% of Buffer B over a volume of 150 mL. Fractions containing pure PhzE were identified by SDS-PAGE, desired fractions were pooled and the concentration of the protein was determined with the Bradford assay (Bradford 1976). The pooled solution was dialyzed two times each time for 3 hrs against 2 L Dialysis buffer at 4 ºC, then incubated at 4 ºC overnight with addition of 2 mg TEV protease per 40 mg PhzE to remove the N-terminal His_6-tag. The protein was then

filtered through a 0.2 µm filter, concentrated to 8-10 mg/ml (18-20 mg/mL for PhzD) in an Amicon chamber using a 30 kDa (10 kDa for PhzD) cut off membrane. Size-exclusion chromatography was then performed for buffer exchange and removal of aggregates with a Superdex 200 (Superdex 75 in case of PhzD) gel-filtration column using Gel filtration buffer (flow rate= 2 mL/min). The purified protein was concentrated to 8-10 mg/mL (18-20 mg/mL for PhzD), aliquoted and stored at -80 °C.

3.3.3 Production of seleno-*L*-methionine labeled PhzE

Seleno-*L*-methionine was incorporated into PhzE protein by the methionine biosynthesis suppression method (Doublié 1997). *Escherichia coli* Rosetta 2 pLysS cells transformed with a pET-19mod plasmid containing N-terminal His$_6$-tagged PhzE were grown at 37 °C overnight in 30 mL Terrific Broth media containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol with vigorous shaking. The culture was then centrifuged at 6000g, 4 °C for 20 min. Cell pellet was then immediately re-suspended with 2 L M9 media containing 50 µg/mL ampicillin and 17 µg/mL chloramphenicol. The culture was grown at 37 °C in a shaker till OD$_{600}$ reached 0.6. Additional amino acids (100 mg/L *L*-lysine, 100 mg/L *L*-phenylalanine, 100 mg/L *L*-threonine, 50 mg/L *L*-isoleucine, 50 mg/L *L*-leucine, 50 mg/L *L*-valine and 60 mg/L seleno-*L*-methionine) were supplemented and the culture was then induced with 0.5 mM IPTG. Cells were further incubated at 20 °C for 16 hours then harvested by centrifugation (20 min, 6000×g, 4 °C). The seleno-*L*-methionine labeled PhzE was purified following the same protocol for native protein purification.

3.4 Analytical Methods

3.4.1 Agarose gel electrophoresis

1% agarose gel was used throughout this study. The gel was prepared by adding 2 g agarose into 200 mL TAE buffer. The mixture was then heated in a microwave to allow complete dissolving of agarose. 10 µL RedSafe Nucleic acid staining solution (20,000x) was added to the solution subsequently. The gel was cast into an agarose gel chamber and was cooled down and polymerized after 1 h.

After the gel polymerized, an appropriate amount of DNA sample mixed with loading buffer was loaded into the wells. A 1kb DNA ladder was loaded as standard. The sample was then run at 10 V/cm till band separation was complete and the result was examined by a Gel Doc XR System equipped with a camera.

3.4.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

Ten pieces of SDS-PAGE gel plates were prepared at a time using a casting chamber. The separating gel solution was poured into the chamber and was covered by a 50% isopropanol solution immediately. After the gel polymerized, the covering solution was carefully replaced by the stacking gel solution (See Table 3.13). SDS gel combs were inserted before the stacking gel polymerized in order to generate sample wells.

The protein sample was mixed with appropriate amount of sample buffer and was incubated at 95 ºC for 10 min prior to loading. A low molecular weight protein marker (phosphorlylase b: 97 kDa, albumin: 67 kDa, ovalbumin: 43 kDa, carboanhydrase: 30 kDa, trypsin inhibitor: 20.1 kDa and lysozyme: 14.4 kDa) was loaded as standard. The gel electrophoresis was carried out at 70 mA in SDS running buffer till the marker reached the lower edge of the gel. The gel was then stained with SDS staining solution followed by destaining with SDS destaining solution overnight.

	Stacking gel solution	Separating gel solution
4× Stacking Gel Buffer (mL)	7.5	-
4× Separating Gel Buffer (mL)	-	15
30% acrylamide (37.5:1) (mL)	4.5	30
Milli-Q water (mL)	17.7	14.5
10% APS (µL)	300	500
TEMED (µL)	30	50
Total (mL)	30	60

Table 3.13: Composition of SDS gel solutions.

3.4.3 **Determination of protein concentration**

Protein concentration was determined using a NanoDrop ND-1000 spectrophotometer. The extinction coefficient ε values of proteins were calculated by Protparam from the ExPASy online server (Gasteiger et al., 2005). For PhzE, ε was determined as 40715 $M^{-1}\cdot cm^{-1}$ (Absorption value 583 at 1 g/L protein concentration), and for PhzD, ε was determined at 43555 $M^{-1}cm^{-1}$ (Absorption value 1888 at 1 g/L protein concentration).

3.4.4 **Analytical GF-HPLC and RP-HPLC**

Analytical gel filtration (GF) and reverse phase (RP) High-performance liquid chromatography (HPLC) experiments were carried out on a Waters 600 system equipped with a Waters 600S controller, a 717 plus autosampler and a 2487 Dual λ Absorbance Detector. For GF-HPLC, a Superdex 200 HR 10/30 column (separation range 10-600 kDa) was used. GF buffer was used as running buffer and the flow rate was set to 1 mL/min.

For RP-HPLC, a PRONTOSIL 120-5-C18-AQ 5 µm reverse column was used. Unless otherwise indicated, solvent A was 0.1% TFA in H_2O and solvent B was 0.1 TFA in Acetonirile. The flow rate was set to 1mL/min and the chromatography was recorded under a typical gradient program as shown in Table 3.14. The integrated software package from Waters was used to process experimental data.

Time (min)	Solvent A%	Solvent B%
0-2	100	0
2-17	100-50	0-50
17-27	50-0	50-100
27-29	0	100
29-37	100	0

Table 3.14: Gradient program for RP-HPLC experiements.

3.4.5 **MALDI-TOF-MS**

Matrix Assisted Laser Desorbtion Ionization - Time of Flight (MALDI-TOF) experiments were performed with a Voyager-DE Pro Biospectrometry workstation (Applied Biosystems, Weiterstadt, Germany) to measure the mass of PhzE. Protein sample was first diluted with H_2O to a final concentration of 0.1 mg/mL and then mixed with equal volume of matrix solution (Table 3.15). 2 μL sample mixture were then spotted on a 100-well MALDI sample plate and dried on the bench for 10 min. MALDI-TOF spectra were measured with the following instrument settings: acceleration voltage = 25000 V, grid voltage = 91 %, guide wire = 0.3 % and extraction delay time = 1000 ns. To record and process the spectrum, the program Data Explorer TM (Voyager software package 5.1) was used.

Composite	Volume (μL)
1% (v/v) TFA	70
Milli-Q H_2O	630
Acetonitrile	300

Saturated with sinapinic acid, vortexed and the undissolved materials were spinned down

Table 3.15: Preparation of matrix solution for MALDI-TOF-MS.

3.4.6 **ESI-MS**

Electrospray Ionisation Mass Spectrometry (ESI-MS) was applied to measure the mass of ADIC. The measurement was carried out with an LCQ ESI mass spectrometer (Finnigan, SanJose, USA). Mass spectrometry was carried out in positive ion detection mode and data were collected in the m/z range between 0 and 600. Data processing and mass calculation were performed using the Xcalibur software package.

3.5 X-ray crystallography methods

3.5.1 Crystallization

3.5.1.1 Crystallization of ligand-free PhzE

Crystallization trials for PhzE were performed using the NeXtal JCSG PACT screening suite. The sitting drop method was applied with a Mosquito Crystallization Robot, where 0.1 µL protein (8.7 mg/mL and 4.3 mg/mL) was mixed with 0.1 µL reservoir. The plates were then incubated at 20 ºC. Initial crystals were observed in the drops with the higher protein concentration (8.7 mg/mL) and the reservoir composition of 0.1 M 1,3-bis[tris(hydroxymethyl)methylamino]propane (BTP) pH 7.5, 0.2 M potassium thiocyanate and 20% (w/v) PEG 3350.

To optimize the size and shape of ligand-free PhzE crystals, the hanging drop vapor diffusion method was applied with drops consisting of 2 µl protein and 1 µl precipitant solution at 20 ºC. Diffraction-quality crystals were obtained with a protein solution containing 8.7 mg/mL PhzE, 1 mM $MgCl_2$, 10 mM glutamine, and a reservoir consisting of 0.1 M BTP pH 7.0, 0.2 M potassium thiocyanate and 22% (w/v) PEG 3350. The colorless crystals possess a hexagonal shape and grew to a size of 0.5 × 0.3 × 0.3 mm in about one week. Seleno-L-methionine labeled ligand-free PhzE crystals were obtained at similar conditions (with additional 5 mM 2-mercaptoethanol in the protein solution). Crystals of ligand-free PhzE (Se-Met PhzE) were briefly washed in a cryo solution consisting of 0.1 M BTP pH 7.0, 0.2 M potassium thiocyanate, 25% (w/v) PEG 3350 and 5% (w/v) PEG 400 prior to plunging into liquid nitrogen.

3.5.1.2 Crystallization of ligand-bound PhzE

To produce ligand-bound PhzE crystals, the protein was first incubated with 50 mM $MgCl_2$, 20 mM chorismate and 25 mM L-Glutamine for 30 min on ice. 1 µl of reservoir was applied to 1 µl protein solution and the mixture was subsequently equilibrated against the reservoir containing 0.1 M HEPES buffer pH 7.1, 0.2 M $MgCl_2$, and 15% isopropanol at 4 ºC. Small crystals were then transferred to a freshly prepared drop by the macro-seeding technique, and the cube-shaped crystals grew to a full size of 0.4

mm^3 in about 3 days. Zinc-free crystals were obtained by first incubating PhzE with 10 mM EDTA on ice for 30 min. EDTA was then removed by a buffer exchange against normal GF buffer using illustra^{TM} NAP-5 column (GE Healthcare). Afterwards, the protein was re-concentrated to 9.5 mg/ml using a Vivaspin 500 column (Sartorius Stedim Biotech). The crystallization conditions of Zinc-free PhzE were the same as that of ligand-bound PhzE. The cryo solution for ligand-bound crystals contains 0.1 M HEPES buffer pH 7.1, 0.2 M $MgCl_2$, 15% isopropanol and 20% glycerol.

3.5.1.3 *Crystallization of PhzD and PhzD-D38A*

Crystals of PhzD and the PhzD-D38A mutant were obtained from 1 μL + 1 μL hanging drops equilibrated against a reservoir containing 0.1 M sodium cacodylate buffer pH 6.5, 0.2 M sodium acetate and 25% (w/v) PEG 4000 at 20 ºC. In order to remove bound buffer/precipitant molecules at the active center, crystals were incubated in a solution containing 0.1 M Bis-Tris pH 6.5, 20% (w/v) PEG 4000, 0.1 M NaCl for 60 min. Soaking experiments were carried out by adding 1 mM ADIC into the drop. The crystals were further incubated for 60 min and were briefly washed in a cryoprotecting solution consisting of 0.1 M Bis-Tris buffer pH 6.5, 40% (w/v) PEG 4000 and 0.1 M NaCl prior to data collection at 100 K.

3.5.2 **Data Collection**

Data were collected at 100 K either in house on an Rigaku Micro-MAX-007 HF generator with a MAR345 image plate detector, or at the Swiss Light Source (SLS, Villigen, Switzerland) on beam line X10SA equipped with a MAR225 CCD detector (PILATUS 6M pixel detector). Data collection strategies are summarized in Table 3.16.

3.5.3 **Data preparation**

Apart from native PhzD dataset, which was integrated using imosflm from the CCP4 Program Suite, all datasets were integrated using XDS (Kabsch 2010) and scaled using XSCALE (XDS Package). The final data in mtz-format were generated by XDSCONV (XDS package).

3.5.4 **Structure determination**

3.5.4.1 *MAD phasing of ligand-free PhzE*

Phases of PhzE in the apo form were determined from a Multiple-wavelength Anomalous Dispersion (MAD) dataset of Se-Met labeled PhzE collected at 3.6 Å. SHELXC (Sheldrick, 2010) was used to extract anomalous signals from the dataset and 25 Se atoms, each belonging to one of the two chains in the asymmetric unit, were located with SHELXD (Sheldrick, 2010). Initial phases were then generated with SHARP (delaFortelle & Bricogne, 1997) and improved by solvent flattening, using the program SOLOMON (Abrahams & Leslie, 1996) and DM (Cowtan, 1994) of the CCP4 suite. The final resulting phase information was then transferred to the native dataset and extended to full resolution with DM (Cowtan, 1994).

Project	Dataset	λ(Å)	Space group	Images	Oscillation (º/image)	Beamline and Detector	Note
PhzE	Open	1.0000	$P6_222$	80	0.5		MAD dataset
	SeMet_in	0.9796	$P6_222$	150	0.75		
	SeMet_pe	0.9790	$P6_222$	280	0.75		
	SeMet_re	0.9780	$P6_222$	150	0.75	X10SA +MAR225	
	Closed_Zn	0.9999	$P2_12_12$	258	0.4		Closed form with Zn^{2+}
	Closed	0.9790	$P2_12_12$	200	0.5		
PhzD	native	1.5418	$P4_32_12$	433	0.25	Rigaku+ MAR345	
	native+ADIC inhouse	1.5418	$C222_1$	300	0.5	Rigaku+ MAR345	
	native+ADIC	1.0000	$C222_1$	1000	0.25	X10SA +PILATUS	
	native_emp inhouse	1.5418	$C222_1$	369	0.5	Rigaku+ MAR345	

native_emp	1.0000	C222$_1$	1000	0.25	X10SA +PILATUS
D38A+ADIC inhouse	1.5418	P2$_1$2$_1$2$_1$	202	0.5	Rigaku+ MAR345
D38A+ADIC	1.0400	P2$_1$2$_1$2$_1$	480	0.25	X10SA +PILATUS
D38N+ADIC inhouse	1.5418	P2$_1$2$_1$2$_1$	204	0.5	Rigaku+ MAR345
D38N+ADIC	1.0400	P2$_1$2$_1$2$_1$	480	0.25	X10SA +PILATUS

Table 3.16: Data collection statistics.

3.5.4.2 *Structure determination of ligand-bound PhzE by molecular replacement*

The structure ligand-bound form of PhzE was solved by molecular replacement method with PHASER (McCoy et al. 2007b) and MOLREP (A. Vagin & Teplyakov 1997). Coordinates of one MST domain and GATaseI domain from ligand-free PhzE were used as separate search models. Solution was successfully identified by searching for two copies of MST domain and GATaseI domain, respectively, in one asymmetric unit.

3.5.4.3 *Structure determination of PhzD by molecular replacement*

The native PhzD structure was determined by molecular replacement method using program PHASER (Airlie J McCoy et al. 2007b) and MOLREP (A. Vagin & Teplyakov 1997). Coordinates from one chain of the published PhzD structure from *Pseudomonas aeruginosa* was used as search model (Parsons et al., 2003a). One copy of the protein molecule was found in one asymmetric unit for space group type P4$_3$2$_1$2 and C222$_1$, while two copies were found in P2$_1$2$_1$2$_1$ crystal form.

3.5.5 **Model building and refinement**

5% of total reflections from all datasets used for refinement were choosen at random for calculation on R$_{free}$ and remained the same for the same crystal forms throughout this study.

3.5.5.1 Model building of ligand-free PhzE

In case for ligand-free PhzE, the initial C_α positions were traced in the program O (Jones et al., 1991). After cycles of refinement against experimental data using REFMAC5 (Murshudov et al. 1997), the improved model was further corrected manually in COOT (Emsley et al. 2010). At the later stage of refinement, water molecules were located by COOT and TLS-refinement was introduced using each MST, GATaseI domain and the linker region as a separate TLS body. The final model of ligand-free PhzE was generated after one round of refinement in phenix.refine (Adams et al., 2010).

3.5.5.2 Model building of ligand-bound PhzE

After molecular replacement, the resulting coordinates were put in REFMAC5 for one round of rigid body refinement. The output model was then used as initial model for ligand-bound PhzE and was corrected manually in COOT. Coordinates for benzoate and pyruvate were retrieved from the COOT monomer library. The restraint library for glutamyl-cysteine was generated with PRODRG (Schüttelkopf & van Aalten, 2004) and was introduced to the model as a non-native amino acid residue. REFMAC5 was used for refinement throughout the model building process.

3.5.5.3 Model building of PhzD

After molecular replacement, the model was manually corrected in COOT and refined with REFMAC5. To better interpret the experimental data regarding the ligands (DHHA or ADIC) before incorporating ligand into the structure, the $|F_o-F_c|$ density map of the ligand was generated. The restraint libraries for DHHA and ADIC were generated using the Dundee PRODRG server.

3.5.6 Search for an ammonia channel

The software Caver (Petrek et al., 2006) was used to search for the ammonia channel that connects the two active sites of PhzE. In this study, the Caver 2.0 PyMOL plugin was employed. Before the calculation was performed, the model of ligand-bound PhzE was modified in COOT. In order to visualize the full channel, the side chain of the gatekeeper residue E251 was shifted to its second conformation adopted in the ligand-

free structure. The model was then displayed with PyMOL and the sulfur atom of the glytamyl-cysteine moiety from chain A was given as the starting point of the channel.

3.5.7 **X-ray Fluorescence scan**

To identify the metal bound to the ligand-bound PhzE, an X-ray fluorescence scan experiment was carried out. The crystal of untreated ligand-bound PhzE (protein has not been treated with EDTA prior to crystallization) was mounted on the beamline X10SA of the SLS equipped with a Ketek Si-drift fluorescence detector. The X-ray fluorescence spectrum of the crystal at Zinc and Nickel absorption K-edge was recorded (Figure 4.23 B).

3.5.8 **Small Angel X-ray Scattering (SAXS) measurement of PhzE**

SAXS experiments were performed at beamline X33 of the DORIS III storage ring (DESY and EMBL Hamburg, Germany) (Roessle et al., 2007). The scattering data were recorded by means of an image plate with online readout (MAR345, MarResearch, Norderstedt, Germany). The automated sample handling robot was used for loading protein solution in the X-ray beam (Round et al., 2008). The scattering patterns were measured using a sample - detector distances of 2.4 m, covering the range of momentum transfer $0.1 < s < 4.5$ nm^{-1} ($s = 4\pi \sin(\theta)/\lambda$, where θ is the scattering angle and $\lambda = 0.15$ nm is the X-ray wavelength). In order to check for inter protein interactions PhzE was measured at 3 and 6 mg/ml concentration. Repetitive measurements of 120sec of the same protein solution were performed in order to check for radiation damage. No aggregation was found during the initial 120 sec exposure. This initial exposure frame was taken for further analysis. The data were normalized to the intensity of the incident beam; the scattering of the buffer was subtracted and the difference curves were scaled for concentration. Corresponding datasets were merged according to the data quality. All data processing steps were performed using the program package PRIMUS (Konarev et al., 2003). The forward scattering I(0) and the radius of gyration Rg were evaluated using the Guinier approximation (Guiner & Fournet 1955) assuming that at very small angles ($s < 1.3/Rg$) the intensity is represented by $I(s) = I(0) \exp(-(sRg)2/3)$. These parameters were also computed from the entire scattering

patterns using the indirect transform package GNOM (Svergun, 1993), which also provides the distance distribution function *p(r)* of the particle as defined:

$$p(r) = 2\pi \int I(s) sr \sin(sr) ds$$

The molecular mass of PhzE was calculated by comparison with the forward scattering from the reference solution of bovine serum albumin (BSA). From this procedure a relative calibration factor for the molecular mass (MM) can be calculated using the known molecular mass of BSA (66.4 kDa) and the concentration of the reference solution by applying

$$MM_p = I(0)_p / c_p \times \frac{MM_{st}}{I(0)_{st} / c_{st}}$$

where $I(0)_p$, $I(0)_{st}$ are the scattering intensities at zero angle of the studied and the BSA standard protein, respectively, MM_p, MM_{st} are the corresponding molecular masses and c_p, c_{st} are the concentrations. Errors have been calculated from the upper and the lower I(0) error limit estimated by the Guinier approximation.

Low-resolution models of PhzE were built by the program DAMMIN (Svergun, 1999), which represents the protein as an assembly of dummy atoms inside a search volume defined by a sphere of the diameter D_{max}. Starting from a random model, DAMMIN employs simulated annealing to build a scattering equivalent model fitting the experimental data $I_{exp}(s)$ to minimize discrepancy:

$$\chi^2 = \frac{1}{N-1} \sum_j \left[\frac{I_{exp}(s_j) - c I_{calc}(s_j)}{\sigma(s_j)} \right]^2$$

where *N* is the number of experimental points, *c* a scaling factor and $I_{calc}(s_j)$ and $\sigma(s_j)$ are the calculated intensity and the experimental error at the momentum transfer s_j, respectively. Final *ab initio* shape models for PhzE were obtained applying a two-fold symmetry as constrain to the DAMMIN program. The values of the fitting procedure are in the range of 1.50 ± 0.02. Ten independent DAMMIN reconstructions are

superimposed using the program package SUBCOMP and DAMAVER (Svergun 1999) and a most probable PhzE model derived after filtering to the molecular mass of the PhzE dimer.

See chapter 7.3 and 7.4 for more detailed introduction to the principles of protein crystallography and SAXS.

3.6 Biochemical methods

Unless otherwise stated, prior to use in the biochemical measurements, all PhzE samples were treated with EDTA to remove bound metals as described below.

Concentrated PhzE (8-10 mg/mL) was first incubated with 10 mM EDTA on ice for 30 min, EDTA was then removed by a buffer exchange against normal GF buffer using illustra™ NAP-5 column (GE Healthcare). Afterwards, the protein was re-concentrated using a Vivaspin 500 column (Sartorius Stedim Biotech).

3.6.1 pH optimum of PhzE

PhzE's activity as a function of pH was analyzed in order to identify the optimum condition for biochemical assays. The reaction mixture contained 50 mM buffer substance (citric acid for pH 6.0 and 6.5, Tris for pH 7.0, 7.5, 8.0, 8.5 and 9.0), 1 mM $MgCl_2$, 10 mM L-glutamine and 40 µM chorismate. The reaction was initiated by adding 1 µL PhzE (8.7 mg/mL) and monitored by continuous spectrophotometric assay in a 1 mL quartz cuvette (Hellma Optik GmbH) at the wavelength of 280 nm with a Beckman DU640 spectrophotometer at 25 °C. The reaction curves at different pH were prepared and compared using GraFit data analysis software (Erithacus Software, UK).

3.6.2 Analysis the Mg^{2+} dependence of PhzE

The reaction solution was prepared in a 1 mL quartz cuvette (Hellma Optik GmbH) to contain 100 mM HEPES-Na pH 7.0, 25 mM L-glutamine, 50 µM chorismate and various amounts of $MgCl_2$ (from 0 to 50 mM). The reaction was initiated by adding 1 µL EDTA-treated PhzE (9.0 mg/mL) to a final concentration of 129 nM, and was followed spectrophotometrically for 5 min at 280 nm with a Beckman DU640 spectrophotometer

at 25 °C. The optimum Mg^{2+} concentration for PhzE was determined by comparing the reaction curves using GraFit data analysis software (Erithacus Software, UK).

3.6.3 UV spectra of the PhzE reaction

To monitor the accumulation of ADIC, a spectrophotometric scan of the PhzE reaction was performed in the middle to near UV range (from 230 nm to 360 nm). In this experiment, the reaction mixture containing 100 mM HEPES-Na pH 7.0, 20 mM *L*-glutamine, 1 mM $MgCl_2$ and 100 µM chorismate was prepared in a 1 mL quartz cuvette (Hellma Optik GmbH). 1 µL PhzE (8.7 mg/mL) was then added to the mixture to a final PhzE concentration of 124 nM. After brief mixing with a plastic stirrer, the quartz cuvette was put into a Beckman DU640 spectrophotometer and UV spectra were recorded after 0, 5, 10, 20, 30 and 60 min.

3.6.4 Determination of the extinction coefficient of ADIC

1 mL reaction mixture containing 100 mM HEPES-Na at pH 7.0, 25 mM *L*-glutamine, 1 mM $MgCl_2$ and 200 µM chorismate was prepared. Reaction was initiated by adding 1 µL PhzE (8.7 mg/mL) to a final concentration of 124 nM. After each 40 and 80 min, 50 µL of reaction mixture was injected and analyzed on RP-HPLC (see description in analytical methods). The area under the peak of chorismate and ADIC on the spectra was calculated by the integrated software. Due to the fact that the decrease in amount of chorismate is equivalent to the increase in amount of ADIC (Formula 3.1), given the $\varepsilon_{280} = 2,750$ $M^{-1} \cdot cm^{-1}$ for chorismate (Gould & Eisenberg, 1991), it was possible to calculate the ε_{280} of ADIC using the formula below:

$$\varepsilon_{ADIC} = \varepsilon_{chr} \times (A_{ADIC80} - A_{ADIC40}) / (A_{chr40} - A_{chr80})$$

Formula 3.1: Calculation of the extinction coefficient value of ADIC. A: Calculated area under respective peak.

3.6.5 Michaelis-Menten kinetics of PhzE

Activity assays of native PhzE and the mutants were performed by a modification of a method described previously (Morollo & Bauerle, 1993). Reaction mixtures containing

100 mM HEPES-Na (pH 7.0), 1 mM $MgCl_2$, 20 mM glutamine, the enzyme, and variable amounts of chorismate were prepared in a 1 mL quartz cuvette (Hellma Optik GmbH). For kinetic experiments, the enzyme was used at 14.3 pM (final concentration) and chorismate varied within the range of 1 µM and 50 µM. The initial rate of ADIC formation was monitored by a continuous spectrophotometric assay at the wavelength of 280 nm with a Beckman DU640 spectrophotometer at 25 °C. The extinction coefficient for chorismate at 280 nm ε_{280} = 2,750 $M^{-1} \cdot cm^{-1}$ was taken from the literature (Morollo & Bauerle, 1993) and the value for ADIC at 280 nm ε_{280} = 11,490 $M^{-1} \cdot cm^{-1}$ was measured in this study (see chapter 3.6.4). K_M and k_{cat} values were determined by fitting the data to the Michaelis-Menten equation using GraFit data analysis software (Erithacus Software, UK). Experiments were repeated at least three times.

3.6.6 Inhibition analysis of PhzE

Inhibition of PhzE by transition metals Zn^{2+}, Ni^{2+} and Mn^{2+} were tested. The reaction solution was prepared in1 mL quartz cuvette (Hellma Optik GmbH) containing 100 mM HEPES-Na pH 7.0, 10 mM *L*-glutamine, 1 mM $MgCl_2$ and 50 µM chorismate. Additionally, 1 mM $ZnCl_2$, $NiSO_4$ or $MnSO_4$ (1 mM $MgCl_2$ for native reaction) was supplemented respectively to each of the reaction set-ups. The reaction was then initiated by adding 1 µL PhzE (4.8 mg/mL) to a final concentration of 69 nM, and was followed with a continuous spectrophotometric assay at 280 nm with a Beckman DU640 spectrophotometer at 25 ºC.

3.6.7 Inhibition of PhzE by *L*-tryptophan, DHHA and PCA

To test the potential inhibitors of PhzE, control reactions were carried out in 1 mL quartz cuvette (Hellma Optik GmbH) containing 100 mM HEPES-Na pH 7.0, 20 mM glutamine, 1 mM $MgCl_2$ and 50 µM chorismate. In addition, *L*-tryptophan (20, 50 and 100 µM), DHHA (100 and 200 µM) or PCA (200 µM and 2 mM) were added to the reaction solution respectively. 1 µL PhzE (8.5 mg/mL) was then added and the reaction was monitored spectrophotometrically at 280 nm with a Beckman DU640 spectrophotometer at 25 °C.

3.6.8 **PhzE and PhzD coupled enzymatic assay**

A PhzE/D coupled assay was designed to show that PhzD could convert the PhzE product into DHHA. The reaction solution was prepared to contain 100 mM HEPES-Na pH 7.0, 25 mM glutamine, 1 mM $MgCl_2$ and 100 µM chorismate. 1 µL PhzE (8.3 mg/mL) was then added to the solution and mixed gently with a plastic stirrer. The reaction mixture was subsequently kept under room temperature for 10 min. 1 µL PhzD (2.7 mg/mL) was then supplemented and after 10 more min, the reaction was quenched by adding TFA to a final concentration of 0.5% (v/v). 30 µL of the resulting solution was injected and to RP-HPLC (see analytical methods). The chromatogram of PhzE/D coupled reaction was recorded and analyzed with HPLC integrated program and GraFit data analysis software (Erithacus Software, UK).

3.6.9 **Isothermal titration calorimetry**

Isothermal Titration Calorimetry (ITC) measurements were employed to determine the dissociation constants of chorismate to PhzE with and without Mg^{2+}. Measurements were performed at 25 °C with a VP-ITC system (MicroCal LLC). For measurement without Mg^{2+}, PhzE was diluted to 100 µM with a buffer containing 50 mM Tris (pH 7.5) and 150 mM NaCl. For measurement with Mg^{2+}, the buffer contained an additional 1 mM $MgCl_2$. Chorismate was dissolved in the same buffer as used for the protein to a final concentration of 1 mM. All buffers and solutions were degassed prior to use. The system was equilibrated for at least 30 min before the measurement started, and then 2 µl substrate solution was injected initially, followed by injections of 8 µL at 4 min intervals. Raw data containing enthalpy changes corresponding to the association of chorismate to PhzE were recorded and analyzed with Origin version 7 (OriginLab Corporation).

A more detailed introduction of ITC is given in chapter 7.5.

3.6.10 **Production and purification of ADIC**

ADIC was produced enzymatically in a 2 mL reaction solution containing 5 mM chorismate, 120 mM Glutamine, 10 mM $MgCl_2$ and 50 mM HEPES pH 7.1. The reaction was initiated by adding PhzE into the mixture to a final concentration of 5 µM. After one

hour, the reaction mixture was shock-frozen with liquid nitrogen in 200 µL aliquots and stored at -80 ºC.

The purification of ADIC was performed by RP-HPLC and identified by ESI-MS as described (see chapter 3.4). 200 µL of the reaction mixture were injected at each run. The peak according to the retention time of ADIC was collected manually in 2 mL eppendorf tubes and was shock-frozen immediately in liquid nitrogen.

4 RESULTS AND DISCUSSION

This section has been arranged into two parts to present the results of two proteins: the ADIC-synthase PhzE and the isochorismatase PhzD, respectively. In part one, the crystal structures of the ligand-free PhzE and the ligand-bound PhzE are presented and the conformational movement upon ligand binding was analyzed, followed by the data of biochemical and biophysical studies of the protein function. In part two, the crystal structures of PhzD are presented and discussed.

Part I: Structures and function of PhzE

4.1 Determination of PhzE crystal structures

Crystallographic studies were originally conducted on PhzE from the *Pseudomonas* strains. However, it was impossible to obtain diffraction quality crystals of PhzE from *P. aeruginosa* and *P. chlororaphis*. Therefore, the crystallization trials were extended to PhzE from *Burkholderia lata* 383.

4.1.1 Sequence analysis

The sequence of PhzE was retrieved from the complete genome sequence of *Burkholderia lata* 383 (US DOE Joint Genome Institute) and was submitted and analyzed by different bioinformatic tools. Results from the Protparam online proteomics server (Gasteiger et al., 2005) revealed that PhzE has a molecular mass of 69882.1 Da and theoritical pI of 5.34. PSORT prediction indicated that PhzE probably locates in bacterial cytoplasm (Nakai & Kanehisa, 1991).

A Blast (National Center for Biotechnology Information web tool) search against Non-redundant protein sequence database indicated that PhzE is conserved among phenazine producers and it contains two putative conserved domains. The N-terminal part of the protein resembles a chorismate binding domain and the C-terminal part has a GATaseI domain. Among the top hits from blast results, anthranilate synthases with

fused MST/GATaseI domains were also identified to possess high similarity to PhzE. PhzE from three other bacterial origins, as well as anthranilate synthase with fused domains from two bacterial strains are listed in Table 4.1 and a complete sequence alignment of all six proteins (including PhzE from *B. lata* 383) was performed with Clustal X (Larkin et al., 2007) (Figure 4.1).

Protein	Origin	Query coverage (%)	Sequence identity (%)	E value
PhzE	*P. aeruginosa*	98	57	0
PhzE	*P. chlororaphis*	97	55	0
PhzE	*P. fluorescens*	97	55	0
Anthranilate synthase (fused MST/GATaseI domains)	*S. venezuelae*	95	49	3e-149
Anthranilate synthase (fused MST/GATaseI domains)	*S. coelicolor*	95	48	2e-136

Table 4.1: Results from Blast search. 5 hits are listed according to their max score (not shown). The E value describes the quality of the alignment; the lower the E value, the more significant is the match.

A Blast search of PhzE sequence against Protein Data Bank (PDB) was launched to search for structures that are potentially related to PhzE. Five proteins from different MST enzyme classes in the top hits are listed in Table 4.2. A complete sequence alignment was performed with Clustal X (Larkin et al., 2007) and is shown in Figure 4.2. All these five proteins are synthases (bacterial anthranilate synthase, isochorismate synthase, salicylate synthase, para-aminobenzoate synthase and aminodeoxychorismate synthase) and belong to MST enzyme family. Interestingly, all of them contain either only the MST domain or separated MST and GATaseI domains. This fact, together with the relatively high E value, low percentage of query coverage and low sequence identities, suggested that PhzE might have a distinct quaternary structure.
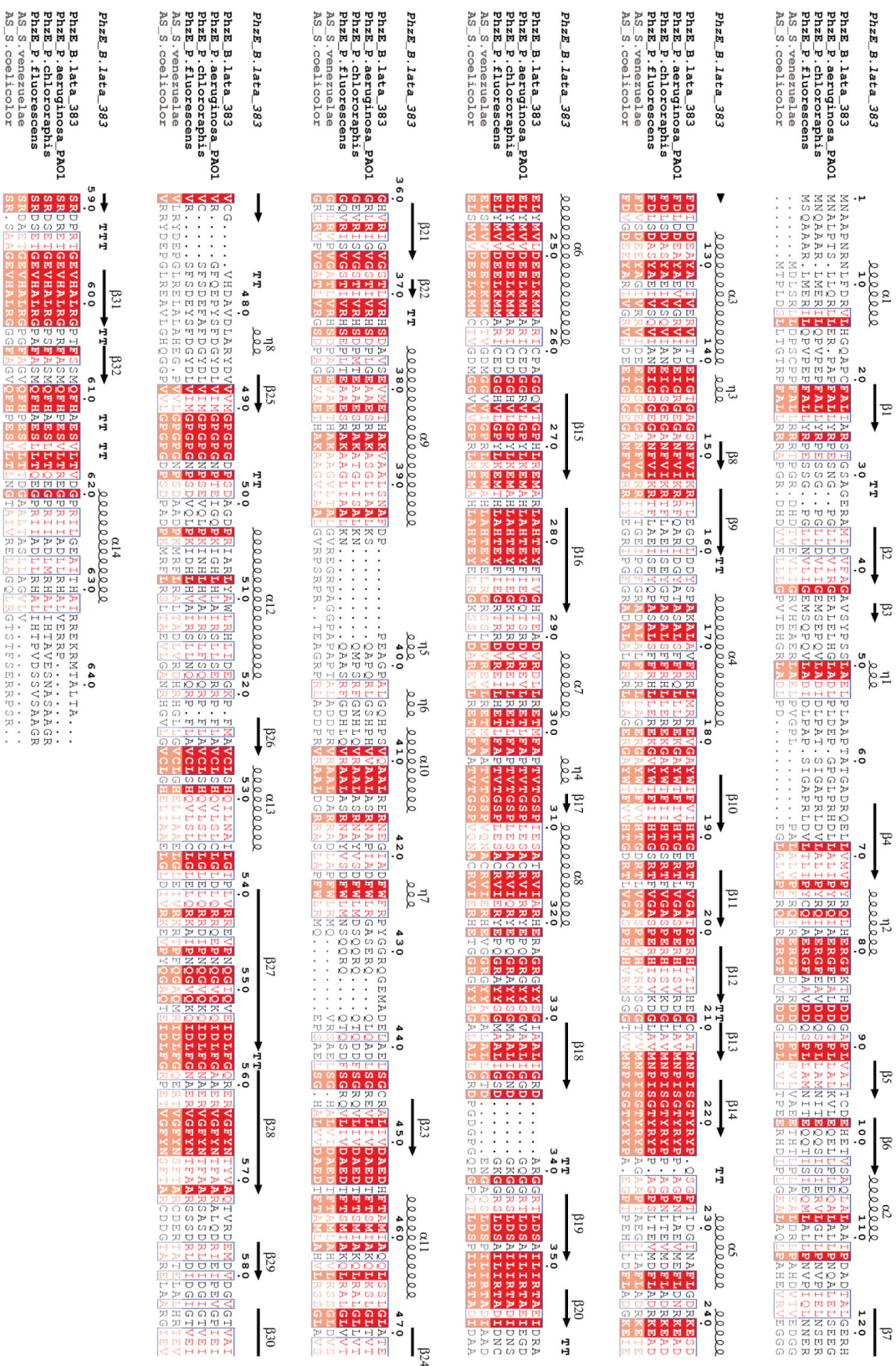
Figure 4.1: Sequence alignment between PhzEs and fused anthranilate synthases. Secondary structural elements of PhzE *B. lata* 383 (see chapter 4.2) is shown (on top of the alignment). Two of the related MST/GATase1-fused anthranilate synthases are shown in grey. The figure was prepared using ESPript 2.2 (Gouet et al., 2003)
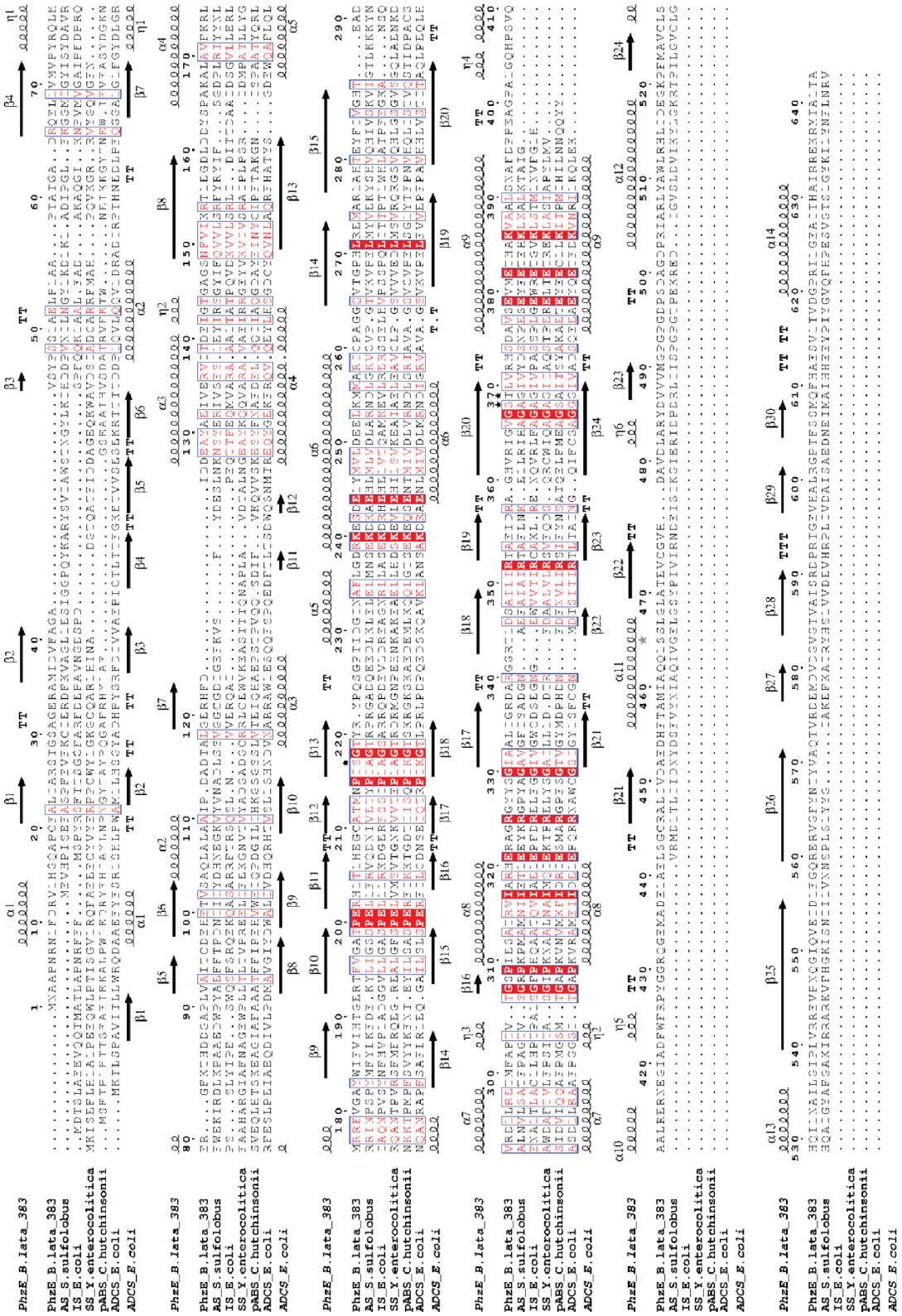
Figure 4.2: Structure-based sequence alignment of various MST enzymes. Secondary structural elements of PhzE (see chapter 4.2) and its most distant homologue from the alignment (ADC synthase) are shown and compared (on top and at bottom of the alignment). Sequence of the AS is composed of two chains, the TrpE subunit and the TrpG subunit, respectively. The figure was prepared using ESPript 2.2 (Gouet et al., 2003)

Protein	Origin	PDB ID	Query coverage (%)	Sequence identity (%)	E value
Anthranilate Synthase	*S. solfataricus*	1QDL	33	26	9e-12
Isochorismate Synthase	*E. coli*	3HWO	34	24	8e-5
Salicylate Synthase	*Y. enterocolitica*	2FN0	50	24	4e-4
Para-aminobenzoate Synthase	*C. hutchinsonii*	3H9M	41	23	6e-4
Aminodeoxychorismate (ADC) Synthase	*E. coli*	1K0E	33	25	0.041

Table 4.2: Results from a Blast search against PhzE. 5 hits are listed according to their max score (not shown). The E value describes the quality of the alignment, the lower the E value, the more significant the match is.

4.1.2 Cloning, over-expression of *phzE* and protein purification

phzE gene was amplified from a pSKII plasmid containing phz operon, and was subcloned into a pET19mod vector (Figure 4.3).
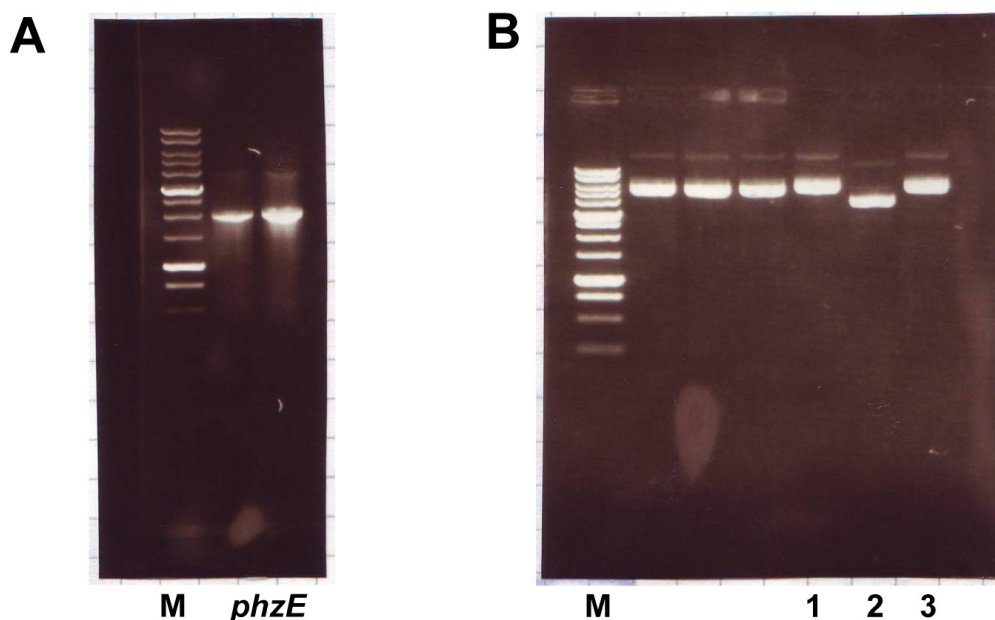

Figure 4.3: Cloning of *phzE*. (A) PCR amplification of *phzE*, checked on agarose gel. (B) Plasmid bearing *phzE* shown in lane 1 and 3, empty vector is shown in lane 2. M: 1 kb DNA ladder.

The 643-amino-acid protein PhzE was expressed as a TEV-cleavable N-terminal 6×His-tag fusion protein. The protein was purified with Ni^{2+} affinity chromatography (Figure. 4.4).
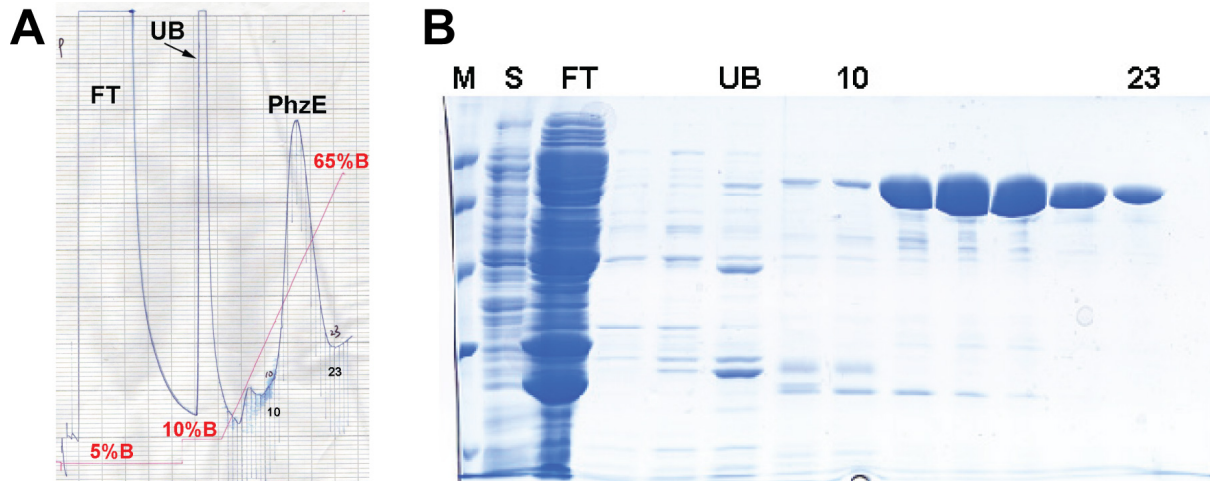

Figure 4.4: Purification of PhzE by Ni^{2+} affinity chromatography. (A) Chromatogram of the purification. FT: Flow through from the column. UB: Elution of unspecifically bound proteins. Percentage of Buffer B in the elution buffer is indicated in red. (B) SDS-PAGE gel analysis. M: LMW marker. S: Supernatant after fluidizer. 10 and 23: the 10th and 23rd fractions of the elution, as marked in the chromatogram.

After affinity chromatography, TEV cleavage was performed at 4 °C in dialysis buffer for 16 hours, TEV enzyme was applied at a fixed molar radio to PhzE (1:20). After concentrated in an Amicon chamber (Millipore) using a 30 kDa cut off membrane, gel filtration chromatography was performed for removing TEV enzyme and the final buffer exchange. PhzE was eluted from a Superdex 200 HR (26/60) column (GE Healthcare) as a single peak with a retention time (RT) of 80 min (Figure 4.5). According to the calibration curve of the column (flowrate = 2 mL/min), the molecular weight of the molecule should be in the range between 120 and 180 kDa, indicating that PhzE (Mw ≈ 70 kDa) is a dimer in solution. In the end, at least 50 mg ≥95% pure PhzE was obtained from 1 L culture. The protein was then concentrated to 8-10 mg/mL in an Amicon chamber using a 30 kDa cut off membrane.
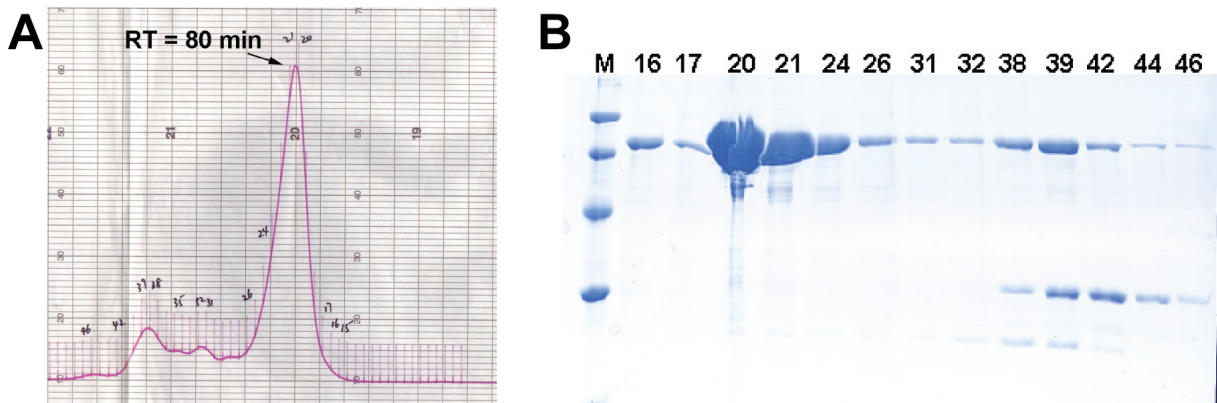
Figure 4.5: Gel filtration chromatography of PhzE. (A) The gel filtration chromatogram of PhzE. PhzE was eluted as a single peak with a retention time of 80 min. (B) SDS-PAGE gel result from gel filtration. M: LMW marker. The numbers indicate fractions according to the chromatogram. The impurities in fraction 38-46 might be TEV enzyme.

4.1.3 Analytical gel filtration and MALDI-TOF

4.1.3.1 Analytical gel filtration

In order to determine the oligomeric state of PhzE in solution more precisely, analytical HPLC experiment was performed. A gel filtration standard (Bio-Rad) was used as reference (Figure 4.6).
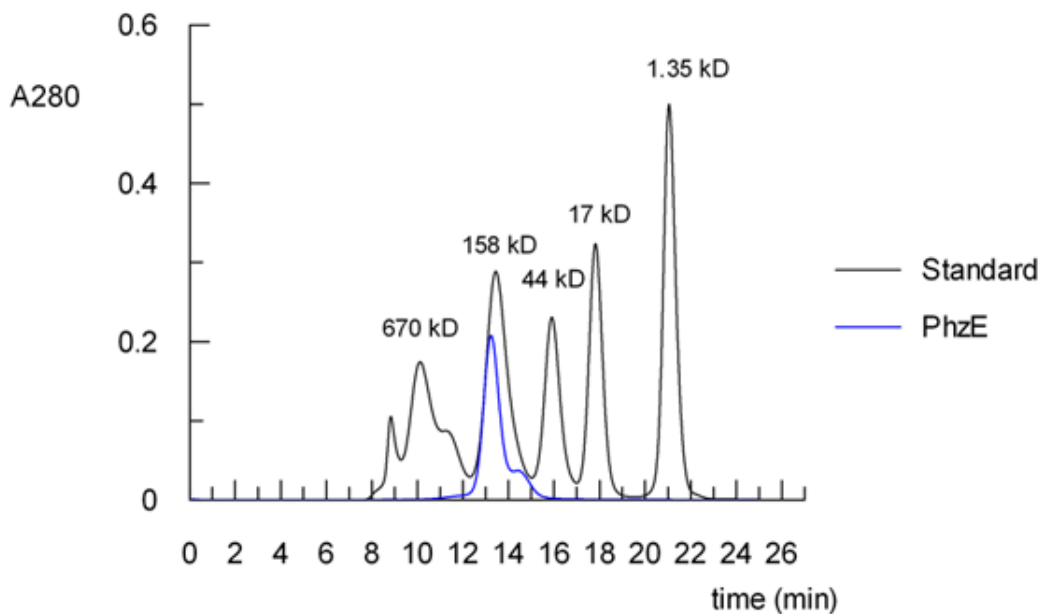

.**Figure 4.6: Analytical Gel filtration of PhzE. Flowrate was set to 1 mL/min.**

The elution statistics of the standard are shown in Table 4.3. The equation used for calculating PhzE molecular weight in solution was derived by single exponential fitting of the standard curve (Fomula 4.1). PhzE has a retention time of 13.23 min and the calculated molecular weight of PhzE is 167 kDa. As the single PhzE chain has 643 amino acids and has a calculated molecular weight of 70 kDa and taking account of the N-terminal 6-His tag, PhzE in solution presents as dimer.

Component	Molecular weight (Da)	Retention time (min)
Thyroglobulin (bovine)	670,000	10,11
γ-globulin (bovine)	158,000	13,45
Ovalbumin (chicken)	44,000	15,90
Myoglobin (horse)	17,000	17,82
Vitamin B12	1,350	21,05

Table 4.3: Elution statistics of the gel filtration standard.

$$Mw = 60767 \times e^{-0.45t}$$

Fomula 4.1: The fomula used to calculate molecular weight of PhzE. Mw: molecular weight in Dalton. t: retention time of PhzE.

4.1.3.2 MALDI-TOF of PhzE

Matrix Assisted Laser Desorption Ionization – Time of Flight (MALDI–TOF) mass spectrometry experiments were performed to examine the quality of purified PhzE and the efficiency of His_6-tag cleavage by TEV.

For His_6-tagged PhzE, two major peaks were observed from the spectrum (Figure 4.7 **A**), the higher one has a calculated mass of 72058 Da regarding the mono-protonated PhzE $[PhzE+H]^+$ and the lower one has a calculated mass of 36022 Da, representing the double-protonated protein $[PhzE+2H]^{2+}$. Apart from these two peaks, there were only minor peaks with relative intensity below 20% observed in the mass range between 10 kDa and 200 KDa.

After being treated with TEV enzyme, the mass of PhzE was examined again with MALDI–TOF. Two major peaks were detected in the mass range between 10 kDa and 200 KDa, the higher peak [PhzE+H]$^+$ has a mass of 70114 Da and lower one [PhzE+2H]$^{2+}$ has 36022 Da (Figure 4.7 **B**). From the results of the spectra, a 1944 Da mass difference between the His$_6$-tagged PhzE and TEV treated one could be observed. According to the map of pET19m vector, the amino acid sequence of N-terminal cleavable His$_6$-tag is N'–MGHHHHHHAENLYFQG–C' with a calculated mass of 1950 Da. Residues glycine, serine and histidine were left at the N-terminus of PhzE after TEV cleavage, the calculated molecular mass of N'-GSH-PhzE is 70125 Da. These results together indicate that the TEV cleavage was efficient.
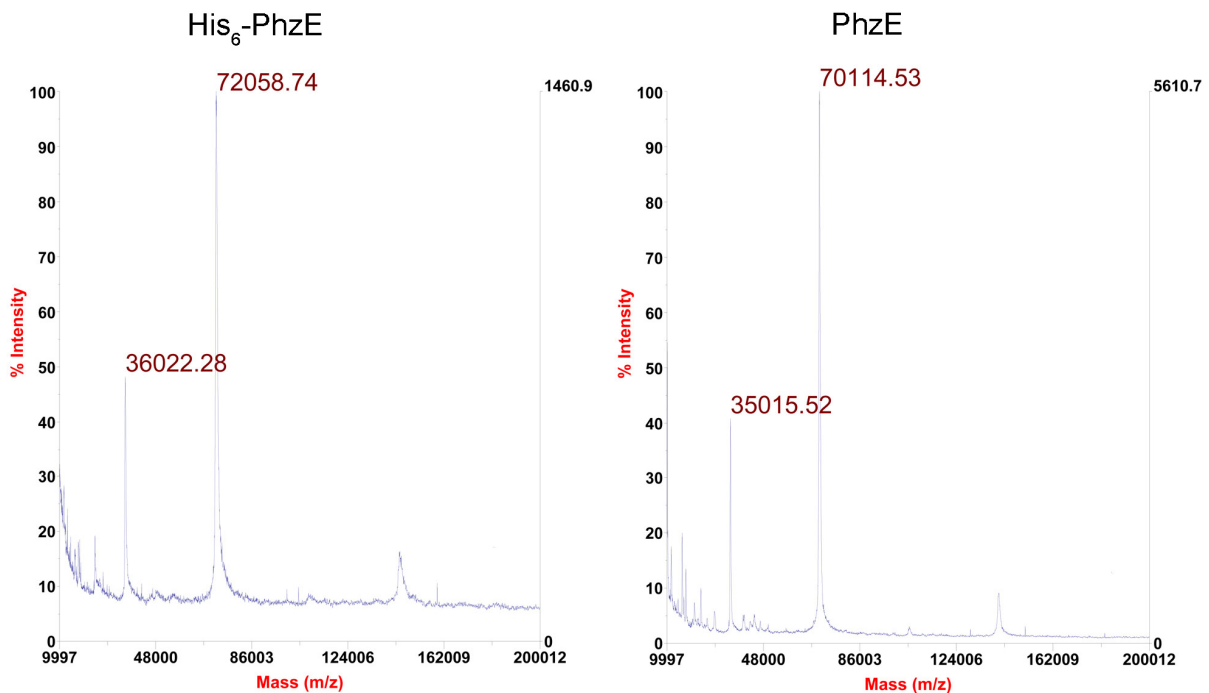

Figure 4.7: MALDI-TOF spectra of PhzE. A) PhzE with His$_6$-tag uncleaved. B) PhzE after TEV cleavage of His$_6$-tag.

4.1.4 **Crystallization of PhzE**

4.1.4.1 Crystallization of ligand-free PhzE

From the initial crystallization screen using the NeXtal PACT suite, hexagonal-shaped crystals (Figure 4.8 **A**) were obtained in wells containing PEG 3350 and potassium thiocyanate as precipitants (BTP buffer pH 8.5, 20% PEG 3350, 2.0 M potassium thiocyanate). The size and shape of these crystals was improved by optimizing these conditions (Figure 4.8 **B** and **C**). Crystals with best quality were obtained with mother-liquor containing BTP buffer pH 7.0, 22% PEG 3350 and 2.0 M potassium thiocyanate. These crystals were flash-cooled in a cryo-solution (BTP buffer pH 7.0, 25% v/v PEG 3350, 2.0 M potassium thiocyanate and 5% v/v PEG 400) and exposed to an X-ray beam at 100 K to test diffraction. They diffracted to approximately 3.5 Å on a home-source rotating anode and to 2.9 Å at beam X10SA of the Swiss Light Source (SLS, Villigen, Switzerland).
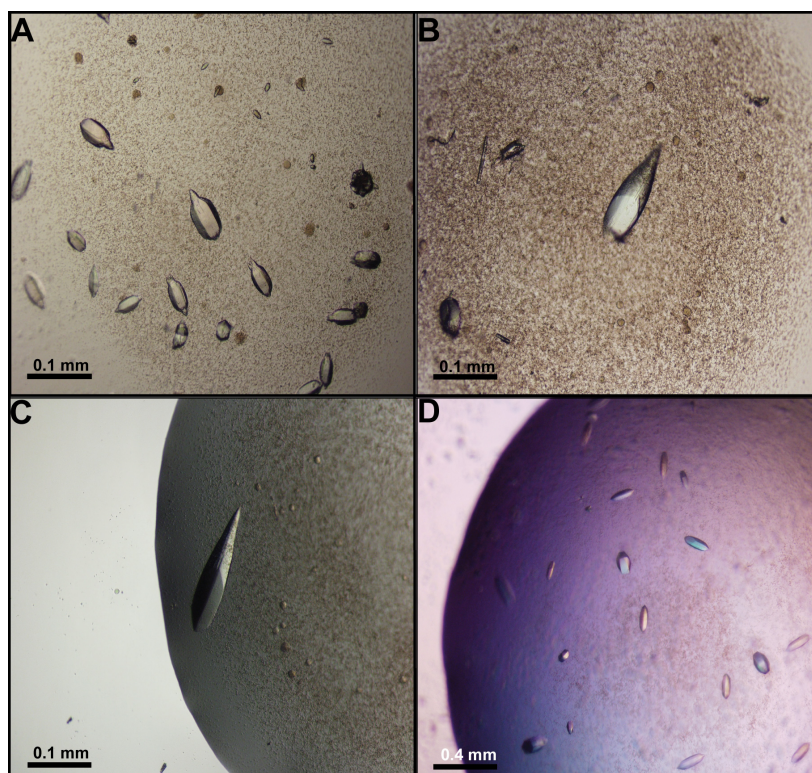

Figure 4.8: Ligand-free PhzE crystals. (A) Crystals obtained from initial screen conditions (B) and (C) Optimized crystals of native PhzE. (D) Selenomethionine labeled PhzE crystal.

Selenomethionine substituted protein was prepared for structure solution by the Multi-wavelength Anomalous Dispersion (MAD) method and crystallized using similar conditions (with additional 5 mM 2-Mercaptoethanol in the protein solution) as the native protein. These crystals diffract to 3.7 Å at SLS X10SA (Figure 4.8 **D**).

4.1.4.2 Crystallization of PhzE in the presence of ligands

Since the ligand-free PhzE crystals cracks immediately after soaking with chorismate, in order to obtain liganded PhzE structure, co-crystallization trials were carried out. Very small crystallines were observed from initial screen of PhzE in presence of ligands (Figure 4.9 **A**) and were re-produced in 2 µL hanging drops (1 µl of reservoir + 1 µl protein) (Figure 4.9 **B**).
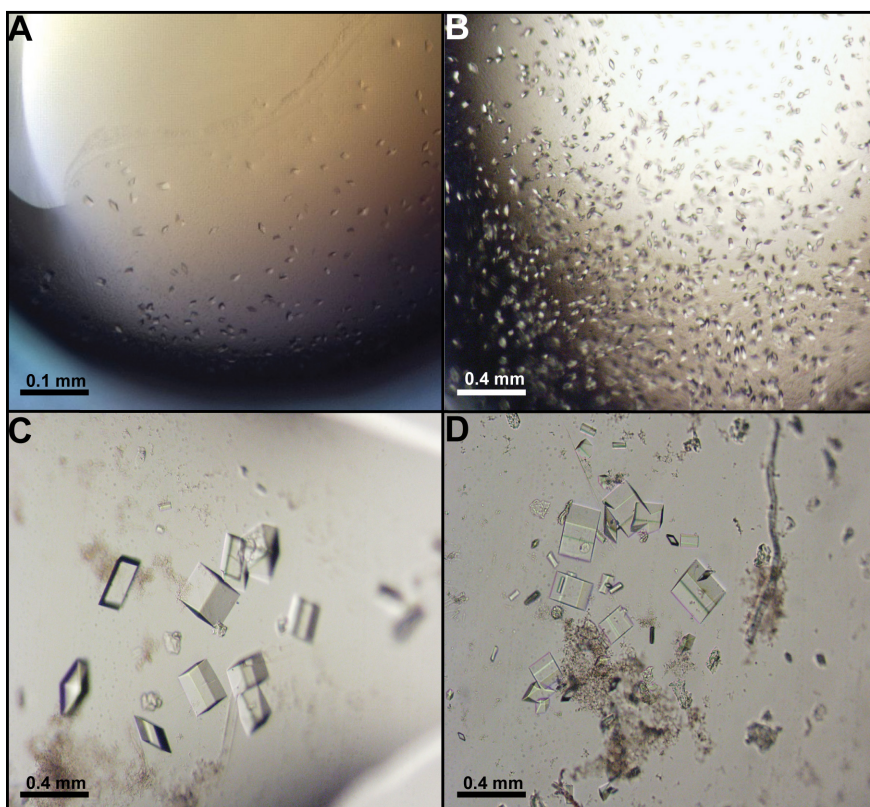

Figure 4.9: Ligand-bound PhzE crystals. (A) Crystallines obtained from initial screening plate. (B) Crystallines reproduced in 2 µL hanging drops. (C) Zn²⁺ containing ligand-bound PhzE crystals after seeding. (D) Zn²⁺-free ligand-bound PhzE crystals after seeding.

Small crystals were then transferred to a freshly prepared drop by macro-seeding technique, and the cubic crystals grew to a full size of 0.4 mm^3 in 3 days (Figure 4.9 **C**). EDTA-treated PhzE crystals were obtained in the same way (Figure 4.9 **D**). Cryo-solution for these PhzE crystals contains 0.1 M HEPES buffer pH 7.1, 0.2 M MgCl$_2$, 15% isopropanol and 20% glycerol. All crystallization experiments were performed at 4°C. The best crystal of ligand-bound PhzE diffracted to approximately 2.7 Å on a home-source rotating anode and to 2.1 Å at beamline X10SA of the Swiss Light Source (SLS, Villigen, Switzerland).

4.1.5 **Data collection statistics**

Diffraction data of ligand-free PhzE were collected at 100 K from 80 non-overlapping 1 s 0.5º oscillation images. The crystal diffracted to 2.9 Å and belongs to space group P6$_2$22 with unit cell dimensions a = b = 172.4 Å, and c = 216.4 Å, α = β= 90º, γ = 120º.

Diffraction data of Zinc-containing ligand-bound PhzE were collected at 100 K from 258 non-overlapping 1 s 0.4º oscillation images. The crystal diffracted to 2.1 Å and belongs to space group P2$_1$2$_1$2 with unit cell dimensions a = 259.7 Å, b = 94.5 Å and c = 53.5 Å, α = β = γ = 90º.

Diffraction data of Zinc-free ligand-bound PhzE crystal were collected at 100 K from 200 non-overlapping 1 s 0.5º oscillation images. The crystal diffracted to 2.6 Å and belongs to space group P2$_1$2$_1$2 with unit cell dimensions a = 258.2 Å, b = 97.8 Å and c = 54.0 Å, α = β = γ = 90º.

MAD data from a Se-methionine-labelled ligand-free PhzE crystal were collected at 100 K at the Se-K-edge from 280 non-overlapping 0.75º oscillations at λ=0.97895Å (Peak), 150 non-overlapping 0.75º oscillations at λ=0.97957Å (Inflection) and 150 non-overlapping 0.75º oscillations at λ=0.97703Å (Remote). All data were collected on beam line X10SA of the Swiss Light Source (SLS, Villigen, Switzerland), using a MAR225 CCD detector.

Data collection statistics for all PhzE crystals are summarized in Table 4.4

Data collection[1]	Se-MAD[2]			Open PhzE	Closed PhzE + Zn^{2+}	Closed PhzE
	Se-peak	Se-inflection	Se-remote			
Space group	P6$_2$22			P6$_2$22	P2$_1$2$_1$2	P2$_1$2$_1$2
Cell dimensions						
a, b, c (Å)	171.5, 171.5, 217.2			172.4, 172.4, 216.4	259.7, 94.5, 53.6	258.2, 97.8, 54.0
α, β, γ (°)	90, 90, 120			90, 90, 120	90, 90, 90	90, 90, 90
Wavelength	0.97895	0.97957	0.97793	1.00002	0.99986	0.978946
f'	-7.51	-13.20	-2.58	—	—	—
f''	10.75	5.12	4.97	—	—	—
Resolution (Å)	20 – 3.6	20 – 4.0	20 – 4.0	20 – 2.9	20 – 2.1	20 – 2.6
(highest shell)	(3.7 – 3.6)	(4.1 – 4.0)	(4.1 – 4.0)	(3.0 – 2.9)	(2.2 – 2.1)	(2.7 – 2.6)
R_{sym}(I)	17.4 (93.7)	14.4 (55.7)	18.5 (83.2)	7.0 (52.5)	6.6 (30.5)	6.9 (41.2)
R_{merge} F	12.8 (52.7)	13.3 (41.7)	18.2 (62.1)	9.3 (48.8)	10.5 (38.4)	8.9 (41.4)
I / σ(I)	15.2 (3.5)	13.6 (4.1)	11.2 (2.8)	20.1 (3.7)	12.4 (3.7)	17.7 (4.5)
Completeness (%)	99.9 (99.9)	99.9 (100)	99.9 (100)	99.6 (99.9)	98.3 (98.3)	99.8 (100)
Redundancy	13.6 (13.6)	7.3 (7.3)	7.3 (7.3)	4.8 (4.9)	2.2 (2.2)	4.0 (4.1)

Table 4.4: Data collection statistics for PhzE crystals. [1]All datasets were collected from a single crystal. [2]Data collections statistics for MAD data refer to unmerged Friedel pairs.

72

4.1.6 **Phasing statistics**

A MAD dataset collected from a single Se-Met PhzE crystal was integrated and scaled with the XDS package and 25 out of 36 expected selenium atoms were successfully located with SHELXD. The coordinates of the Se atoms were then extracted and initial structural phases were calculated with the program SHARP, using experimentally determined f' and f'' values (Figure 4.10). MAD data was used at the resolution range between 20 – 4.0 Å. The initial phase set with an overall FOM of 0.463 was subsequently improved by solvent flattening with the programs SOLOMON and DM.

The solvent content was 62% as calculated with Mattews_coef from the CCP4 suite. Statistics of the first cycle of solvent flattening showed a high correlation value (0.3340) indicating good quality of the initial map. Phasing statistics are summarized in Table 4.5.
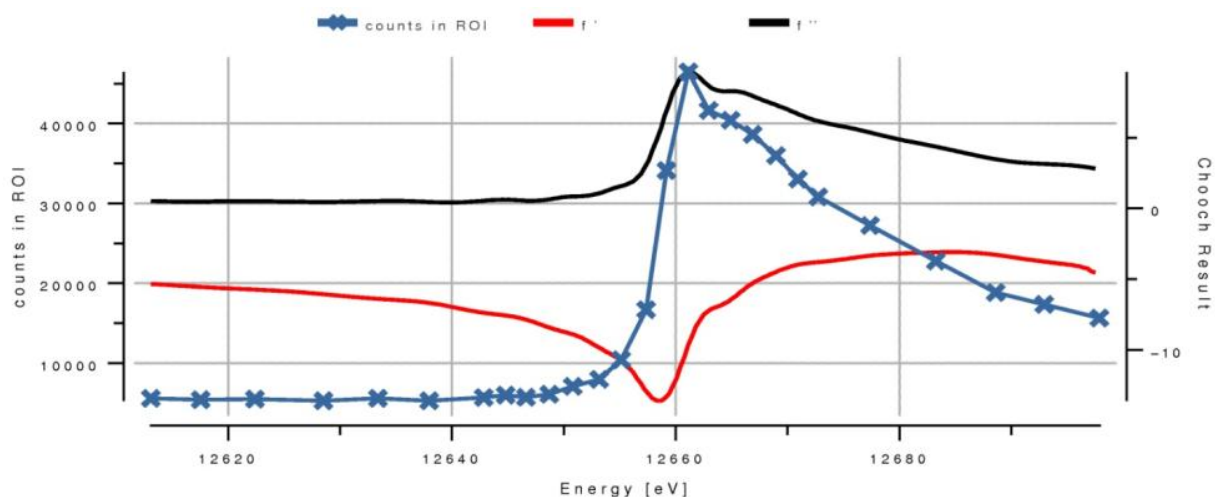

Figure 4.10: Determination of f' and f'' values by X-ray fluorescence scan.

Fluorescence Scan

	E (eV)	f'	f''
Peak	12663.81	-7.5	10.8
Inflection	12657.10	-13.2	5.1
Remote	126798.44	-2.6	5.0

Phasing statistics from the last sharp cycle

FOM	0.4631
Highest resolution shell (4.10-4.00Å)	0.1957
Phasing power (acentrics) anomalous	1.432
Highest resolution shell (4.10-4.00Å)	0.513
Cullis R-factor (acentrics) anomalous	0.720
Highest resolution shell (4.10-4.00Å)	0.958

Statistics after first cycle of solvent flattening

Contrast	0.5660
Overall R-factor	0.4935
R-factor	0.4720
Overall Correlation	0.3340

Table 4.5: Phasing statistics of PhzE

4.1.7 **Model building and Refinement statistics**

The electron density map of apo PhzE after solvent flattening showed well defined and connected structural elements for protein molecules. An initial model of apo PhzE was generated by tracing C_α positions in program O (Jones et al., 1991). After cycles of refinement using REFMAC5 (Murshudov et al., 1997), clear electron density could be observed for a homo-dimer of PhzE with exception of the termini and two flexible

regions (missing residues in the model: chain A: M1 – P5, G429 – L439, R634 – A643; chain B: M1 – N6, T28 – E33, E433 – L439, K637 – A643).
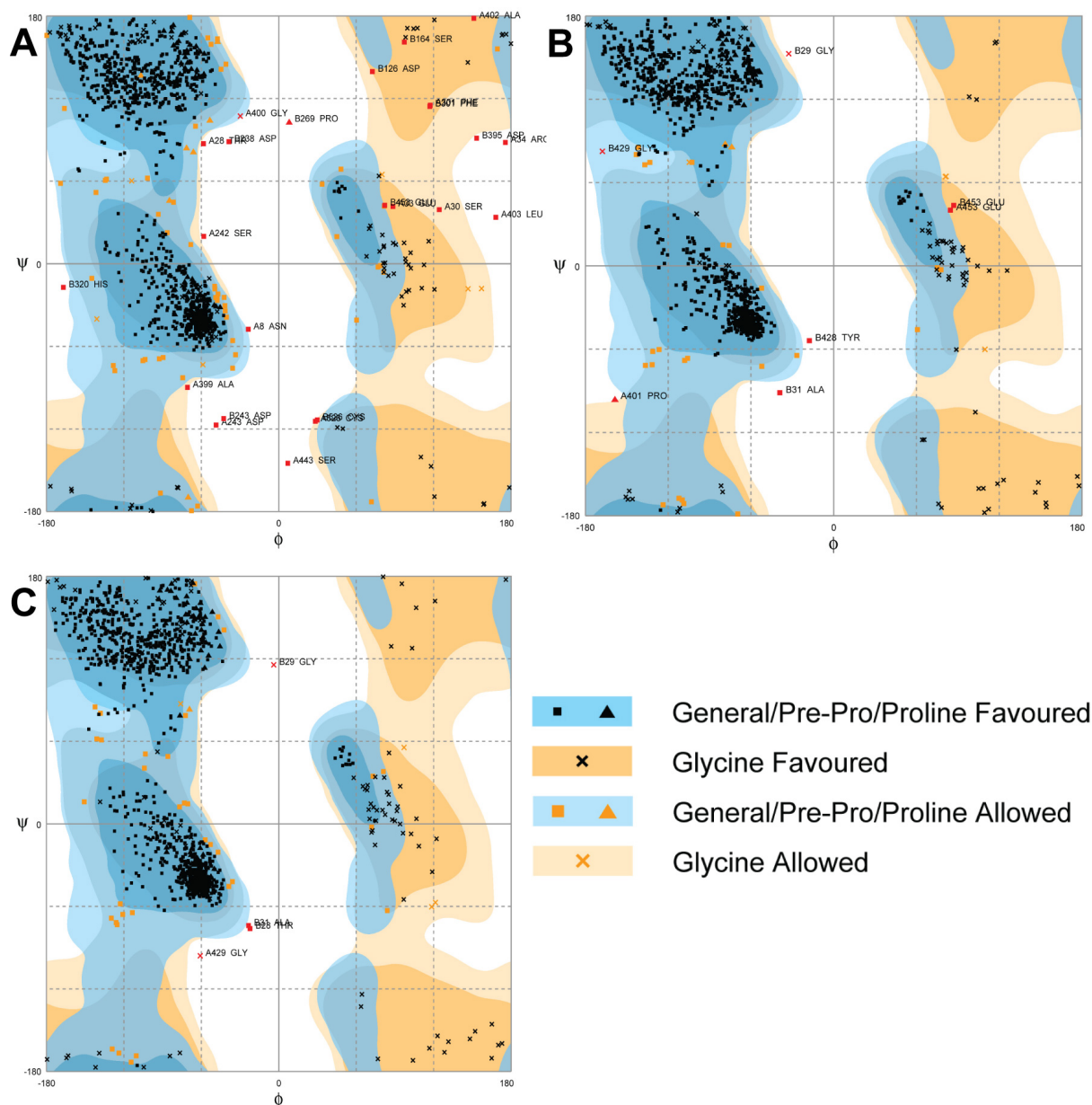

Figure 4.11: Ramachandran plot showing torsion angels of peptide bonds in all three forms of PhzE structures. (A) Ligand-free PhzE. (B) Ligand-bound PhzE with Zn²⁺. (C) Ligand-bound PhzE metal-free. Plots are generated by program RAMPAGE from CCP4 suite.

The improved model was further corrected manually in COOT (Emsley et al., 2010), followed by several rounds of REFMAC refinement. TLS (Translation, Libration and

<u>S</u>crew-rotation displacement) refinement was introduced during REFMAC cycles using the MST domain, the GATase1 domain and the linker region as separate TLS bodies. Water molecules and possible alternative conformations of protein residues were then added manually to the model according to the F_o-F_c map of the electron density. One round of refinement was carried out by phenix refine (Afonine et al., 2005), giving the final model of ligand-free PhzE with R_{work} = 18.7% and R_{free} = 23.1%. The Ramachandran plot (Figure 4.11) showed that 92.6% of all residues lie in the favorite region (5.4% in allowed region).

The structure of liganded PhzE+Zn^{2+} was solved by PHASER (Airlie J McCoy et al. 2007b) using a modified ligand-free PhzE structure (without water, and with separated MST and GATase1 domains) as search model. The structure was refined to R_{work} = 14.6% and R_{free} = 20.0%. The Ramachandran plot (Figure 4.11) showed that 97.0% of all residues lie in favorite region (2.5% in allowed region). Since the structure of liganded PhzE metal-free form was almost identical to the Zn^{2+}-bound form except for the Zn^{2+} ion, it was therefore solved by putting the structure of ligand-bound PhzE+Zn^{2+} directly in REFMAC5 with one round of rigid body refinement. It was finally refined to R_{work} = 16.8% and R_{free} = 23.5%. The Ramachandran plot (Figure 4.10) showed that 96.5% of all residues lie in favorite region (23.2% in allowed region). Refinement statistics of PhzE are shown in Table 4.6.

	Ligand-free PhzE	Ligand-bound PhzE + Zn^{2+}	Ligand-bound PhzE metal-free
Resolution (Å)	20 – 2.9	20 – 2.1	20 – 2.6
(highest shell)	(3.0 – 2.9)	(2.15 – 2.1)	(2.65 – 2.6)
No. reflections	204850 (42476)	324405 (146170)	173817 (43071)
R_{work}	18.7 (25.5)	14.6 (17.4)	16.8 (22.0)
R_{free}	23.1 (32.4)	20.0 (22.4)	23.5 (34.3)
No. atoms			
Protein	9088	10439	9119
Ligand/ion	0	45	42
Water	80	812	332

***B*-factors**			
Protein	31	30	30
Ligand/ion	–	33	34
Water	56	53	28
R.m.s deviations			
Bond lengths (Å)	0.017	0.023	0.016
Bond angles (°)	1.693	1.901	1.576

Table 4.6: Refinement statistics of PhzE structures.

4.2 **Structural analysis of PhzE**

Unless otherwise specified, the ligand-bound PhzE is discussed here using the structure of that partially contaminated with Zn^{2+} bound to the GATase1 active site. Numbering of the secondary structural elements is according to Figure 4.1.

4.2.1 **Overall structure of ligand-free and ligand-bound PhzE**

The first crystal structure of PhzE determinded in this study was in the absence of ligands, which will be referred as the open form. The structure of PhzE in the presence of chorismate, *L*-glutamine and Mg^{2+} was also determined, which will be referred as the closed form. The "open" and "closed" nomenclature is due to the fact that ligand binding of PhzE induces a structure movement of the protein that leads to an open to close motion of the MST active site. Details will be shown and discussed later in this chapter.

As indicated by size-exclusion chromatography, PhzE forms homodimers in solution. Indeed, a 140 kDa PhzE homodimer in the asymmetric unit of both crystal forms was found. In the open form structure, PhzE dimer has a dimension of 120×70×35 Å, and adopts a butterfly-shaped quaternary arrangement in which the two N-terminal chorismate-binding MST domains (residues 1 – 395) interact with each other and lie near the dimer's two fold symmetry axis. The linker regions of both protomers are partially disordered (missing part in chain A: G429 – L439 and chain B: E433 – L439). However, considering the crystal packing and tracing of the linkers, it is only possible that the GATase1 domain (residues 442 – 643) from chain A is in intimate contact with the MST domain from chain B and vice versa, forming two inter-chained MST/GATase1

functional dimer (Figure 4.12 **A** and **B**). The unusual domain intertwinement is retained in the closed form structures (Figure 4.12 **C**), in which the dimer has a dimension of 145×50×35 Å. If the MST domains from chain A of both structures are superposed, a positional change of approx. 60 Å at the outer edge of GATase1 domain from the same chain could be observed (Figure 4.12 **D**). Comparing to the open structure, in the closed form the dimer is more outstretched.
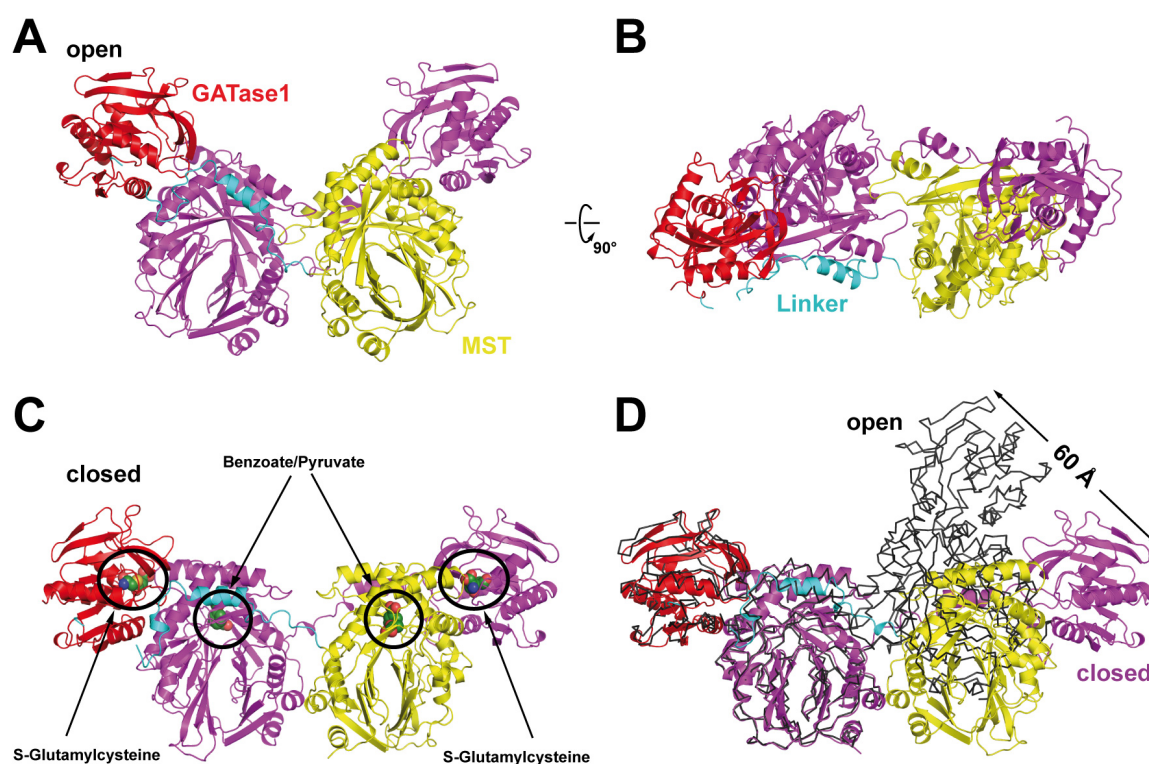

Figure 4.12: Overall structure of PhzE. (A) Ligand-free open form. One chain has been colored in magenta, the other chain is colored according to domains (MST: yellow, GATase1: red, Linker: cyan). (B) Same as in A, top view. (C) Ligand-bound closed form. Chains are colored the same way as in A and B. Circles indicate the position of active sites in the MST and GATase1 domain. Ligands are shown in ball-and-stick representation. (D) Comparison of the open (black lines) and closed structure (cartoons) by superimposition of MST domain from chain A of each crystal form. This figure was prepared with PyMOL (Schrödinger LLC).

4.2.2 **SAXS measurement of the PhzE envelope**

The crystal structures of PhzE revealed an unprecedented quaternary structure showing domain intertwinement between MST and GATase1. In order to exclude the possibility that this unusual domain arrangement is a crystallography artifact, therefore, Small Angle X-ray Scattering (SAXS) experiments were performed to obtain the structural information of PhzE in solution. SAXS is capable of measuring the shape and size of macromolecules and can usually provide structural information at a resolution range between 50 – 250 Å. Due to the fact that PhzE is a marcro-molecule (140 kDa homodimer) which can be presented homogenously in solution, it was possible to obtain enough information from the scattering curve and generated the SAXS-envelope of PhzE dimer. The crystal structures of PhzE were then aligned with the SAXS-envelope with the program Situs 2.5 (Wriggers, 2010) (Fig 4.13). The experimental data was also fit with calculated scattering curve of both crystal structures and χ^2 values were calculated (Formula 4.2). Even though the scattering was recorded in the absence of ligands, the $\chi^2 = 1.5$ for the closed form is much lower comparing to $\chi^2 = 3.2$ for the open form, indicating that the ligand-bound structure of PhzE fits the SAXS envelope much better than the crystal structure obtained in the absence of substrates.

$$\chi^2 = \frac{1}{M} \sum_{i=1}^{M} \left(\frac{I_{exp}(q_i) - cI(q_i)}{\sigma(q_i)} \right)^2$$

Formula 4.2: Fomula to calculate χ^2 value of the SAXS measurement. *$I_{exp}(q)$* is the experimental data and *$I(q)$* is the computed model, *$\sigma(q)$* is the experimental error, *M* is the number of points in the measurement and *c* is the scale factor.
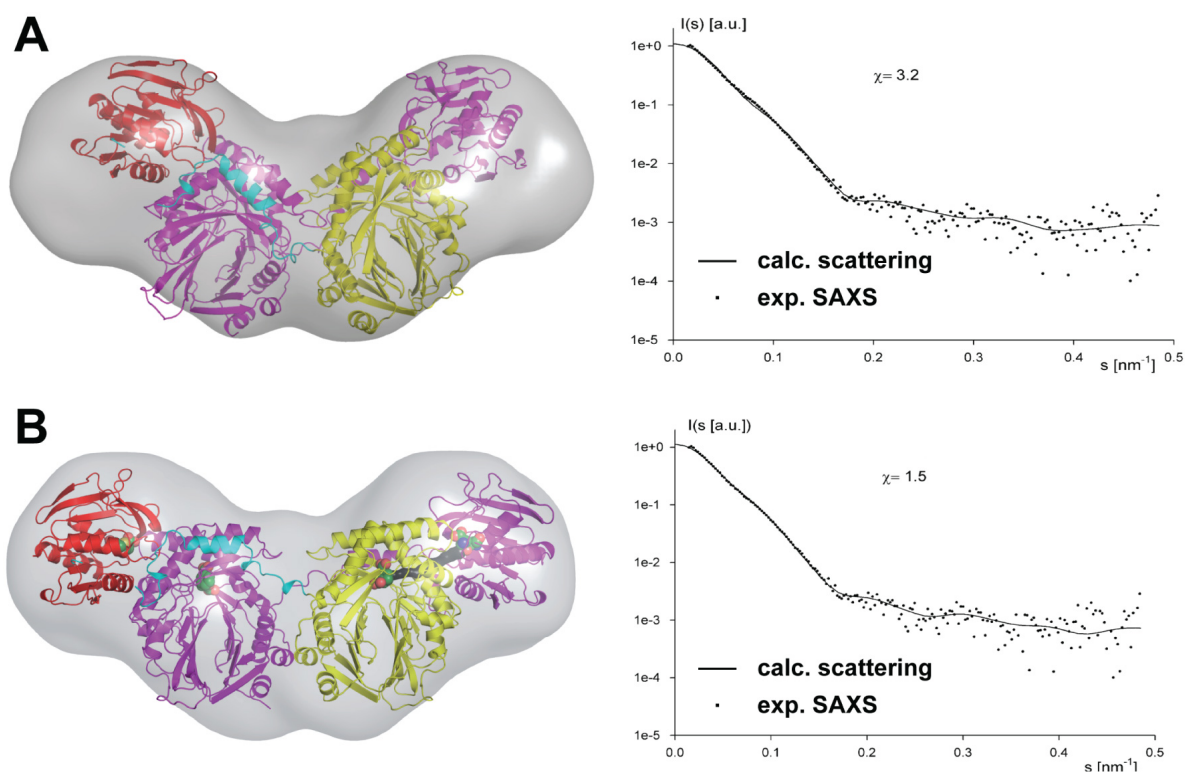
Figure 4.13: Small-angle x-ray scattering of PhzE. The SAXS envelope is shown in grey. (A) Fit of the ligand-free open PhzE. (B) Fit of the ligand-bound closed PhzE. The graphs show experimental and calculated scattering curves. The closed form fits the SAXS data better, even if the scattering was recorded in the absence of ligands.

4.2.3 **Structural comparison of PhzE to AS**

Among all PhzE homologues which contain both an MST (TrpE) and a GATase1 (TrpG) domain, the heterotetrameric AS is the only enzyme class that has been structurally characterized. Therefore, the overall structure of PhzE (closed form) has been compared with three different AS whose structure has been reported.
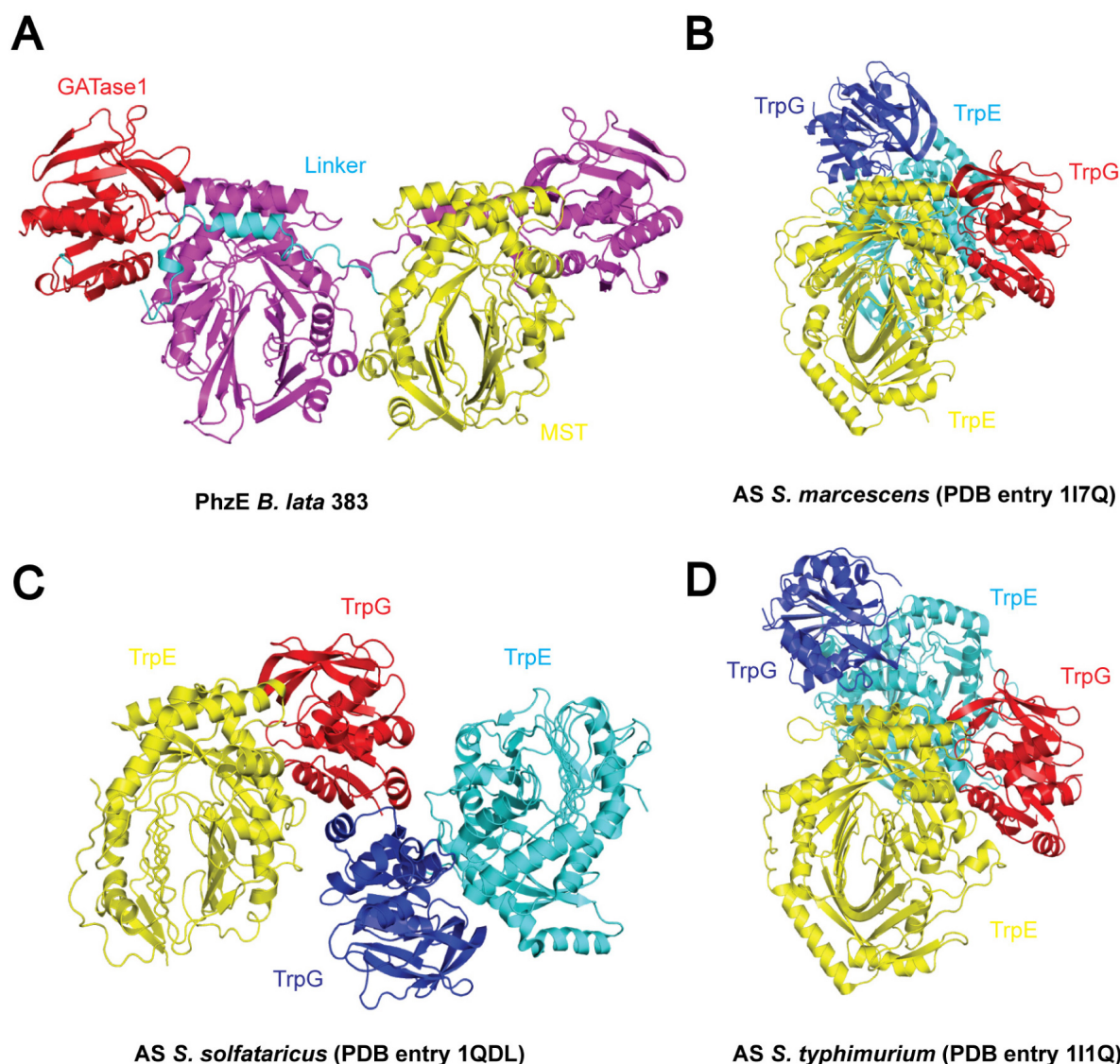
Figure 4.14: Quaternary structure comparison of PhzE with heterotetrameric TrpE/G anthranilate synthases (AS). (A) PhzE from *Burkholderia lata* 383. (B) AS from *Serratia marcescens*. (C) AS from *Sulfolobus solfataricus*. (D) AS from *Salmonella typhimurium*. All structures were superposed with the "yellow" MST domain of PhzE.

The MST (TrpE) domains (subunits) from chain A of all four structures are superposed (Figure 4.14). This shows that the relative position of the MST/GATase1 functional pair in PhzE is similar to that of heterotetrameric TrpE/G AS. In the case of AS from *Serratia marcescens* (PDB ID: 1I7Q) (Spraggon et al., 2001) and from *Salmonella typhimurium* (PDB ID: 1I1Q) (Morollo & Eck, 2001), the two functional dimers are held together through interactions of residues from both TrpE and TrpG

subunits, for AS from *Sulfolobus solfataricus* (PDB ID: 1QDL) (Knochel et al., 1999), the interaction exists only between TrpG subunits. Interestingly, PhzE clearly adopts a different quaternary structure to all three AS, where the functional MST/GATase1 pairs are held together only through the two MST domain interfaces. However, this finding itself doesn't explain the functional diversity between PhzE and AS because the quaternary structures among AS are also different.

4.2.4 **The MST domain of PhzE**

4.2.4.1 *Overall structure of the MST domain*

One hallmark of the MST domain is its distinguished tertiary fold where a large orthogonal β-sandwich structure is formed in the center of the domain, covered by α-helices on the surface (Figure 4.15 **A** and **B**). Secondary structure elements (α1 – α9, β1 – β22) have been numbered starting from N-terminal of the protein (Figure 4.1). In both the open (Figure 4.15 **A**) and closed (Figure 4.15 **B**) conformation, the secondary structural elements of the MST domain are well conserved. However, in the open form, B-Factor value of residues located on α5, α9, β14, β15, β16, as well as in loop regions extended from both termini of α5, are significantly higher than those in the closed form (Figure 4.15 **C** and **D**), indicating that these residues may be involved in the open-form to closed-form movement of PhzE. Indeed, such a movement is confirmed when the open-form and the closed-form structures are compared. It is also proved that the movement is induced by ligand binding of the MST domain (see chapter 4.2.7).
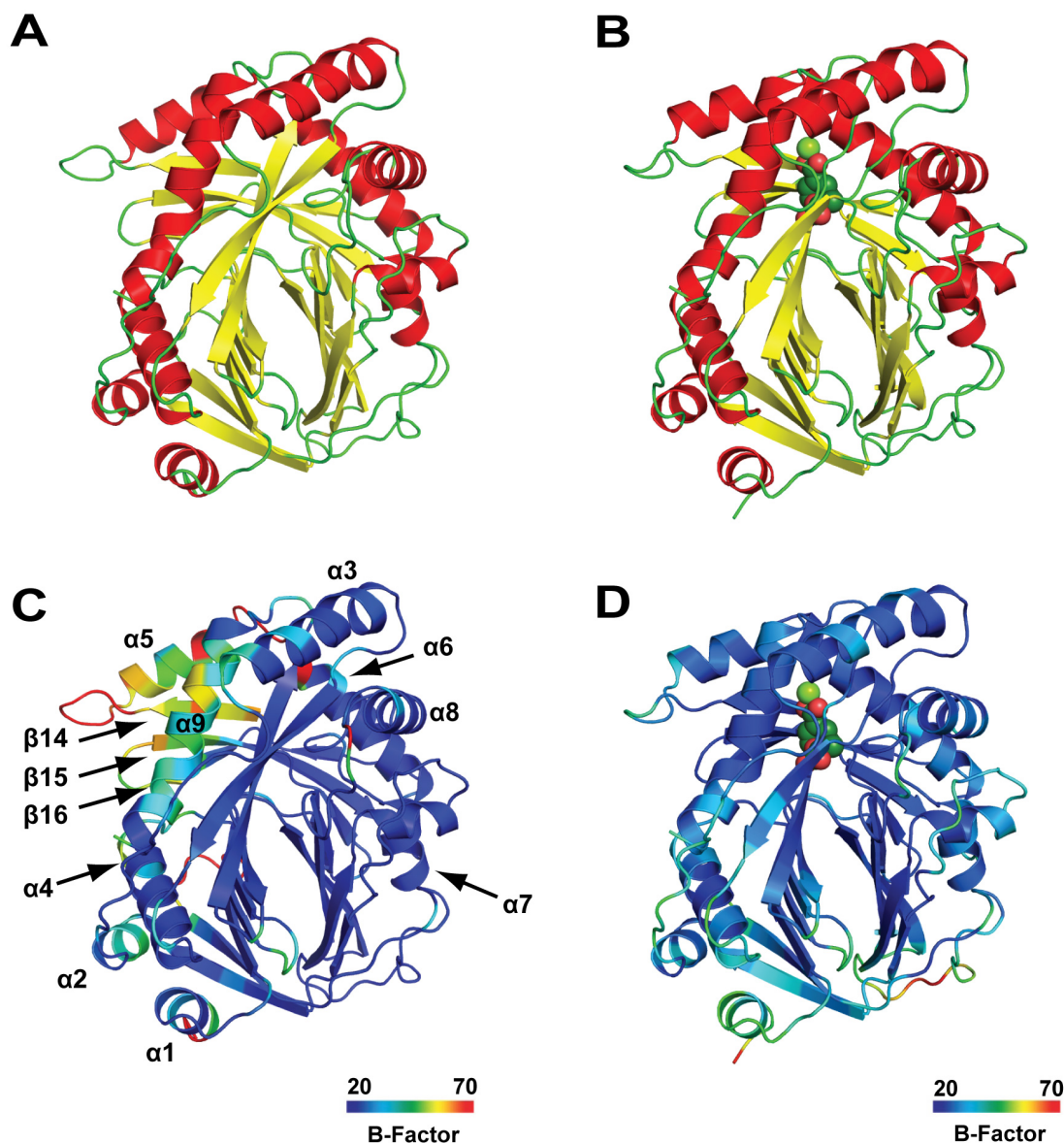
Figure 4.15: Structure of the MST domain of PhzE. (A) Cartoon representation of open MST, colored by secondary structural elements. Helices are colored in red, strands are in yellow. (B) Cartoon representation of closed MST. Bound ligands were colored by elements. Secondary structures are colored in the same way as in A. (C) Cartoon representation of open MST showing Cα B-Factor. (D) Cartoon representation of closed MST showing Cα B-Factor. Bound ligands were colored by elements.

4.2.4.2 Structural comparison of the MST to its homolgues

A search with the DALI server (Holm & Rosenstrom, 2010) identifies proteins (subunits) that share similar fold with the MST domain of PhzE. One protein from each

sub-family of MST enzymes is listed in Table 4.7. Comparison of the overall structure between the MST domain of PhzE and other MST enzymes is shown in Figure 4.16.

PDB ID/Chain ID	RMSD	Sequence identity (%)	Description
1QDL/A	2.8	19	Anthranilate Synthase (AS)
3HWO/A	2.9	19	Isochorismate Synthase (ICS)
2FN0/A	2.5	19	Salicylate Synthase (SS)
3H9M/A	2.5	15	Para-aminobenzoate (PAB) Synthase
1K0E/B	2.7	18	Aminodeoxychorismate (ADC) Synthase

Table 4.7: Structural comparison of MST enzymes to MST domain of PhzE. Only one protein from each sub-family is listed here.

Although the MST domain has a low overall sequence identity to its homologues (< 20%), the tertiary structure of related enzymes is well conserved. Interestingly, it is found that PAB synthase and ADC synthase, two enzymes from the folate biosynthesis pathway, possess similar sequence identity and tertiary structure to those of the MST enzymes. These two enzymes are both two component enzymes that utilize one subunit to generate NH_3 and perform nucleophilic attack at C4 position of the chorismate ring at the other subunit (MST subunit), with or without concominant cleavage of the pyruvate. Since the functional and structural relationships between these two enzymes are surprisingly similar to that of the other two pairs of enzymes in the family (AS/ADICS and SS/ICS), it is suggested that the previously described MST enzyme family should be extended to include PAB synthase and ADC synthase.
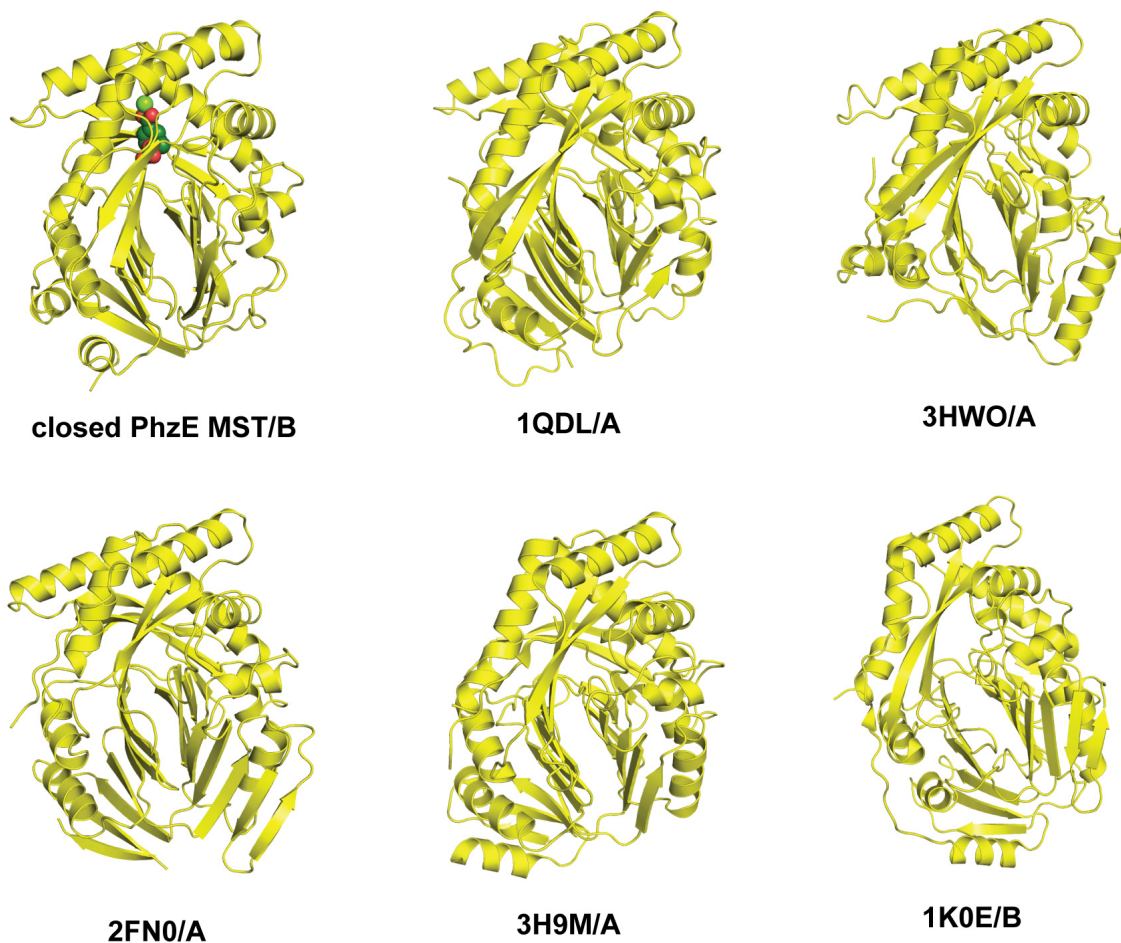
Figure 4.16: Structural comparison of MST domain from PhzE and other MST enzymes. Ligands are colored by element and shown as sphere and stick. Model alignment and figure preparation were both done with Pymol. The enzymes are described in Table 4.7.

4.2.4.3 Active center of MST domain

Although the ligand-bound closed conformation of PhzE was crystallized in the presence of chorismic acid, neither the substrate chorismic acid nor the product ADIC were observed in the structure. Instead, benzoate and pyruvic acid were found in the active center of MST domain (Figure 4.17 **A** and **B**). The omit electron density shows clearly that the carbonyl oxygen atom of pyruvate is 3 Å away from C3 of benzoate, indicating that the C-C bond is broken (Figure 4.17 **B**). The same phenomenon was reported also in the ligand-bound crystal structure of AS from *Serratia marcescens*

(PDB entry 1I7Q) (Spraggon et al., 2001), in which benzoate and pyruvic acid sit in almost the same position as in the closed PhzE structure.
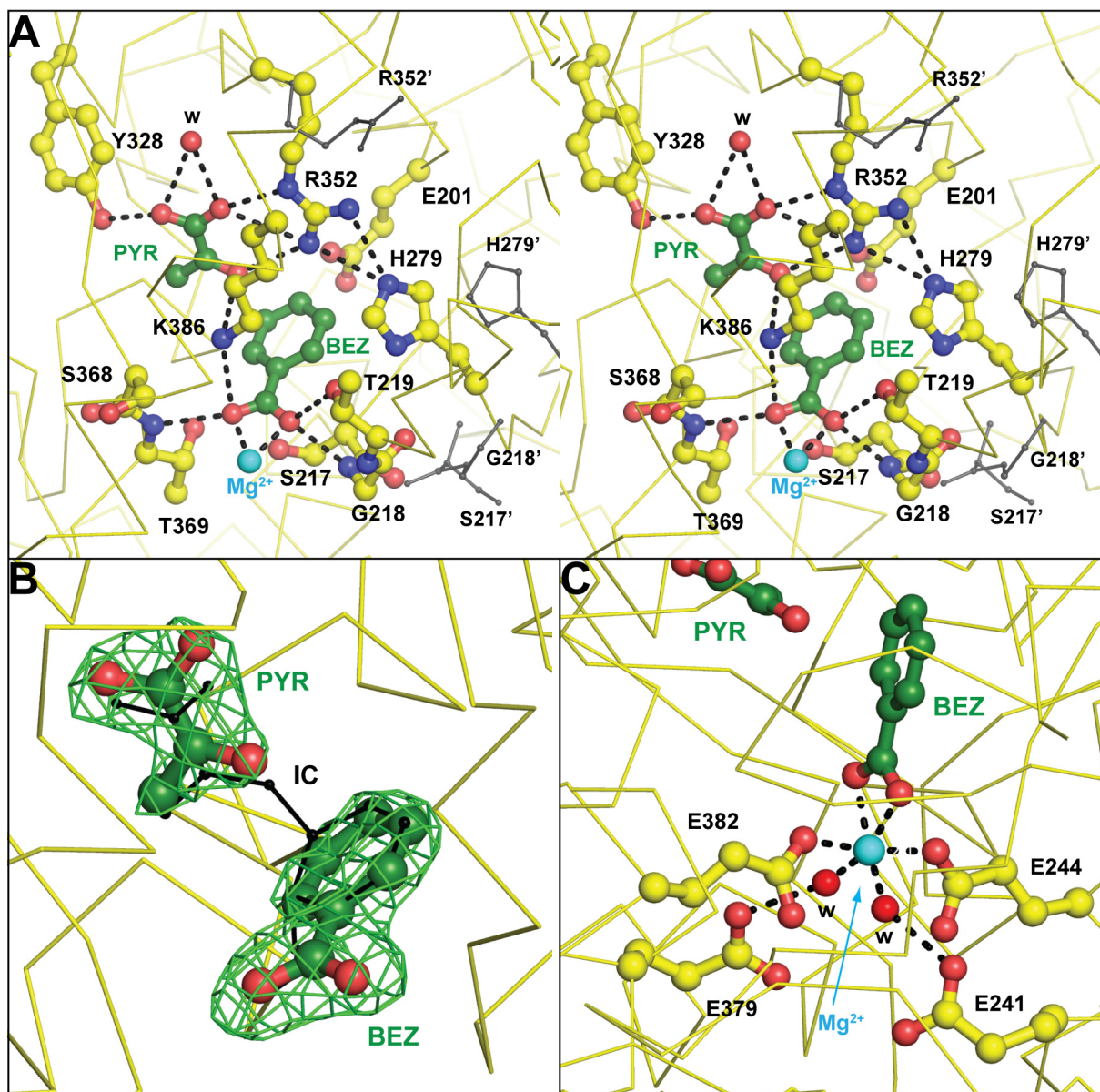

Figure 4.17: Details of the active site of MST domain. (A) Active site of the MST domain. Residues that make hydrogen bonds with pyruvate (PYR) or benzoate (BEZ) are shown in ball-and-stick representation. Mg^{2+} is shown in cyan. The position of important moving residues in the open form is shown in black. (B) Superposition of benzoate and pyruvate to isochorismate (IC, black) at the active center of MST domain. The omit electron density of benzoate and pyruvate is displayed at 3.5 σ. Isochorismate-coordinates are extracted from PDB entry 3WHO. (C) Residues and molecules involved in Mg^{2+} coordination.

Although benzoate and pyruvate trapped at the active center are not natural ligands of the enzyme, structural alignment of PhzE to a recently published structure of isochorismate synthase EntC (PDB entry 3HWO) showed that the position and interaction of benzoate and pyruvic acid are very similar to those of isochorismate in the EntC structure (Sridharan et al., 2010) (Figure 4.17 **B**), indicating that the benzoate and pyruvate ligands can serve as a good template for investigating enzyme/substrate interaction.

In the closed-form structure, R352 adopts a different rotamer conformation compared to that in the open-form structure in order to bind pyruvate. H279 also moves nearly 4 Å with respect to the open form to interact with and stabilize the new rotamer of R352. Pyruvate also makes hydrogen bonds with Y328 as well as with a water molecule close by (Figure 4.17 **A**). The carboxylate group of benzoate is involved in hydrogen bonds with a number of residues near by, including the amide bonds between S217 and G218, S368 and T369 and the side chains of S217, T219 and K386. It also occupies two of the octahedral coordination positions of the Mg^{2+} cation which was found at the active center. Other residues/molecules involved in coordination of Mg^{2+} including E244, E382 and two water molecules, which are held in position through hydrogen bonds with E241 and E379, respectively (Figure 4.17 **C**). Protonation of the leaving hydroxyl group at C4 of chorismic acid is achieved by the side chain of E201, the importance of this residue is indicated by the fact that E201Q was inactive (Table 4.10). The nucleophilic attack by ammonia at C2 is possibly base assisted. The carbonyl groups of I216 and T304, together with side chains of S217 and T369, are well positioned to act as hydrogen bond acceptors aiding the deprotonation of NH_3. Given the information provided, the mechanism of ADIC generation in the MST active center is proposed (Figure 4.18).
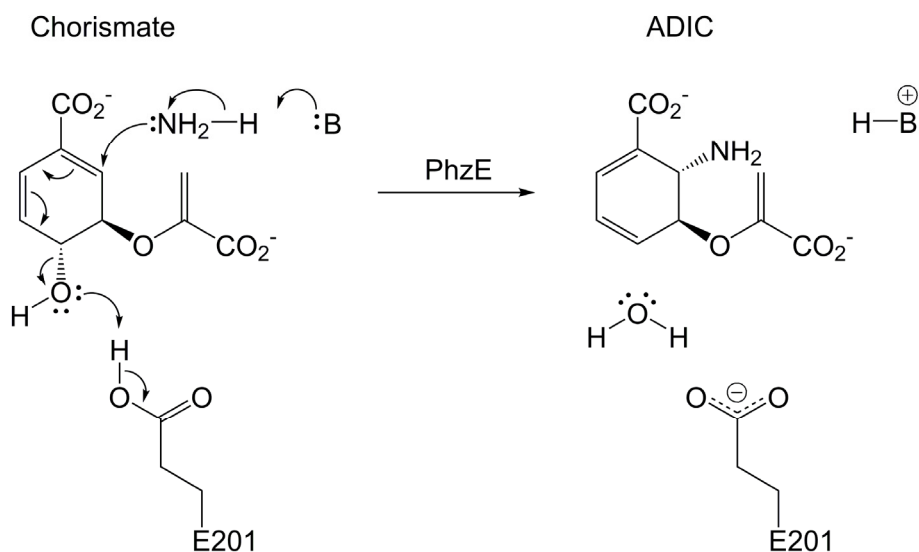
Figure 4.18: Proposed mechanism for the generation of ADIC in the active site of the MST domain. The identity of the base B required for deprotonation of the NH$_3$ nucleophile is unclear.

The structures of PhzE alone do not provide enough information why benzoate and pyruvate but not chorismate or ADIC were observed in the active site of closed MST. However, since neither benzoate nor pyruvate were present in the buffer used for purification or crystallization, and because the closed form PhzE crystals only grows if the enzyme was incubated with all its substrates (chorismate, L-glutamine and Mg^{2+}) for at least 1 hour prior to the crystallization trials, it is believed that the presence of benzoate and pyruvate is the result of the breakdown of its product ADIC. This speculation is also backed up by the fact that it was impossible to crystallize either the native enzyme with exogenous benzoate and pyruvate or any of the inactive mutants in the closed conformation. ADIC is a relatively unstable compound with a half-life of 34 hours in Tris buffer at pH 8.0 and 22 ºC (Morollo & Bauerle, 1993), which is similar to the time scale of crystallization. It is therefore likely that the breakdown of ADIC occurs in the active site. This hypothesis is also supported by observations of the breakdown of IC via a sigmatropic rearrangements resulting in the elimination of pyruvate (DeClue et al., 2005) (Zwahlen et al., 2007). In this reaction, the enzyme is believed to play a role in orienting the methylene group of the enolpyruvyl moiety towards the C2 hydrogen atom of IC, facilitating the transfer of the C2 hydrogen to C9 of the pyruvyl moeity and the simultaneous cleavage of the C–O bond. Based on this, a pericyclic pyruvate

elimination mechanism for ADIC was proposed (DeClue et al., 2005) (Figure 4.19). The generation of benzoate in this case requires reduction. It is unknown if the reducing agent (β-mercaptoethanol) in the crystallization condition is the cause of the reduction. In addition, X-ray radiation damage might be another possible explanation for ADIC breakdown.
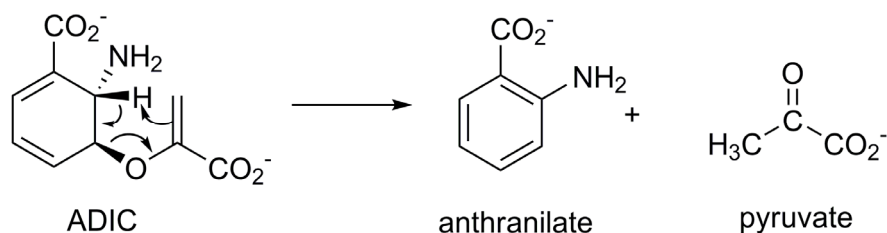

Figure 4.19: pericyclic pyvuvate elimination mechanism in AS.

4.2.5 **GATase1 domain of PhzE**

4.2.5.1 *Overall structure of GATase1 domain*

The GATase1 domain of PhzE belongs to the family of class I GATases, one of the two classes of typical glutamine amidotransferases which use histidine and glutamate residues for the activation of the cysteine thiol group. Due to the catalytic triad (H, E and C), the class I GATases are also called triad GATases (Mouilleron & Golinelli-Pimpaneau, 2007). Similar to other class I GATases, the GATase1 domain of PhzE shares a common open α/β structure fold (Figure 4.20 **A** and **B**). Superimposition of the PhzE GATase1 domain in its open and closed forms reveals that only small structural changes (backbone RMSD = 0.595 Å) accompany ligand binding. The active site is located close to the surface and is solvent accessible in the open structure, but it becomes covered through interactions with the MST domain upon a ligand-induced rigid body movement (Figure 4.27).

4.2.5.2 *Structural comparison of the GATase1 domains (subunits/enzymes)*

The coordinates of the GATase1 domain from chain B of the ligand-bound closed PhzE (residue A440 – K636) were submitted to the DALI server (Holm & Rosenstrom, 2010) and the search results indicated that it is most similar to the TrpG subunit of AS,

which is one of the best studied type 1 glutamine amidotransferase (also known as TrpG-type). The GATase1 domain of PhzE and TrpG subunits of AS share exactly the same glutamine hydrolyzing residues despite their poor sequence identity (Table 4.8). A comparison of the structures of GATase1 and TrpGs is shown in Figure 4.21.

PDB ID/Chain ID	RMSD	Sequence identity (%)	Description
1I7Q/D	2.0	26	Anthranilate Synthase
1QDL/B	2.4	25	Anthranilate Synthase
1I1Q/B	2.2	29	Anthranilate Synthase

Table 4.8: Structural comparison of TrpG subunit from AS to the GATase1 domain of PhzE.
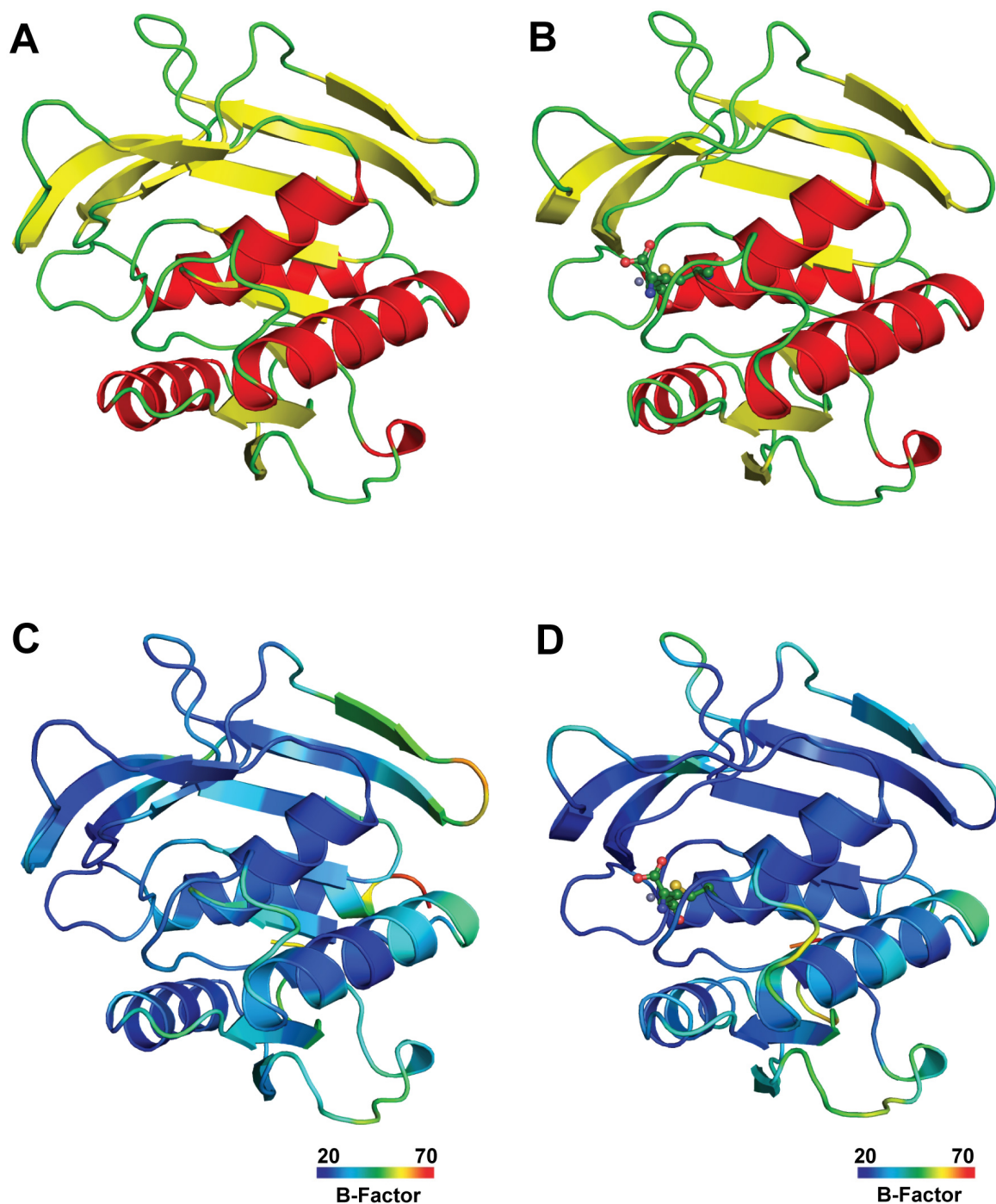
Figure 4.20: Overall structure of the GATase1 domain. (A) Cartoon representation of GATase1 in the open form, colored by secondary structural elements. Helices are colored in red, strands are in yellow. (B) Cartoon representation of closed GATase1. Bound ligands were colored by elements. Secondary structures are colored in the same way as in A. (C) Cartoon representation of open GATase1 showing Cα B-Factor. (D) Cartoon representation of closed GATase1 showing Cα B-Factor. Bound ligands were colored by elements.
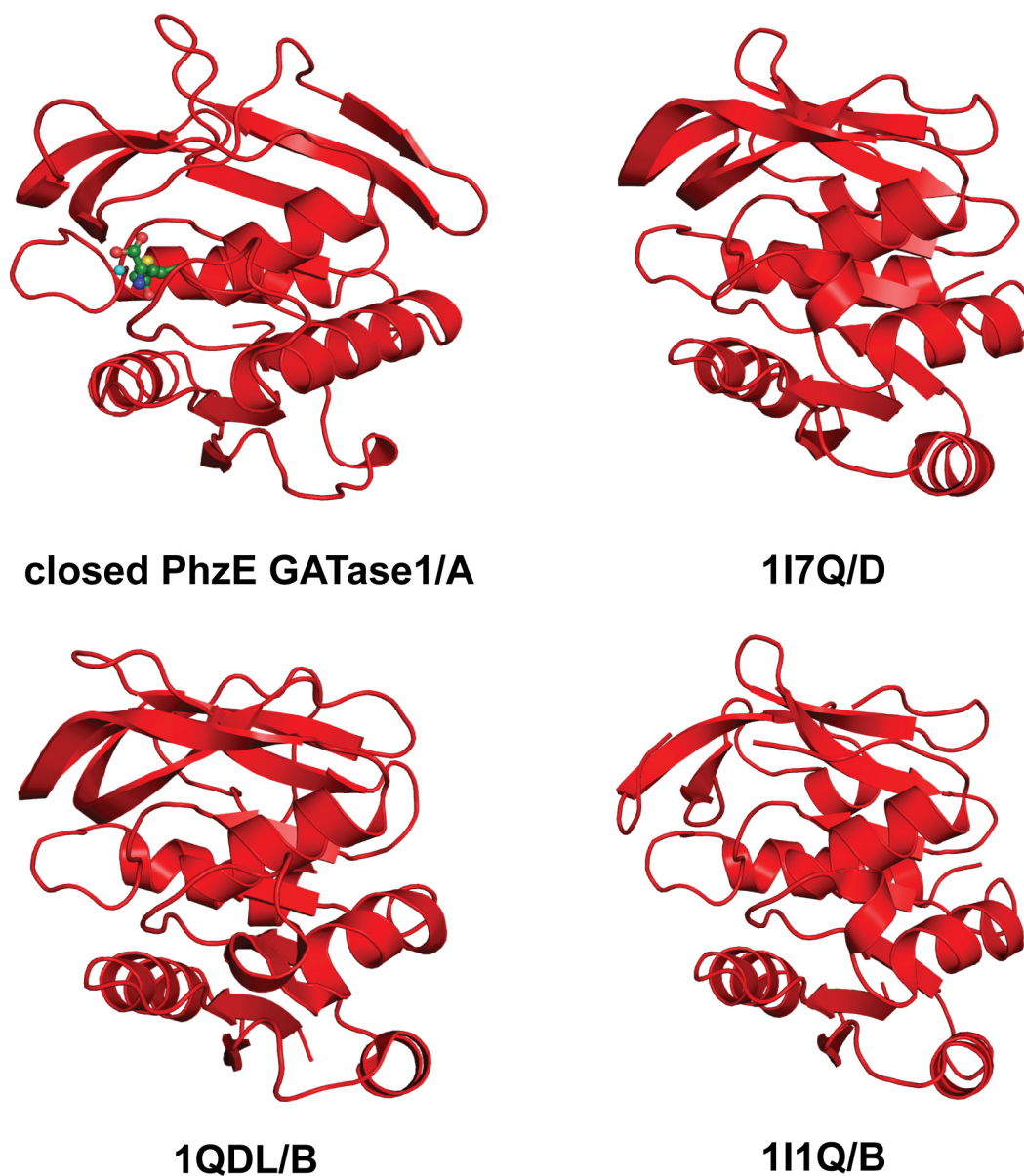
Figure 4.21: Structural comparison of GATase1 domain from PhzE and TrpG. Ligands are colored by element and are shown as balle and stick. Model alignment and figure preparation were both done with Pymol.

4.2.5.3 Active center of the GATase1 domain

In the active center of closed form GATase1, the catalytic C526 is covalently modified by a glutamyl moiety, which is in polar contacts with the side chain of Q530 and Q548, as well as with the backbone amides of G493, T568 and Y569 and the

backbone carbonyl group of G495 (Figure 4.22 **A**). It is believed that the liganded active center observed represents an intermediate state of the glutaminase catalytic cycle, while the ammonia has already been released from glutamine and the C526 is still covalently modified (Figure 4.22 **B**).
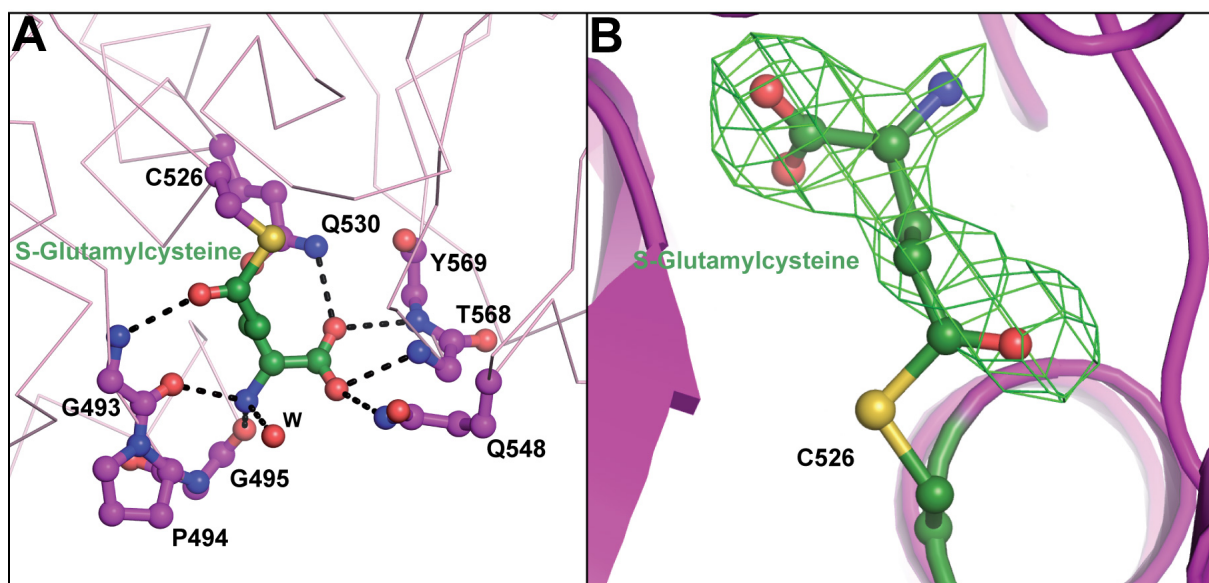

Figure 4.22: Details of active center of the GATase1 domain of PhzE. (A) Covalent modification of C526 and coordinations of the glutamyl moiety. (B) |F$_O$F$_C$| difference electron density of covalently bound glutamyl moiety at 2 σ.

Interestingly, in the data set with highest resolution (2.1 Å) collected for the closed form crystal in this study, additional electron density including anomalous diffraction in the vicinity near C526, H611 and the backbone carbonyl group of W566 was observed. The electron density could be removed by treating PhzE with EDTA prior to crystallization. According to its tetrahedral coordination and evidence revealed by x-ray fluorescence experiment, the anomalous density was identified as Zn^{2+} cations (Figure 4.23 **A**). However, Zn^{2+} has not been intentionally introduced during the entire experiments from protein preparation to crystallization. One of the conceivable sources of Zn^{2+} could be the $NiSO_4$ used during the first step of protein purification, however, this chemical contains only less than 0.002% zinc as impurity, giving an overall Zn^{2+} concentration less than 2 μM in the solution used to charge the Ni-NTA column. This concentration is much lower than the cytosolic concentration of Zn^{2+} in *E. coli*, which is in the millimolar range, but cytosolic zinc almost all bind to zinc-binding proteins.

Therefore, the origin of the bound Zn^{2+} remains unveiled in this study. Therefore, it is conceivable that PhzE may have to compete with other zinc binders in the cells in order to acquire Zn^{2+} as observed in the crystal structure here. It is indeed confirmed that PhzE is inhibited by divalent cations including Zn^{2+}, Mn^{2+} and Ni^{2+} (Figure 4.33), which contributes to the argument that on top of transcriptional regulation by quorum sensing, divalent transitional cations may add a second layer in the overall control of phenazine biosynthesis.
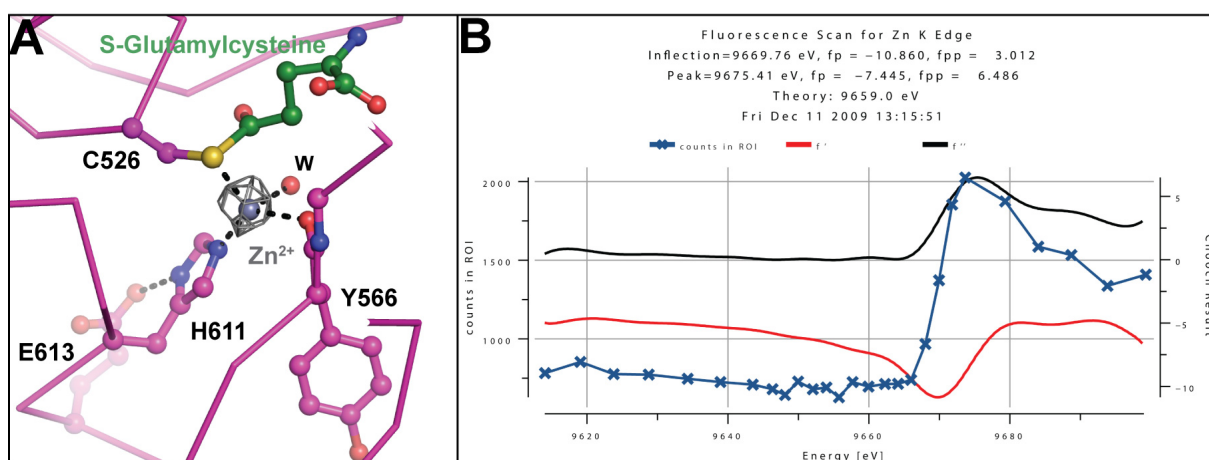

Figure 4.23: Zn^{2+} at the active center of GATase1 domain. (A) Tetrahedrally coordinated Zn^{2+} (grey sphere) at the GATase1 active center. Anomalous difference electron density of data collected at 12.4 keV at 5σ is displayed as a grey mesh. (B) X-ray fluorescence spectrum of a PhzE crystal in the closed conformation at the zinc K-edge. The protein has not been treated with EDTA prior to crystallization.

Because the zinc-free and zinc-bound structures of closed PhzE are identical except for the Zn^{2+}, the discussion in this study is based on the higher resolution Zn^{2+} containing data.

4.2.6 **The linker region**

The linker comprises 46 amino acids between P396 and E441, part of which is glycine and proline-rich but contains also helical secondary structures. Due to the high flexibility, the linker was not completely traceable. In the open form, G429 – L439 in chain A and E433 – L439 in chain B are missing, whereas in the closed form, Q432 – L439 are missing in both chains. The linker spans over a distance of 60 Å in order to connect the MST and GATase1 domains. Structure comparison of open and closed

linker shows that the backbone conformation from P396 to R414 differs most strongly (Figure 4.24). It is believed that the flexible linker plays important roles in assisting the large structural changes between the two crystal forms of PhzE, which will be discussed below (see chapter 4.2.7).
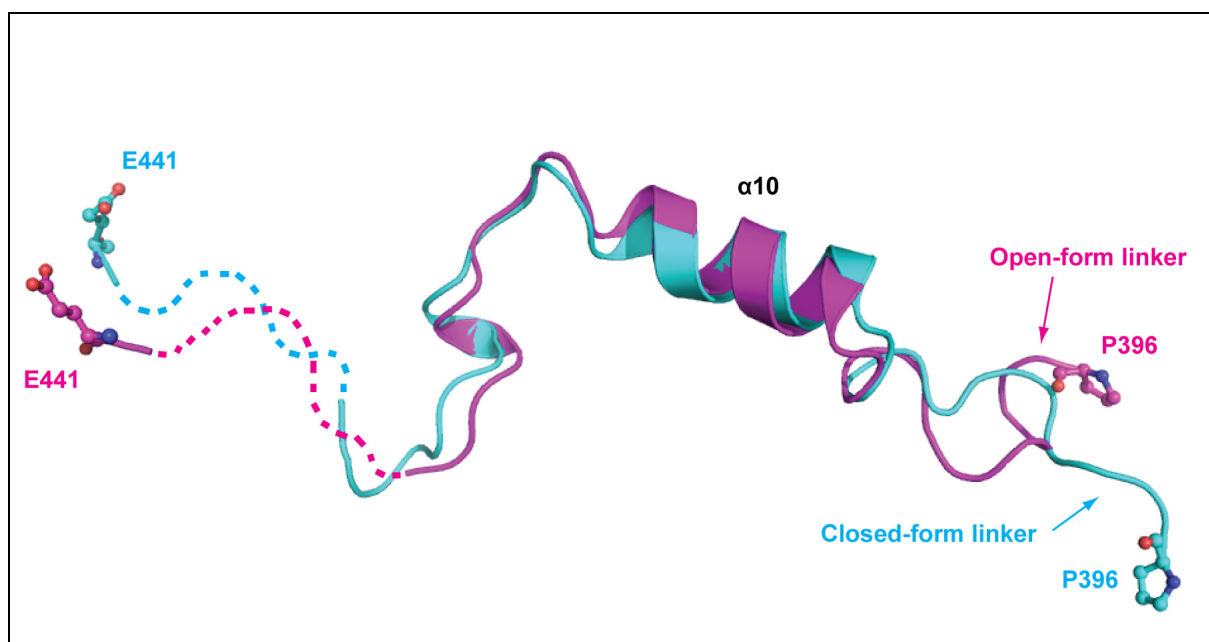

Figure 4.24: The linker region. The cartoon colored in magenta is the linker from open PhzE chain A, in cyan is from closed PhzE chain A. Dashed lines indicate the flexible part of the linker that is untraceable (G429 – L439 in case for open conformation and Q432 – L439 for closed form). Note that the loop region starting with P396 near N-terminal part shows the largest structural difference between two structure forms. The numbering of α-helix is according to figure 4.1.

4.2.7 Ligand-induced structural changes of PhzE

The two crystal structures of PhzE in its ligand-free open form and its ligand-bound closed form provide insight into the ligand-induced conformational changes that accompany the enzyme's catalytic cycle.

Concerning substrate binding, it is believed that Mg^{2+} and chorismic acid bind first, enabling the GATase1 domain to capture glutamine since neither glutamine nor Mg^{2+} was observed in the open form structure despite being present in the crystallization solution and chorismic acid alone did not show any affinity towards PhzE in ITC

experiments unless Mg^{2+} was present (see chapter 4.3.1). This is an indication that the large structure changes are mainly induced by Mg^{2+} and chorismic acid binding.

In open form PhzE, the repulsion between negatively charged helix α9 (E379 and E382) and highly acidic stretch D238 to E244 stabilizes the domain structure as shown by electrostatic surface around the entry of the active site calculated with APBS (Figure 4.25 **A**) (Baker et al., 2001). Binding of Mg^{2+} and chorismic acid is believed to trigger the structural changes, and the octahedral coordination sphere of Mg^{2+} involves a bidentate interaction with the carboxylate group of benzoate, as well as the side chains of E244 and E382, and two water molecules that are held in position by E241 and E379. With respect to the open structure, these acidic residues approach each other, leading to a clamp-closing like movement of α9 on one side and a large structural element from P215 to L249 on the other side (Figure 4.25 **B**). Of these amino acids, D238 – E241 also change their backbone conformation from coil to α-helical, which is the most significant structural rearrangement within MST domain. This movement is likely to further stabilize the closed form conformation and is also observed in AS (PDB entries 1I7S and 1I7Q) (Spraggon et al., 2001) and IS (PDB entries 3BZM and 3HWO) (Parsons et al., 2008) (Sridharan et al. 2010) to a lesser extent. It is thus conceivable that neutralization of the negative charges in the active site by binding of Mg^{2+} partially explains the requirement for Mg^{2+} of MST enzymes. In addition, the conserved K240, which is pointing outward in the open structure, reorients towards the active center in corresponding to the glutamic acids and interacts with the side chains of E241, E244 and E379, assisting in stabilizing the clamp in the closed conformation (Figure 4.25).
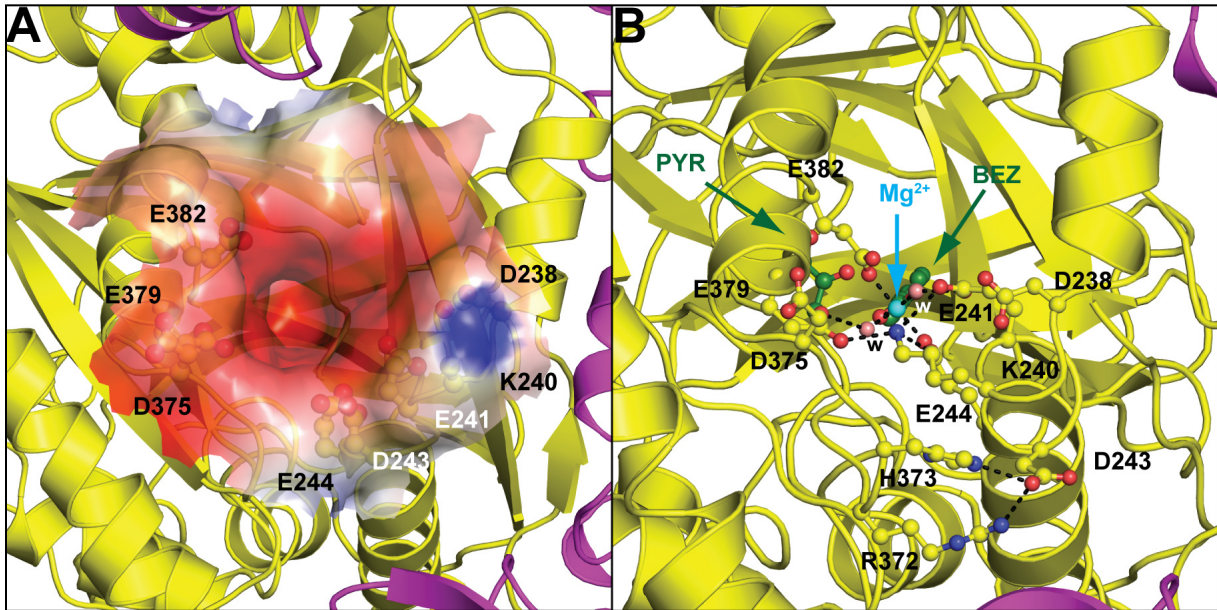
Figure 4.25: Positioning of acid residues near the active center of MST domain. (A) Acidic residues and electrostatic surface around the entry to the active site of MST domain in the open form. (B) Interaction between ligands and surrounding acidic residues in the active center of closed MST domain.

As one jaw of the clamp, P215 – L249 contains a strand β14 from one of the β-sheets and therefore pulls the two strands (V266 – I284) β15 and β16 and residues A393 – G419 from the linker together with it. Anchoring of P215 – L249 also involves polar interactions between the carboxylate group of benzoate and the backbone amide between S217 and G218, leading to a peptide flip in comparing to the open conformation (Figure 4.26). The arrangement of these structural elements (P215 – L249 and V266 – I284) in the closed form is incompatible with the open form. In the open form the more compressed MST/MST interface would lead to a steric conflict to the R211 – P227 from the second chain. When the open and closed structures are superposed by one MST domain from the same chain, there is an approx. 50º rotation between the two MST/GATase1 functional units (Figure 4.27 **A**). However, these large structural rearrangements are almost completely compensated though backbone changes of linker residues D395 – R414 within the functional units, allowing much less changes in the functional MST/GATase1 dimer interface with respect to the MST/MST reorientation.
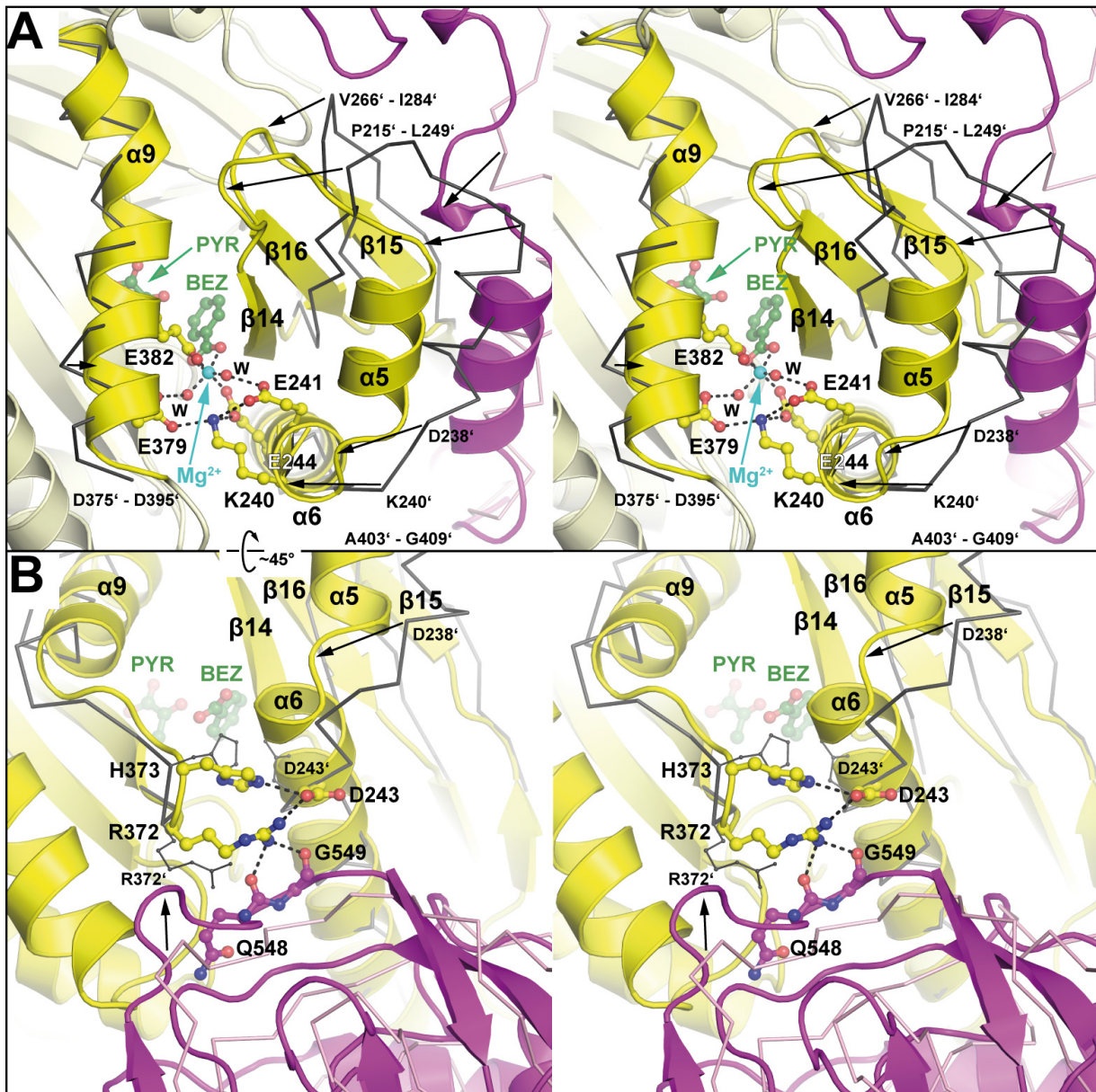
Figure 4.26: Ligand binding induces structural changes in PhzE. (A) Stereo-plot of the clamp-like closing of the MST active site upon ligand binding. Closed form structural elements are shown in yellow, their position in the open form is shown as black ribbons. Residues from the linker of the other monomer are colored in magenta. Black arrows indicate the directions of structural movements. (B) Stereo-plot showing the formation of new hydrogen bonds in the MST/GATase1 interfaces of the ligand-bound structure. Secondary structural elements are numbered according to Figure 4.1.

The GATase1 domain seems to undergo a rigid body respositioning that is triggered by the refolding within D238 – E241, where the side chain of D243 shifts from its position in the open form outwards to form hydrogen bonds with R372 and H373. As a

98

consequence, R372 and H373 reorient to interact with the backbone carbonyl group of Q548 and G549 (Figure 4.26 **B**), which might be sufficient to rotate the GATase1 for 8º to generate a 8 Å movement of positions at outer edge of the domain. Furthermore, as comparison to the solvent exposed GATase1 active site in the open form, a 5 Å shift of the loop G491 – D501 is believed to shield the active center from the solvent in the closed form (Figure 4.27 **B**).
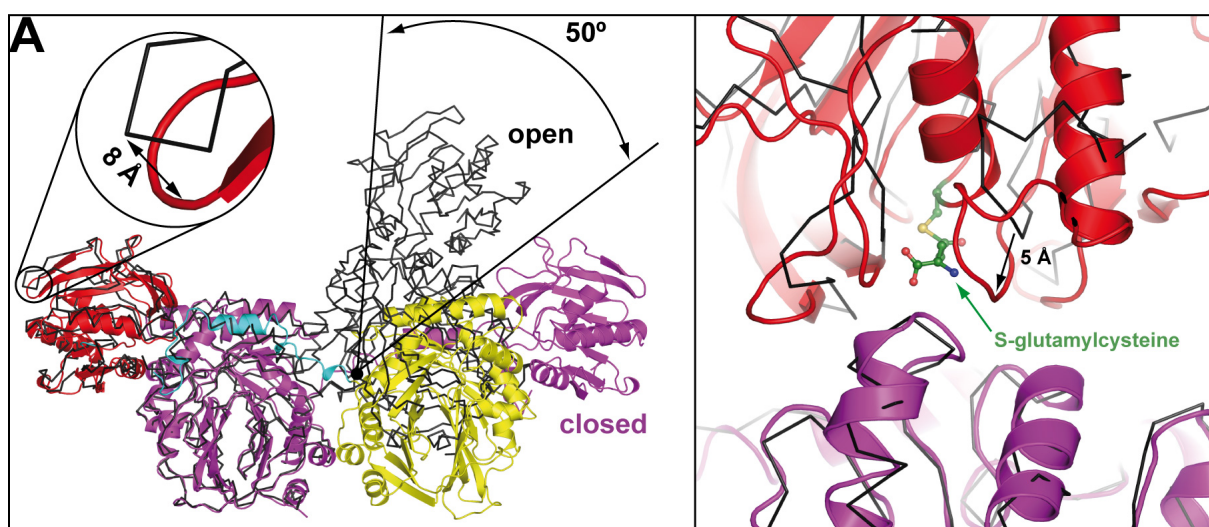

Figure 4.27: Comparison of the open (black lines) and closed structure (cartoons). Structures are superposed by the MST domain from chain A. (A) Overall structure comparison, showing the 8 Å movement of outer edge of GATase1 domain from chain A and the 50º rotation between two MST/GATase1 functional units. (B) Sheltering of the GATase1 active site. A 5 Å movement of the loop G491 – D501 is indicated by black arrow. Covalently modified C526 is shown in ball and stick presentation and colored by element. The open PhzE structure is shown in black lines and the closed PhzE is colored in red (GATase1 domain) and magenta (MST domain).

Another observation concerns the change of MST/MST interface when the ligand-free and ligand-bound structures are compared. Interestingly, the SAXS measurements confirmed that the ligand-bound crystal structure fits the SAXS envelope much better than the ligand-free structure does, even if the measurements were performed in the absence of ligands (Figure 4.13). Therefore, the relatively larger area buried by MST/MST interface in the ligand-free open structure (471 Å^2 per monomer) seems to be an artefact caused by crystal packing forces. It is conceivable since in the ligand-bound closed structure, only 219 Å^2 of solvent-accessible surface area is buried by the MST/MST interface per monomer, and given the very weak interaction between the two

MST domains, it seems not surprising that the relative position of the two MST/GATase1 functional units in the open structure is possibly a distortion induced by crystal packing. In addition, the MST/MST interactions in the open-form structure involve residues R221 – P227, which undergo ligand-induced movements such that the larger interface area in the ligand-free structure can not be retained in the ligand-bound structure.

4.2.8 The ammonia transporting channel

Since the ammonia required for ADIC formation at MST active center is generated at the GATase1 active site, an ammonia transporting mechanism is required to establish an internal pathway for two important reasons. Firstly, it makes sure that the toxic free ammonia can be directly passed through without being released to the environment. Secondly, since the local concentration of free ammonia in the solvent would be too low for the ADIC synthase reaction, an internal ammonia channel guarantees that enough ammonia is transported to the MST active center. Indeed, analysis with CAVER (Petrek et al., 2006) reveals two parts of an internal ammonia channel that connect the active centers of the MST and GATase1 domain in the closed structure of PhzE. This channel is interrupted by the side chain of E251, which forms two hydrogen bonds with N149 (Figure 4.28 **A**). In open PhzE, E251 adopts a different rotamer to interact with the side chain of K254, and the required space for this rotamer is still available in the closed conformation (Figure 4.28 **B**). This implies that E251 acts as a gatekeeper of the ammonia channel between the two active centers. Therefore, the closed conformation likely represents a late state of the catalytic cycle in which NH_3 has already been transported through the channel to the chorismate binding site. The ammonia channel is approx. 25 Å long, starting from sulfur of the glutamylcysteine and ends at C2 of the *si*-face of the chorismate, establishing a stereoselective modification to generate ADIC. The nucleophilic attack at C2 is probably assisted by the back bone carbonyl group of I216 and T304 as well as the side chain of S217 and T369, which are all in position acting as hydrogen bond acceptors, facilitating the deprotonation of NH_3. Protonation of the leaving hydroxyl group at C4 of chorismate is achieved by E201. It is

proved by biochemical studies that mutation of either E201 or the gate residue E251 leads to an inactive enzyme (Table 4.10).
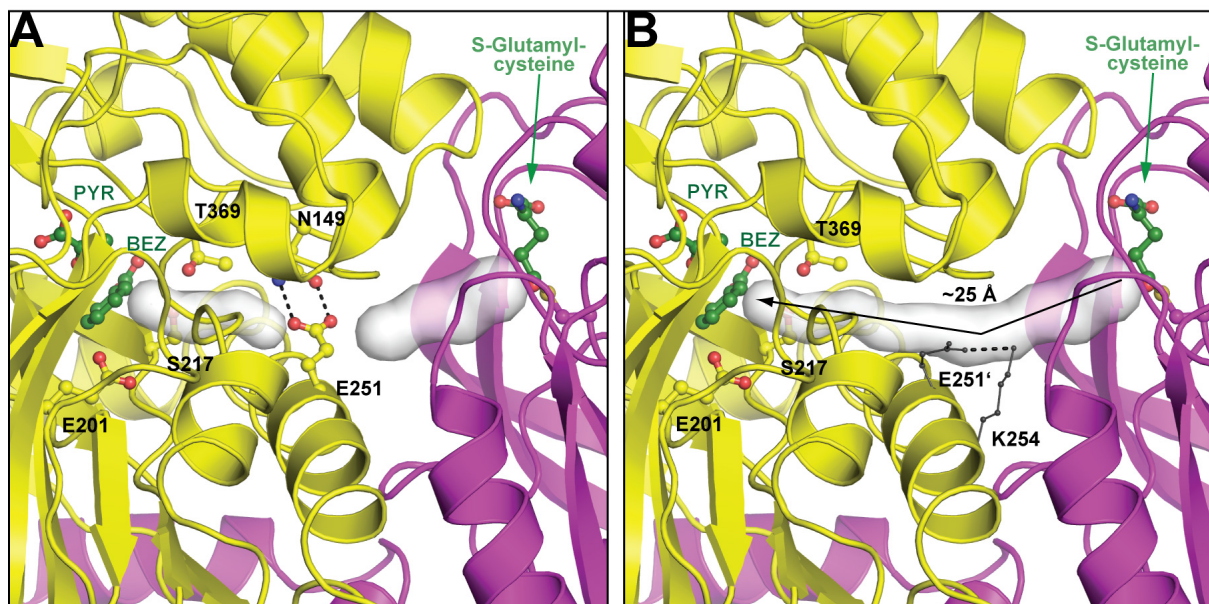

Figure 4.28: The ammonia transporting channel. (A) Two parts of the ammonia channel between the active centers of MST domain and GATase1 of the second chain, which is blocked by E251 in the closed form. (B) The full ammonia channel. Note that E251 likely acts as a gatekeeper and adopts its rotamer in the open structure, interacting with K254. The arrow indicates the trajectory of NH$_3$ to the *si*-face of C2 of chorismate.

4.3 Functional analysis of PhzE

4.3.1 ITC measurement of chorismate-PhzE binding

ITC experiments were performed to investigate the binding of chorismate to PhzE. The results show that in the absence of Mg^{2+}, the binding between chorismate and PhzE was almost undetectable (Figure 4.29 **A**). However, when Mg^{2+} was supplied to protein and ligand solution prior to the measurement, it was possible to measure the binding affinity between chorismate and PhzE (K = 2.0 × 10^5 M^{-1} and K$_d$ = 5 µM) (Figure 4.28 **B**), indicating that binding of chorismate to PhzE requires the presence of Mg^{2+}. Interestingly, the binding process is endothermic (entropy-driven), which means that the binding process needs to absorb heat from the environment. It is also an indication for

the structural changes that accompany the binding process. The stoichiometry shows that one PhzE homo dimer binds to two chorismate (N = 0.915).
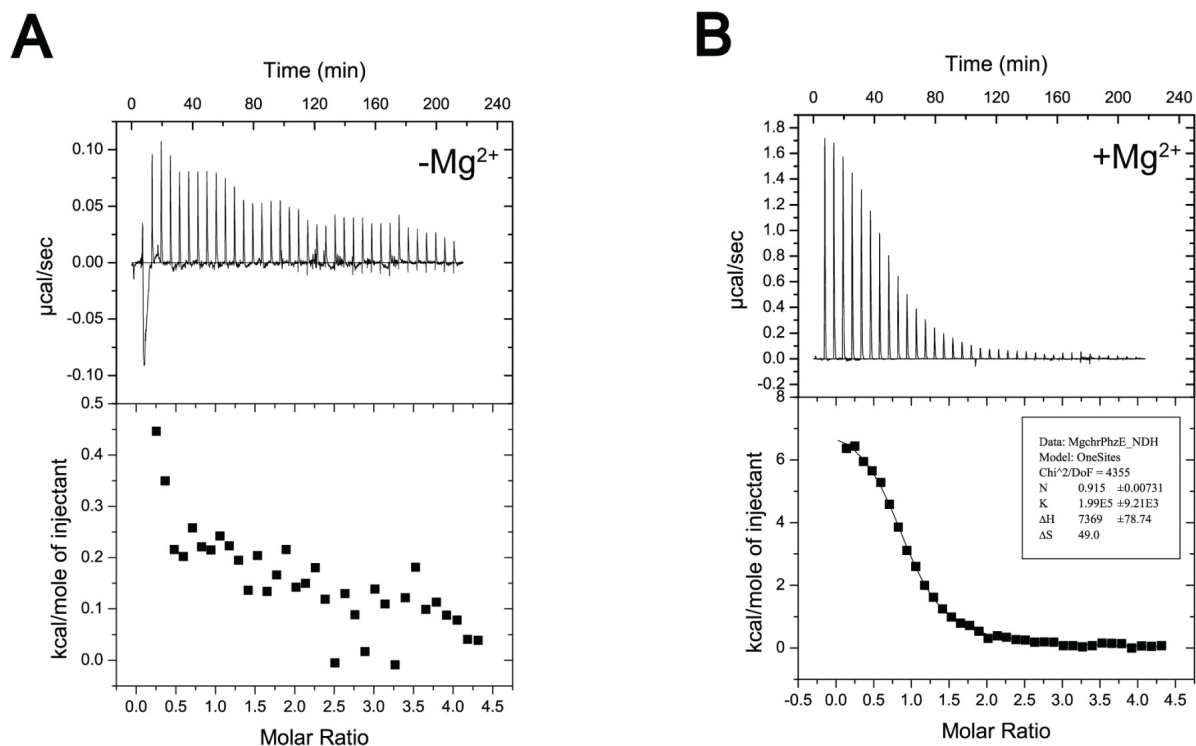
Figure 4.29: ITC titration experiment of chorismate (A) without Mg^{2+} and (B) with Mg^{2+}.

4.3.2 **Analysis of PhzE activity**

The activity of PhzE was first assayed by a UV wavelength scan. Dynamic changes in the UV absorption spectra of the reaction mixture were recorded, indicating the accumulation of a new molecular species with an extinction coefficient greater than that of chorismate, this product has been proposed as ADIC in previous report (Morollo & Bauerle, 1993) (Figure 4.30 **A**). The absorbance of the reaction solution reached a maximum level after 20-30 minutes, indicating the accumulation of the new molecules reached a steady-state level. It also shows that both chorismate and the new molecule have a significant absorbance at 280 nm. Control experiments show that chorismate, glutamine, Mg^{2+} and PhzE are all necessary for the appearance of the new compound.
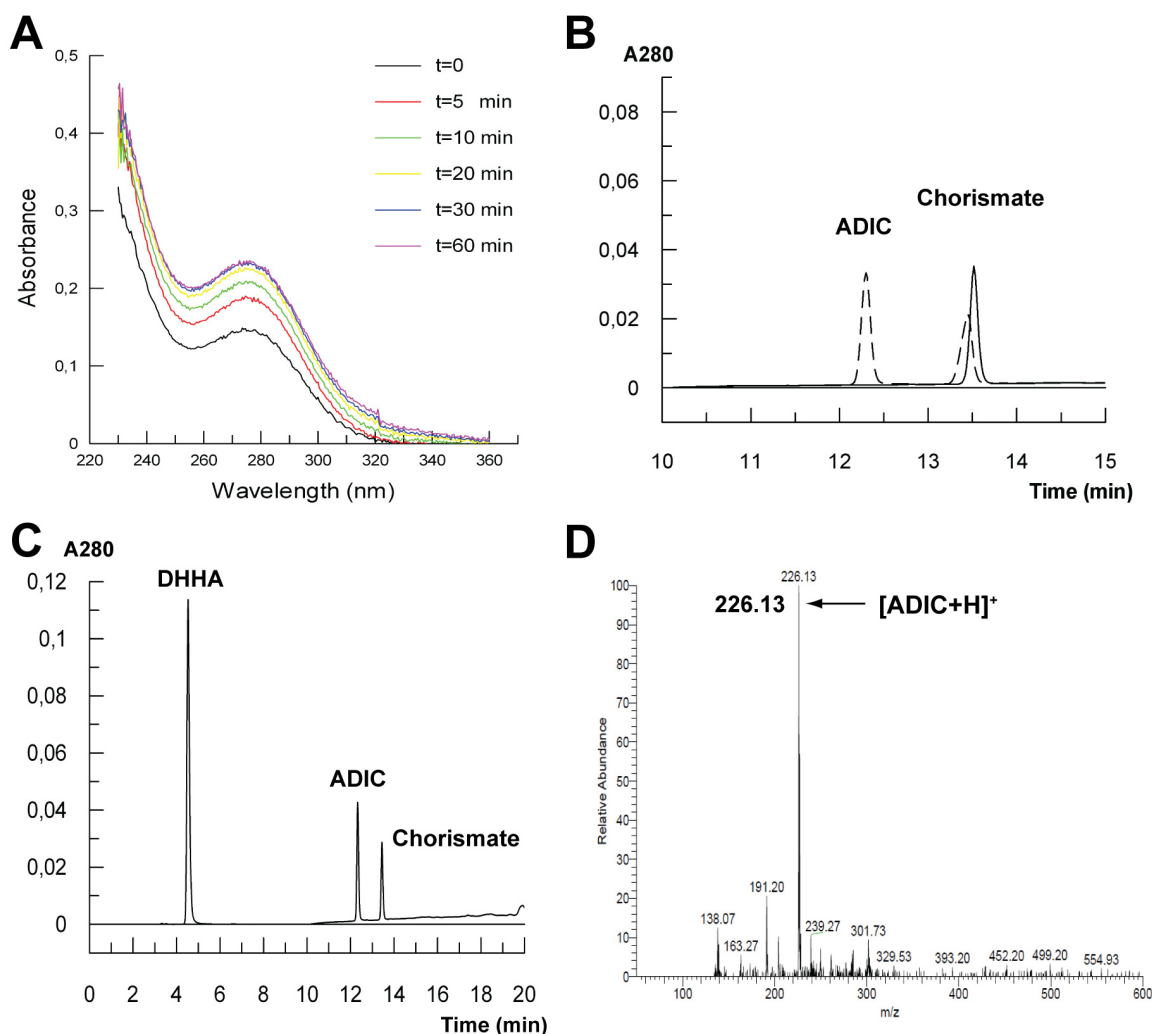
Figure 4.30: Characterization of PhzE reaction and product. (A) UV spectra of PhzE catalyzed reaction. Buildup of product was monitored by increasing of absorption at 280 nm. (B) HPLC chromatogram of of reaction mixture prior to the addition of PhzE (—) and 10 min after addition of PhzE (- -). (C) HPLC chromatogram of PhzE/D reaction. PhzD was added to PhzE reaction solution 10 min after reaction started. Conversion of ADIC to DHHA was observed. (D) ESI-MS spectrum of purified ADIC.

HPLC fractionation of the reaction mixture revealed that the product of PhzE reaction was eluted earlier than chorismate. The retention time of the product was around 12.4 minutes comparing to 13.5 minutes for the substrate chorismate (Figure 4.30 **B**). In addition, a coupled enzymatic assay showed that the product can be further converted to DHHA by PhzD, an isochorismatase that catalyzes the step following PhzE in the phenazine biosynthesis pathway (Figure 4.30 **C**). The molecular mass of PhzE's product was confirmed to be 226.13 Da (M+H$^+$), indicating the product is ADIC (Mw =

225 Da) (Figure 4.30 **D**). Furthermore, it was possible to visualize ADIC at 1.9 Å resolution in the crystal structure of an inactive D38A mutant of PhzD (Figure 4.43). These evidences together prove that PhzE is an ADIC synthase.

To further investigate the reaction condition for PhzE, the pH and Mg^{2+} dependency of the enzyme were tested (Figure 4.31). Results indicated that under the condition of the activity assay (final concentration of PhzE ≈ 100 nM), PhzE is most active at pH 7.0 and with Mg^{2+} concentration of 1 mM. At the pH below neutral, PhzE is almost completely inactive.
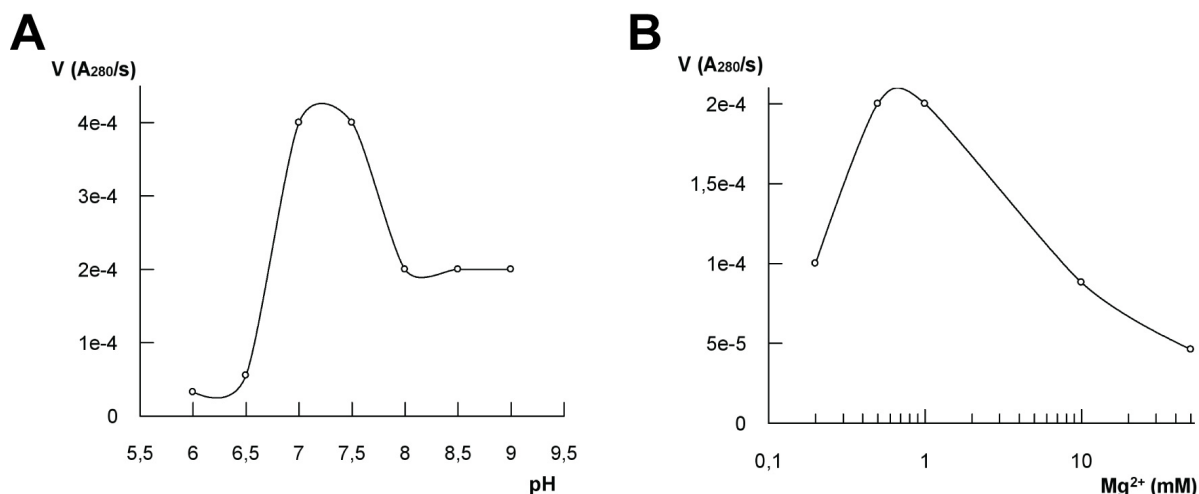
Figure 4.31: pH and Mg^{2+} dependence of the PhzE reaction. (A). pH optimization of PhzE reaction. (B) Optimization of mg^{2+} concentration in PhzE reaction.

4.3.3 Determination of the extinction coefficient value of ADIC

Due to the instability of ADIC in solution (half-life ≈34 hr at 22 ºC), the reported extinction coefficient values of ADIC vary from ε = 11,500 $M^{-1}.cm^{-1}$ for enzymatically produced ADIC (Morollo & Bauerle, 1993) to ε = 5,900 $M^{-1}.cm^{-1}$ for chemical synthetic products (Policastro et al., 1984). Therefore, in order to obtain a more accurate ε value of ADIC for the kinetic measurement, a quatitative HPLC experiment to determine ε_{280} of ADIC was designed. The calculated extinction coefficient value of ADIC is ε = 11,490 $M^{-1} \cdot cm^{-1}$. This value was used for kinetic experiments and quantitative analysis of ADIC throughout this study. Experimental details are described in chapter 3.6.4.

4.3.4 **Kinetic characterization of PhzE**

Kinetic parameters of PhzE were determined and compared with the values reported for other MST enzymes. Since it was indicated by the crystal structures that PhzE was partially contaminated with Zn^{2+} during the purification steps, kinetic measurements for both Zn^{2+}-bound and Zn^{2+}-free PhzE were therefore performed (Figure 4.32).
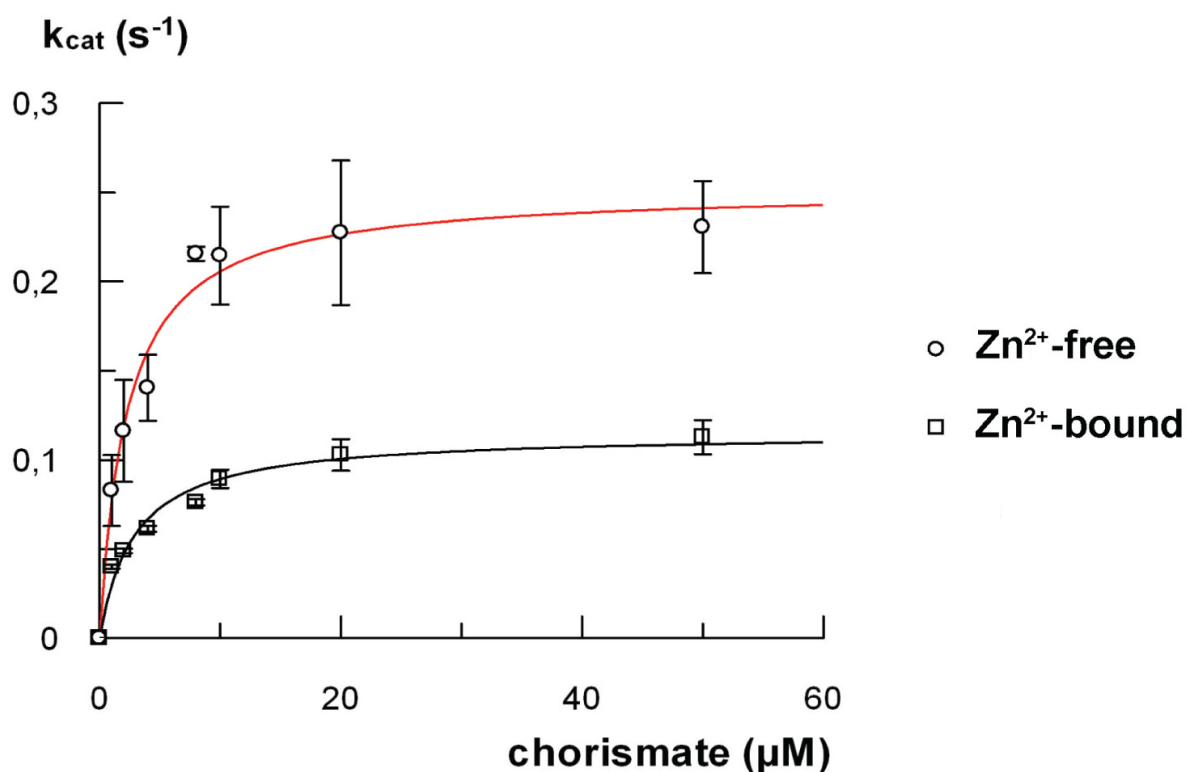
Figure 4.32: Michaelis-Menten kinetic curve of PhzE. Red curve corresponds to Zn^{2+}-free PhzE, while black curve indicates Zn^{2+}-bound PhzE.

After treatment with EDTA to remove bound Zn^{2+}, kinetic parameters for chorismate were $K_m = 2.5 \pm 0.5$ µM and $k_{cat} = 0.3 \pm 0.01$ s^{-1}. The K_m value was similar to that of Zn^{2+} contaminated PhzE, while the k_{cat} value was almost 3 times larger than that of Zn^{2+}-bound enzyme, indicating that the contaminating Zn^{2+} reduced the enzyme activity. The Michaelis-Menten constant K_m and the catalytic efficiency of PhzE is similar to most of the related enzymes, however it has a relatively slow turnover comparing to AS (TrpE/G). Kinetic parameters of PhzE and related enzymes are summarized in Table 4.9.

Enzyme	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1}s^{-1})	Description
PhzE	2.4 ± 0.51	0.3 ± 0.01	1.0×10^5	Metal-free
PhzE (untreated)	3.0 ± 0.48	0.1 ± 0.01	4.3×10^4	Zn^{2+}-bound
TrpE/G	4.0	9.3	2.3×10^6	AS from *S.typhimurium*
Irp9	4.2	0.13	3.2×10^4	SS from *Y.enterocolitica*
EntC	7 ± 0.8	0.62	8.8×10^4	IS from *E.coli* involved in enterobactin biosynthesis
MenF	192 ± 7	3.6 ± 0.08	1.9×10^4	IS from *E.coli* involved in menaquinone biosynthesis

Table 4.9: Kinetic parameters of PhzE and related enzymes with respect to chorismate. Values for related enzymes (TrpE/G, Irp9, entC & MenF) have been taken from the literature.

4.3.5 **Regulation of PhzE activity**

Since the Zn^{2+}-bound PhzE showed reduced activity comparing to Zn^{2+}-free protein, the influence of divalent transition metal cations to PhzE was therefore examined. Indeed, Zn^{2+}, Mn^{2+} and Ni^{2+} were confirmed to be able to inhibit PhzE (Figure 4.33). This result is an indication that on top of the transcriptional regulation by quorum sensing, divalent transition cations might add another layer to the overall control of phenazine biosynthesis.
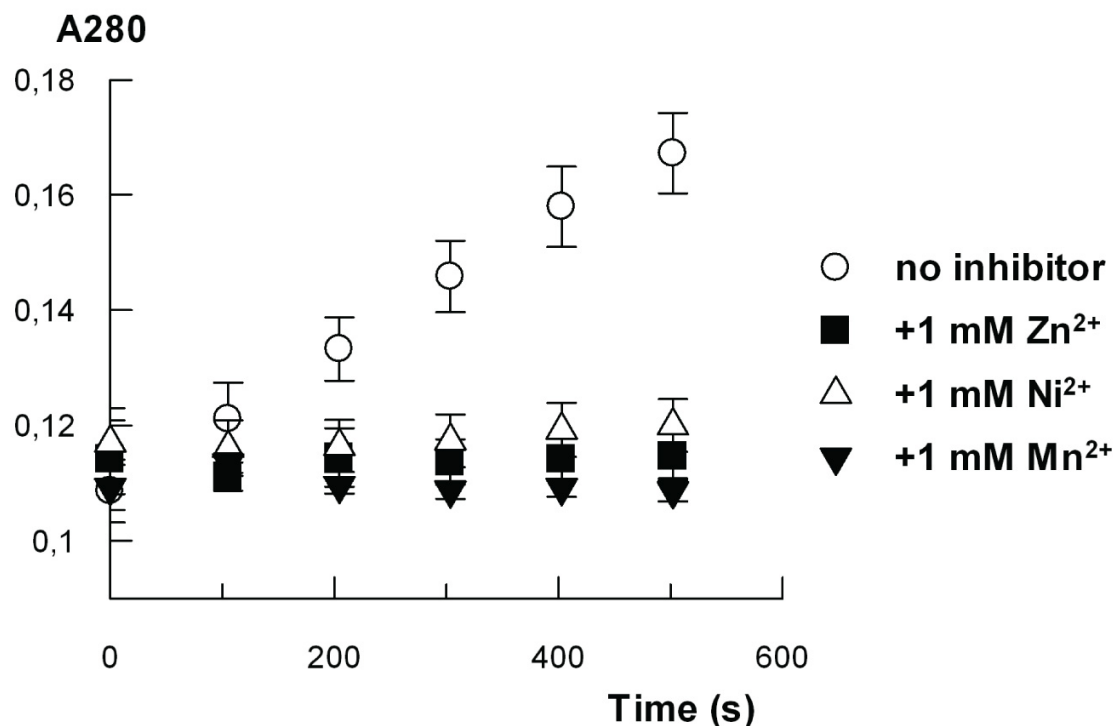
Figure 4.33: Preliminary Inhibition test of divalent transition cations to PhzE.

In AS from *S. marcescens*, the enzyme is feedback inhibited by *L*-tryptophan, the major downstream product of anthranilate in primary metabolism (Spraggon et al., 2001). Whether PhzE is also feeback inhibited by *L*-tryptophan or by the intermediate (DHHA) / product (PCA) of the *phz*-operon was investigated. However, in case of *L*-tryptophan and DHHA, no inhibition was observed (Figure 4.34 **A** and **B**), while only very high concentration (2 mM) of PCA showed partial inhibition to PhzE (Figure 4.34 **C**). This might be a consequence of the structural difference between PhzE and AS, where the allosteric *L*-tryptophan binding site in AS is blocked in PhzE by the side chains of R26 and W184, due to variations at the N-terminus and an insertion near β10 (Figure 4.34 **D**)
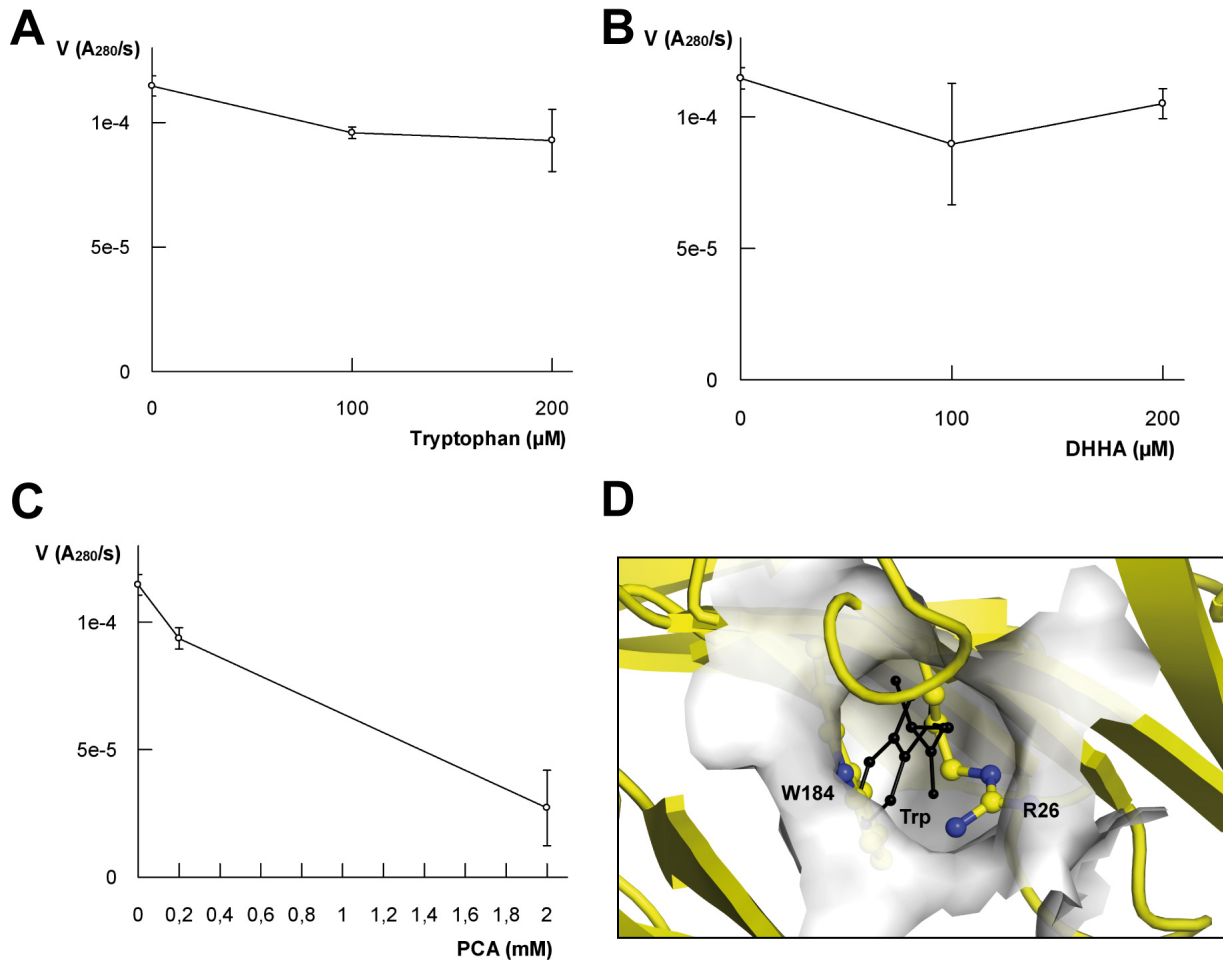
Figure 4.34: Allosteric inhibition of PhzE. (A) Inhibition test of PhzE by *L*-tryptophan. (B) Inhibition test of PhzE by DHHA. (C) Inhibition test of PhzE by PCA. (D) The allosteric tryptophan binding pocket of AS from *S. marcescens* (PDB entry 1I7Q/1I7S) is blocked by R26 and W184 in PhzE. Surface and tryptophan coordinates (black) are from PDB entry 1I7S.

4.3.6 Mutagenesis studies of PhzE

AS synthesizes ADIC from chorismate and subsequently lyses it to anthranilate and pyruvate in the active center of the enzyme, while PhzE releases ADIC as its end product. The reason for this catalytical difference of these two enzymes remains unclear. Structural comparison between PhzE and AS indentifies residues that could possibly convert PhzE into an AS. Three residues S217, S368 and T369 in the first coordination sphere of chorismic acid in which PhzE differs from AS were mutated to alanine, alanine and glycine, respectively. While S368A was expressed in inclusion bodies from *E.coli*,

S217A and T369G retained lower levels of ADIC synthase activity and S368A/T369G double mutant together with S217A/S368A/T369G triple mutant were completely inactive. However, these mutations have not converted PhzE into AS since neither the consumption of substrate chorismate nor production of anthranilate was observed, indicating the fate of ADIC in these enzymes are possibly determined by more subtle mechanisms. Interestingly, previous studies on AS from *S. enterica* showed that the mutation of H398 to methionine almost completely abolished the ADIC lyase activity of the enzyme, resulting in the conversion of AS to ADIC synthase (Morollo & Bauerle, 1993). However, the histidine is also conserved in PhzE as H279 (Figure 4.17 **A**), showing that this residue alone cannot be the determining factor for the catalytic difference between PhzE and AS.

Other residues, including the catalytic residue E201, the gate keeper residue for the ammonia channel E251 and the four glutamic acid that are involved in Mg^{2+} coordination E241, E244, E379 and E382, were also mutated separately to investigate their roles in PhzE activity. Since E201 is crucial for the protonation of C4 hydroxyl group of chorismate, E201Q was inactive as expected. Mutants E251Q and E251A were also both inactive, due to the fact that the disruption of hydrogen bonds of E251 with N149 and K254 probably destroyed the integrity of the ammonia transporting channel. Of the four glutamic acids involved in Mg^{2+} coordination, E244 and E379 are absolutely required for enzyme activity, while E241A and E382A retained lower activity comparing to native PhzE.

Experimental statistics of PhzE and PhzE mutants are summarized in Table 4.10.

Enzyme	K_m (µM)	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1}s^{-1})	Description
PhzE	2.4 ± 0.51	0.3 ± 0.01	1.0×10^5	Metal-free
PhzE (untreated)	3.0 ± 0.48	0.1 ± 0.01	4.3×10^4	Zn^{2+} -bound
E201Q	n.d.	n.d.	n.d.	Catalytic residue
S217A	4.0 ± 1.0	0.11 ± 0.008	2.8×10^4	mutation according to AS
E241A	2.3 ± 0.5	0.03 ± 0.002	1.5×10^4	Involved in Mg^{2+} coordination
E244A	n.d.	n.d.	n.d.	Involved in Mg^{2+} coordination
E251Q	n.d.	n.d.	n.d.	Gate-keeping residue of the ammonia channel
E251A	n.d.	n.d.	n.d.	
S368A	n.d.	n.d.	n.d.	Insoluble protein
S368A/T369G	n.d.	n.d.	n.d.	Double-mutation according to AS
S217A/S368A/T369G	n.d.	n.d.	n.d.	Triple-mutation according to AS
T369G	3.5 ± 1.9	0.01 ± 0.002	3.7×10^3	mutation according to AS
E379A	n.d.	n.d.	n.d.	Involved in Mg^{2+} coordination
E382A	1.7 ± 0.6	0.02 ± 0.002	1.3×10^4	Involved in Mg^{2+} coordination

Table 4.10: Kinetic parameters of PhzE and Phze mutants with respect to chorismate.

Part II: Structural studies of PhzD

In this part, the crystal structure of PhzD and PhzD mutant D38A from *P. fluorescens* are discussed. Since no ADIC was observed in the crystal structure of PhzE and because PhzD catalyzes the step next to PhzE in phenazine biosynthesis pathway, converting ADIC to DHHA, the main purpose of studying PhzD was to prove that ADIC is indeed the product of PhzE by crystallographic soaking experiments.

According to a previous report, the aspartic acid D38 of PhzD from *P. aeruginosa* acts as a general acid to protonate the substrate ADIC (Figure 4.35), and the mutation of aspartic acid to alanine results in completely lost enzyme activity (Parsons et al., 2003). Therefore, the D38A mutant of PhzD from *P. fluorescens* 2-79 was prepared and crystallized as an inactive form of the enzyme and soaked with purified PhzE reaction product, in order to examine if ADIC could be observed in the active center of PhzD.
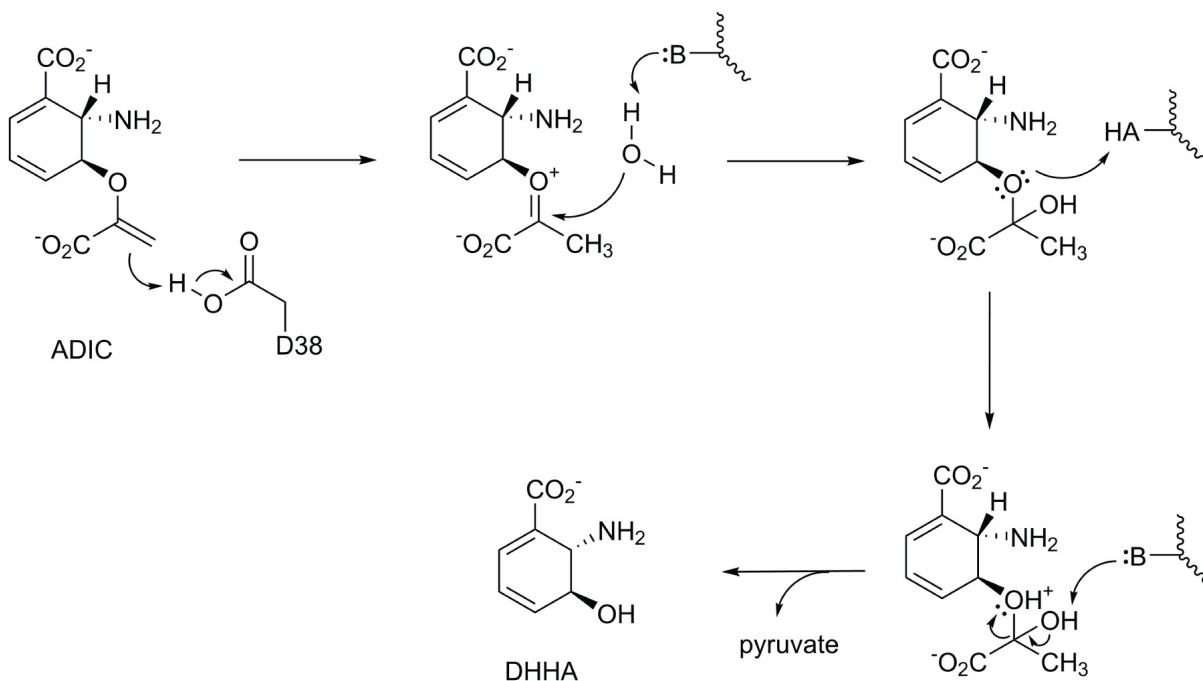

Figure 4.35: Reaction mechanism of PhzD proposed by Parsons et al. (Parsons et al., 2003).

4.4 Structural analysis of PhzD, PhzD-D38A and PhzD-D38N

4.4.1 Sequence alignment of PhzD from different species

Results from the Protparam online proteomics server (Gasteiger et al., 2005) revealed that PhzD *P.fluorescens* 2-79 has a molecular mass of 23074.3 Da and theoritical pI of 5.25. A Blast (National Center for Biotechnology Information web tool) search against the non-redundant protein sequence database indicated that the protein is highly conserved among phenazine producing bacterial species (Table 4.11), and a complete sequence alignment was performed between PhzD from *P.fluorescens* 2-79 (this study) and PhzD from *P. aeruginosa* PAO1 (Parsons et al., 2003) with Clustal X (Larkin et al., 2007).

Protein	Origin	Query coverage (%)	Sequence identity (%)	E value
PhzD	*P. chlororaphis*	100	93	3e-110
PhzD	*P. aeruginosa*	99	78	2e-93
PhzD	*B. lata 383*	99	61	4e-76
Putative DHHA synthase	*S. cinnamonensis*	97	67	2e-75

Table 4.11: Results from Blast search of PhzD from *P.fluorescens* 2-79 against the non-redundant protein sequence database. 4 hits are listed according to their max scores (not shown). The E value describes the quality of the alignment, the lower the E value, the more significant the match is.
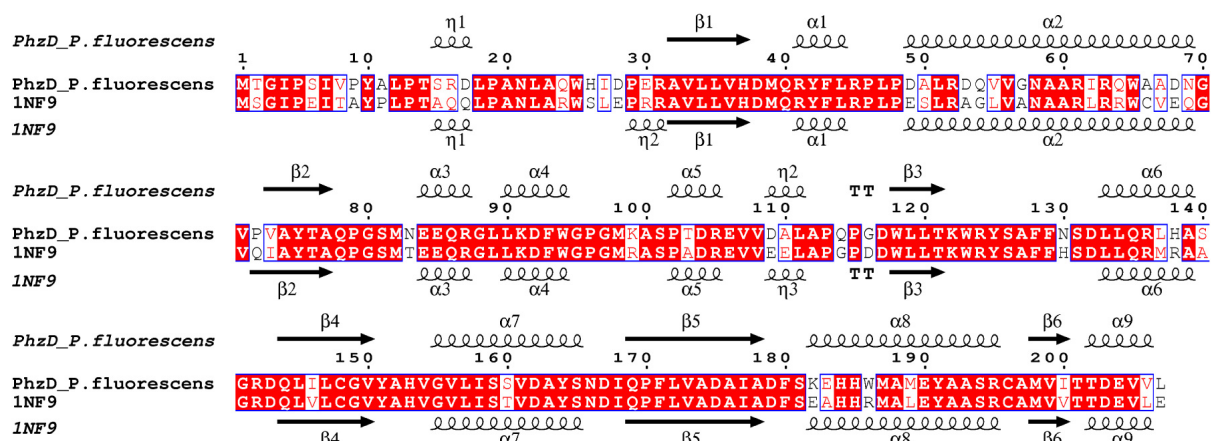
Figure 4.36: Structure-based sequence alignment between PhzD from *P. fluorescens* 2-79 and PhzD from *P. aeruginosa* PAO1 (PDB entry 1NF9). Secondary structural elements are shown (on top and botton of the alignment). The figure was prepared using ESPript 2.2 (Gouet et al., 2003).

4.4.2 **Crystallization and soaking experiments**

4.4.2.1 Crystallization of PhzD, PhzD-D38A and PhzD-D38N

Crystals of native PhzD and the mutants were obtained from the condition indentified preciously, in which 1 µL + 1 µL hanging drops were equilibrated against a reservoir containing 0.1 M sodium cacodylate buffer pH 6.5, 0.2 M sodium acetate and 25% (w/v) PEG 4000 at 20 ºC. Interestingly, Diffraction quality crystals were obtained in three different forms, with space group $P4_32_12$, $C222_1$ or $P2_12_12_1$, respectively, under the same crystallization condition.

Crystals with space group of $P4_32_12$ have a tetragonal shape and grew to a full size of 0.4 × 0.2 × 0.2 mm in 1 day, whereas the cubic crystals with orthorhombic space groups $C222_1$ or $P2_12_12_1$ grew to a full size of 0.4 × 0.4 × 0.1 mm in about 3 days (Figure 4.37). The best crystal of native PhzD diffracted to approx. 1.9 Å while dataset was collected at 100 K in house on a Rigaku Micro-MAX-007 HF generator with a MAR345 image plate detector.
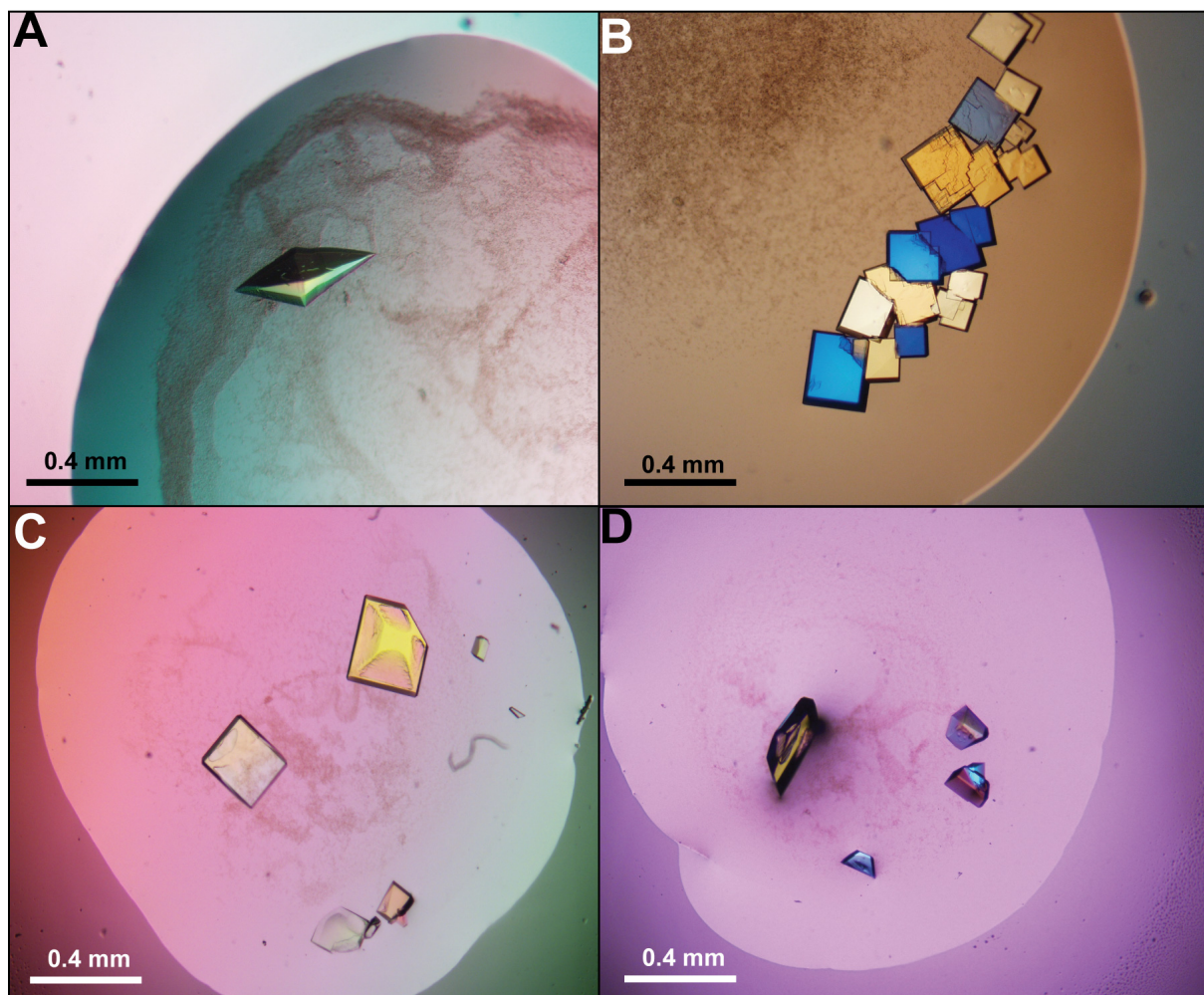
Figure 4.37: Crystals of PhzD and mutants. (A) Tetragonal crystal of PhzD (space group P4$_3$2$_1$2). (B) Cubic crystals of PhzD (space group C222$_1$ or P2$_1$2$_1$2$_1$). (C) Crystals of PhzD-D38A. (D) Crystals of PhzD-D38N.

4.4.2.2 Production of ligand-free PhzD crystals

The ligand-free crystals were obtained by a 'kick-out' experiment, where crystals from the original crystallization condition were transferred to a new drop containing 0.1 M Bis-Tris pH 6.5, 20% (w/v) PEG 4000, 0.1 M NaCl and were incubated for 60 min. Data for apo-PhzD were collected to a resolution of 1.8 Å on an Rigaku Micro-MAX-007 HF generator with a MAR345 image plate detector in house, and to 1.3 Å at the Swiss Light Source (SLS, Villigen, Switzerland) on beam line X10SA equipped with a MAR225 CCD detector (PILATUS 6M pixel detector).

4.4.2.3 Soaking of PhzD and the mutants with ADIC

To perform the soaking experiments, crystals were first prepared as apo-form as described above (see chapter 4.4.2.2). 1 mM ADIC was subsequently added into the drop and crystals were further incubated for 60 min. Data for were collected in house on an Rigaku Micro-MAX-007 HF generator with a MAR345 image plate detector or at the Swiss Light Source (SLS, Villigen, Switzerland) on beam line X10SA equipped with a MAR225 CCD detector (PILATUS 6M pixel detector).

4.4.3 **Data collection statistics**

Except for the native PhzD data, which was collected in house, data for all other crystals were collected both in house and at the SLS. Full data collection statistics are summarized in Table 4.12.

4.4.4 **Structure determination and refinement statistics**

The structures of PhzD *P. fluorescens* and mutants were solved with PHASER (McCoy et al., 2007) using PhzD from *P.aeruginosa* (PDB entry 1NF9) as search model. The structure of native PhzD was refined to R_{work} = 14.8% and R_{free} = 18.7% with a Ramachandran plot (Figure 4.38 **A**) showing that 97.6% of all residues lie in the favorite region (2.4% in the allowed region). For native PhzD soaked with ADIC, the structure was refined to R_{work} = 18.9% and R_{free} = 22.2% with a Ramachandran plot (Figure 4.38 **B**) showing that 96.0% of all residues lie in favorite region (2.0% in the allowed region). For ligand-free PhzD, the structure was refined to R_{work} = 19.0% and R_{free} = 21.0%. Ramachandran plot (Figure 4.38 **C**) indicated that 96.9% of all residues lie in the favorite region (3.1% in the allowed region).

The model of PhzD-D38A soaked with ADIC was refined to R_{work} = 15.1% and R_{free} = 17.5% with a Ramachandran plot (Figure 4.38 **D**) showing that 96.0% of all residues lie in the favorite region (2.0% in the allowed region) and the structure of PhzD-D38N soaked with ADIC was finally refined to R_{work} = 18.8% and R_{free} = 23.3%. The Ramachandran plot (Figure 4.38 **E**) showed that 96.5% of all residues lie in the favorite region (1.5% in the allowed region). Complete refinement statistics are summarized in table 4.13.

	PhzD	PhzD+ADIC in house	PhzD+ADIC SLS	apo-PhzD in house	apo-PhzD SLS	D38A+ADIC in house	D38A+ADIC SLS	D38N+ADIC in house	D38N+ADIC SLS
Space group	$P4_32_12$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions a, b, c (Å)	50.8, 50.8, 151.6	70.6, 73.3, 81.3	70.6, 73.3, 81.3	69.3, 73.8, 74.0	69.3, 73.8, 74.0	68.5, 72.1, 79.0	68.5, 72.1, 79.0	70.6, 73.1, 81.1	70.6, 73.1, 81.1
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å) (highest shell)	50 – 1.87 (1.92 – 1.87)	50 – 1.85 (1.95 – 1.85)	50 – 1.6 (1.7 – 1.6)	50 – 1.8 (1.9 – 1.8)	50 – 1.3 (1.4 – 1.3)	50 – 1.9 (2.0 – 1.9)	50 – 1.35 (1.45 – 1.35)	50 – 2.2 (2.3 – 2.2)	50 – 1.65 (1.75 – 1.65)
$R_{sym}(I)$	1.7 (3.8)	7.5 (31.0)	7.5 (42.1)	5.6 (23.0)	7.2 (47.3)	5.0 (28.0)	7.8 (34.4)	12.6 (45.6)	13.3 (38.5)
$I / \sigma(I)$	53.5 (35.5)	22.2 (5.6)	12.7 (3.7)	30.1 (7.4)	21.0 (3.4)	18.7 (5.0)	13.4 (4.1)	11.3 (4.0)	9.9 (3.4)
Completeness (%)	97.6 (95.2)	99.4 (98.8)	100 (100)	99.9 (100)	99.9 (99.8)	97.7 (98.2)	96.2 (95.7)	99.5 (98.9)	95.2 (96.6)
Redundancy	8.3 (8.1)	5.9 (5.9)	4.7 (4.6)	7.2 (7.0)	8.7 (8.3)	2.5 (2.4)	2.3 (2.1)	2.1 (2.1)	2.2 (2.2)

Table 4.12: Data collection statistics for PhzE crystals.

	PhzD	PhzD/ADIC in house	PhzD/ADIC SLS	apo-PhzD in house	apo-PhzD SLS	D38A/ADIC in house	D38A/ADIC SLS	D38N/ADIC in house	D38N/ADIC SLS
Resolution (Å)²	50 – 1.87 (1.92 – 1.87)	50 – 1.85 (1.95 – 1.85)	50 – 1.6 (1.7 – 1.6)	50 – 1.8 (1.9 – 1.8)	50 – 1.3 (1.4 – 1.3)	50 – 1.9 (2.0 – 1.9)	50 – 1.35 (1.45 – 1.35)	50 – 2.2 (2.3 – 2.2)	50 – 1.65 (1.75 – 1.65)
No. reflections (unique)	139139 (16768)	98307 (16639)	253802 (53743)	126022 (17539)	405912 (46877)	145231 (58140)	387616 (169387)	81640 (38563)	205188 (92948)
R_{work}	14.8 (14.4)	17.6 (22.4)	18.9 (21.4)	16.7 (16.7)	19.0 (21.8)	14.3 (25.2)	15.1 (17.5)	17.1 (16.3)	18.8 (18.3)
R_{free}	18.7 (22.9)	21.3 (27.5)	22.2 (25.3)	20.9 (19.1)	21.0 (28.5)	19.2 (33.1)	17.5 (19.8)	22.4 (25.7)	23.3 (25.6)
No. atoms / B-factors (Å²)									
Protein	1642 / 7	1588 / 25	1588 / 27	1540 / 21	1538 / 18	3239 / 23	3234 / 11	3220 / 31	3215 / 16
Ligand/ion	9 / 8	- / -	- / -	- / -	- / -	36 / 24	- / -	- / -	- / -
Water	330 / 31	98 / 31	100 / 35	143 / 31	157 / 28	454 / 33	564 / 25	230 / 35	411/28
R.m.s deviations Bond lengths / angles (Å / °)	0.012 / 1.386	0.027 / 1.873	0.032 / 2.182	0.026 / 1.973	0.028 / 2.060	0.025 / 1.904	0.028 / 2.210	0.021 / 1.899	0.025 / 2.075

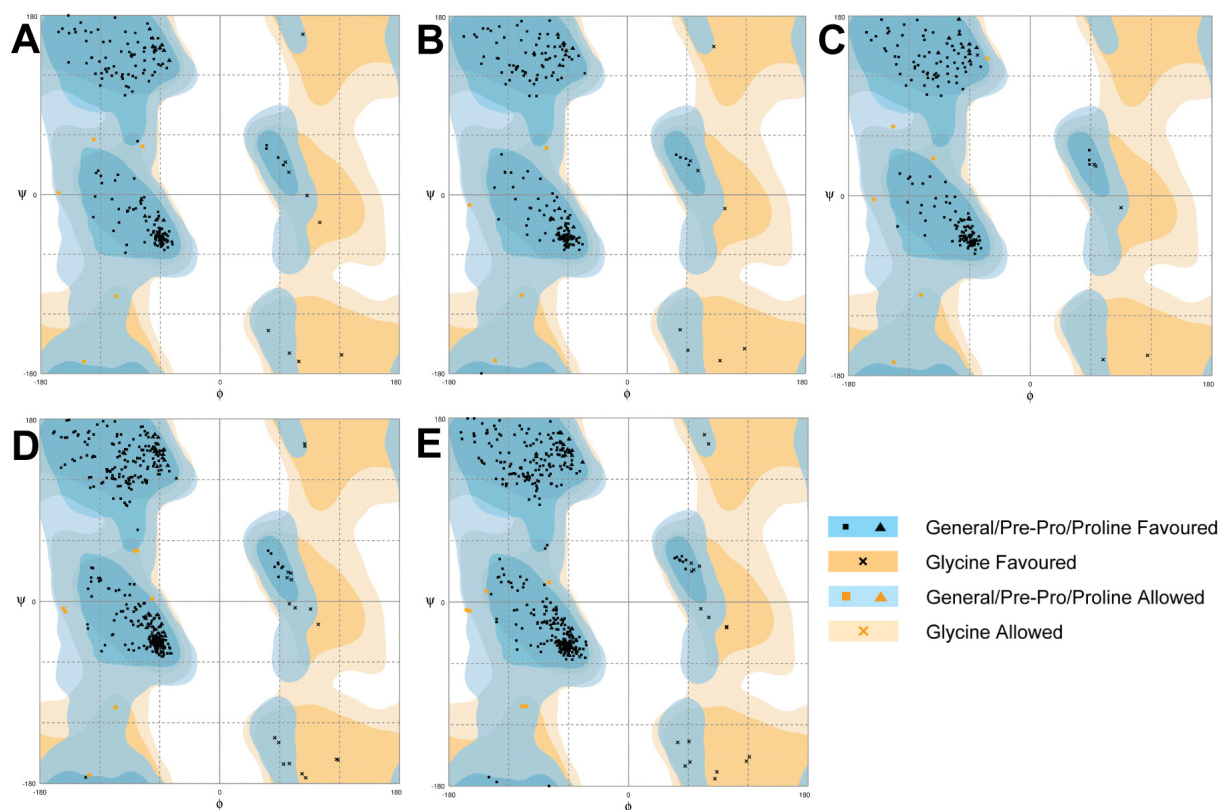**Table 4.13: Refinement statistics of PhzD and mutants.**

Figure 4.38: Ramachandran plot showing torsion angels of peptide bonds of PhzD and mutant structures. (A) Native PhzD. (B) Native phzD soaked with ADIC. (C) Ligand-free native PhzD. (D) PhzD-D38A soaked with ADIC. (E) PhzD-D38N soaked with ADIC. All models except that of native PhzD were refined against datasets collected at SLS. Plots were generated by RAMPAGE from the CCP4 suite.

4.4.5 **Overall structure of PhzD**

The 23 kDa PhzD folds into a α/β single domain, with a six-strand parallel β-sheet lying in the central part of the structure, which is covered by four helices on both sides, a single long one (α2) on one side and three shorter ones (α6, α7 and α8) on the other. As identified by size-exclusion chromatography, PhzD is a dimer in solution. Except for the crystal form $P2_12_12_1$, which contains a PhzD dimer formed by crystallographic symmetry operator in one asymmetric unit, in both of the crystal form $P4_32_12$ and $C222_1$ only one molecule of PhzD was observed in the asymmetric unit. The two-fold dimer interface buries 1467 $Å^2$ surface area of each of the monomers (16% of the overall surface area, calculated with the $P2_12_12_1$ structure).
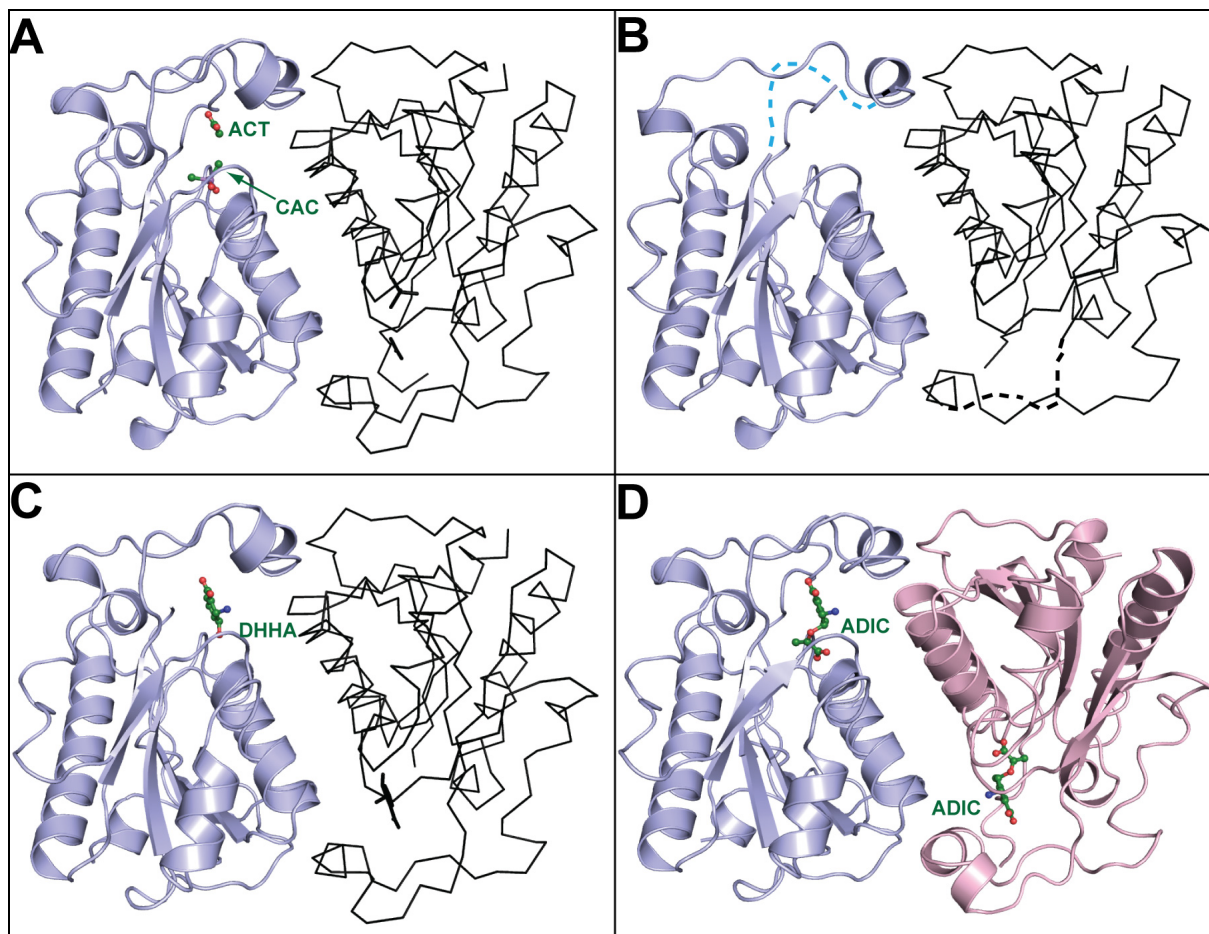
Figure 4.39: Overall structure of PhzD and mutants. (A) Native PhzD in complex with cacodylate (CAC) and acetate (ACT). (B) Ligand-free native PhzD. The flexible loop (P79 – R87) is shown as dashed line. (C) Native PhzD in complex with DHHA. (D) D38A in complex with ADIC. Black ribbon indicates the two-fold crystallographic symmetry mate that resembles the physiological dimer.

All structures of PhzD and the mutants were superposed and compared to the native PhzD structure. The results showed that the overall structures of all PhzD and mutants are almost identical except for the ligand-free structure (table 4.14). Indeed, comparing to ligand-bound structures, electron density is missing for the loop region between P79 – R87 in the ligand-free model, indicating high flexibility of this region (Figure 4.39 **B**). Since the side chain of R87 is acting as a hydrogen bond donor and is absolutely required for ligand binding (see chapter 4.4.6.4), it is therefore believed that the loop P79 – R87 is acting as a gate sitting between the active center and solvent environment to control the entrance/exit of the ligand.

Structure	Chain ID	Overall RMSD (Å)
Ligand-free PhzD	A	0.872
PhzD/ADIC	A	0.318
D38A/ADIC	A	0.478
	B	0.473
D38N/ADIC	A	0.454
	B	0.452

Table 4.14: RMSD of PhzD and mutants. Native PhzD was used as fixed model. Values were generated using SUPERPOSE in the CCP4 suite.

4.4.6 Active center of PhzD

In this section, the crystal structures of PhzD are discussed especially focusing on the details of its active site. Except for the structure of PhzD in complex with cacodylate and acetate, which was refined against in-house data, models for both the PhzD ligand-free and the PhzD soaked with ADIC structures were refined against synchrotron data.

4.4.6.1 PhzD in complex with cacodylate and acetate

The crystallization conditions from which PhzD crystals were obtained contains cacodylate buffer and sodium acetate. Indeed, additional electron density was observed at the active center of native PhzD, which could be refined as cacodylate and acetate (Figure 4.40 **A**).

At the active site, the carboxylate of acetate forms hydrogen bonds with the side chains of R87 and Q78, as well as with two water molecules, which are held in position by the side chain of Y42 and Q78. Cacodylate is in hydrogen bonding distance to the main chain amide of Y151 and the side chain of D38. In addition, a *cis* peptide bond is found between residues Y151 and A152, which has also been observed in related structures of pyrazinamidase from *Pyrococcus horikoshii* and PhzD from *Pseudomonas aeruginosa* (Du et al. 2001) (Parsons et al. 2003) (Figure 4.40 **B**).
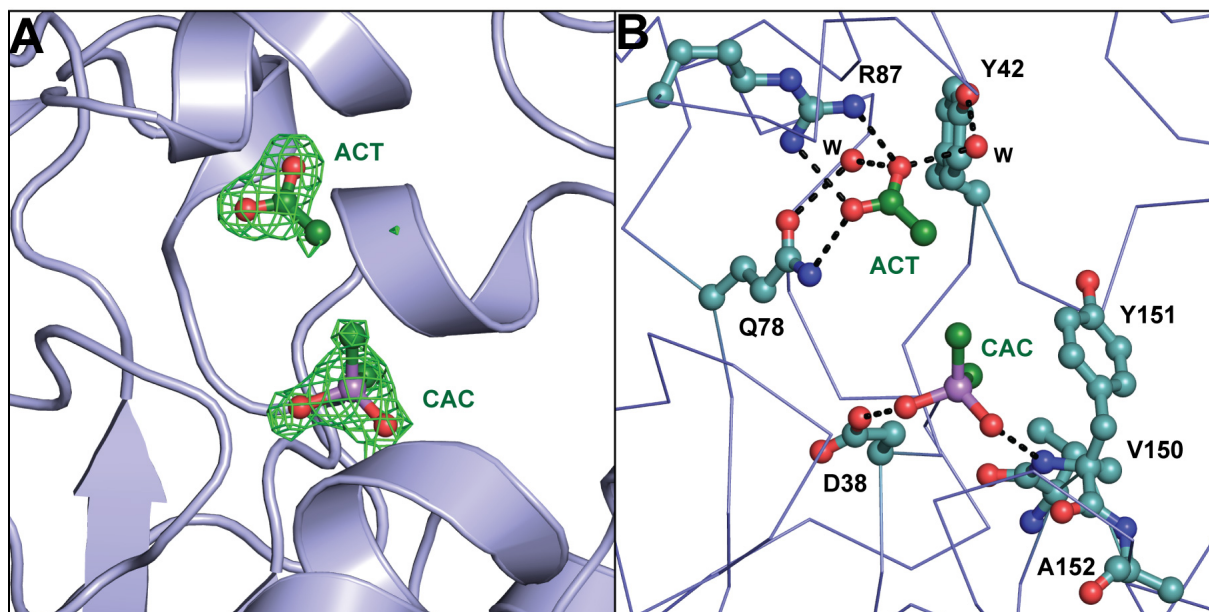
Figure 4.40: Acetate and cacodylate bound to the active site of PhzD. (A) |F$_O$F$_C$| difference electron density of acetate and cacodylate is displayed at 2.0 σ. (B) Details of the active site. Residues that are making hydrogen bonds with acetate and cacodylate are shown in ball-and-strick presentation.

4.4.6.2 Ligand-free PhzD

The main purpose of crystallizing PhzD in this study focuses on the investigation of protein/substrate binding. Since PhzD catalyzes the step after PhzE in the phenazine biosynthesis pathway, converting ADIC to DHHA, experiments were designed to visualize ADIC/DHHA in the active center of PhzD by soaking the crystals with purified PhzE product (ADIC). However, since PhzD was crystallized with non-native substrates bound to its active site (see chapter 4.4.6.1), a direct soaking experiment of PhzD crystals with ADIC was unsuccessful and resulted in ambiguous electron densities at the active site. Therefore, in order to remove acetate and cacodylate from the crystals, a "kick-out" experiment was designed, in which PhzD crystals were transferred to a new drop containing different buffer and precipitant substances (0.1 M Bis-Tris pH 6.5, 20% (w/v) PEG 4000, 0.1 M NaCl) and incubated for 60 min before soaking with ADIC.

To prove that acetate and cacodylate were successfully removed from the active center, the crystal structure of ligand-free PhzD was determined. Superimposition of

ligand-free and ligand-bound PhzD showed that the major difference between these two structures locates at the loop region P79 – R87, which becomes untraceable in the ligand-free structure. Apart from R87, which is sitting right on the loop region, Y42 and Q78, two residues involved in acetate coordination also adopt different rotamers in the ligand-free structure. However, the position of residues involved in cacodylate coordination, namely D38 and Y151, remains almost the same in both structures (Figure 4.41).
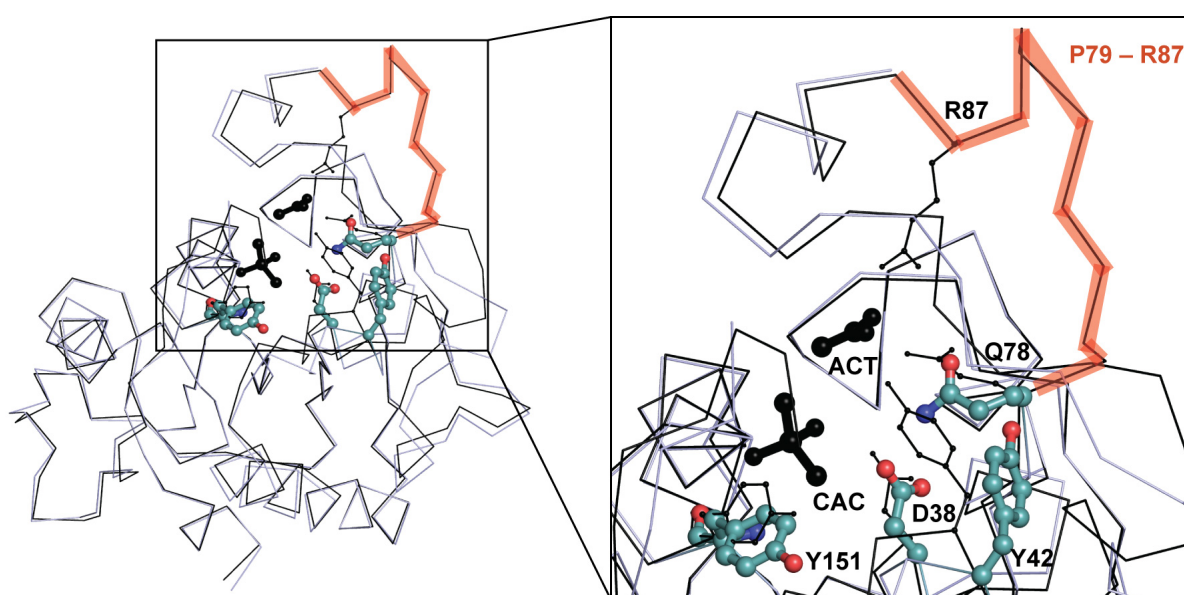

Figure 4.41: Comparison of ligand-free and ligand-bound PhzD. Ligand-bound PhzD is shown as black ribbon and residues involved in ligand binding are presented in black ball-and-stick. Acetate and cacodylate coordinates are from the ligand-bound PhzD structure. Ligand-free PhzD is shown in light-blue ribbon. Residues are colored by element in comparison to those from the ligand-bound structure. The loop region P79 – R87 from the ligand-bound form is colored in dark-orange, and is invisible in the ligand-free structure.

No additional electron density was observed at the active site of ligand-free PhzD, indicating that acetate and cacodylate have been successfully removed. The flexible loop P79 – R87 is believed to play important roles in increasing solvent exposure of the active center which in turns facilitates ligand entry/exit.

Since it was possible to prepare diffraction quality crystals in a ligand-free form, all the soaking experiments of PhzD/D38A/D38N with ADIC were performed using the ligand-free crystals.

4.4.6.3 PhzD in complex with DHHA

Ligand-free PhzD crystals were soaked with purified PhzE product for 60 min. The crystals were then flash frozen in liquid nitrogen and immediately used for data collection. Interestingly, no ADIC was observed, but clear density suggested that the product DHHA was trapped at the active center (Figure 4.42 **A**). This is an indication that although crystallized, PhzD sustained its acitivity and was able to catalyze the conversion of ADIC to DHHA.

The carboxylate group of DHHA is involved in hydrogen bonding with side chain as Q78, R87 and a water molecule which is coordinated by Y42. Due to the position of K122, it is possible that K122 could act as a general base which activates a water molecule to perform the cleavage of pyruvate from ADIC. However, previous report stated that the mutation of D38 to alanine, not K122 to alanine renders the enzyme inactive (Parsons et al. 2003). Beside these interactions, the ring structure of DHHA is also involved in stacking interaction between the side chain of W94 and F43 (Figure 4.42 **B**).
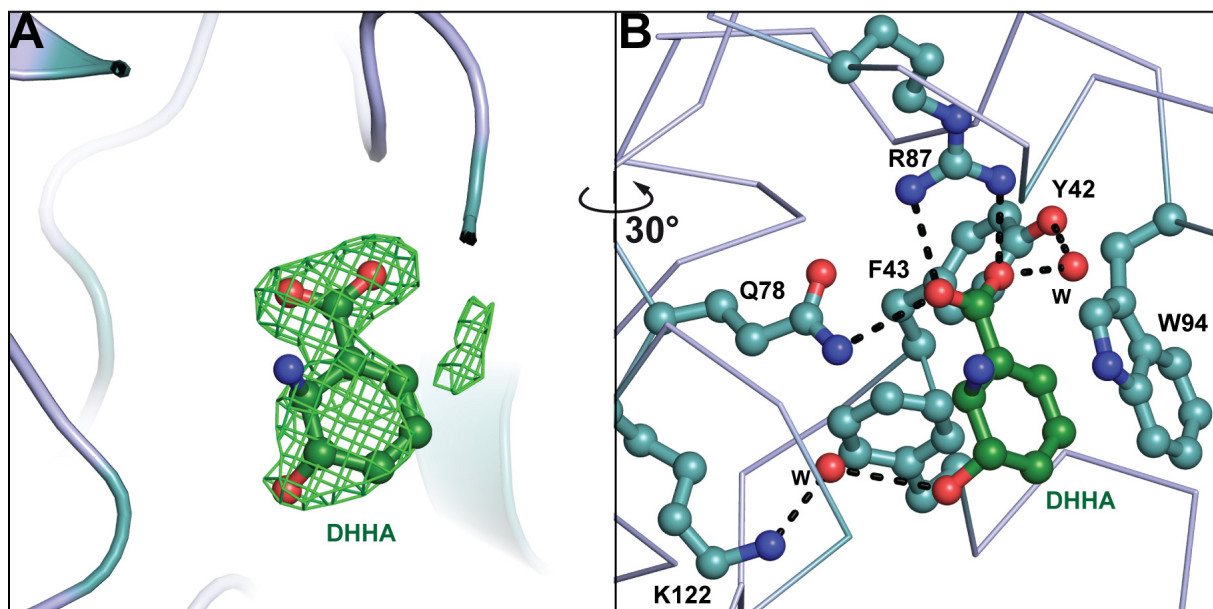

Figure 4.42: DHHA bound to the active site of PhzD. (A) |F$_O$F$_C$| difference electron density of DHHA is displayed at 2.0 σ. (B) Details of the active site. Residues that are involved in coordination of DHHA are shown in call-and-stick presentation.

Because the crystals were soaked with purified PhzE product, the observation of a possibly DHHA at the active site of PhzD may provide indirect evidence supporting that ADIC is indeed produced by PhzE and serves as substrate for PhzD. However, since the 2-amino group of DHHA is ambiguous on the electron density map (Figure 4.42 **A**), this needs to be further proved by more substantial evidence, which would be the direct observation of the substrate ADIC. In order to obtain a crystal of PhzD in complex with ADIC, two inactive PhzD mutants PhzD-D38A and PhzD-D38N were prepared for the soaking experiments.

4.4.6.4 PhzD-D38A in complex with ADIC

Since soaking experiments with PhzD-D38A and PhzD-D38N gave similar results and because the data quality from the PhzD-D38A crystal was better than the one from PhzD-D38N, the discussion here is therefore based on the PhzD-D38A structure.

The $|F_O F_C|$ electron density shows clearly that ADIC is bound to the active center of PhzD. Coordinates of acetate and cacodylate from the native PhzD structure were superposed with ADIC, showing that they are sitting in the position of the C1 carboxylate group and the carboxylate group of the C3 pyruvyl moiety, respectively (Figure 4.43. **A**). Residues involved in ADIC coordination are similar to those that coordinate DHHA. The ring structure of ADIC is held between side chain of W94 and F43, and is hydrogen bonding to Q78 and R87, as well as to three water molecules coordinated by Y42, Q78 and K122. Moreover, the C3 pyruvyl is forming hydrogen bonds with the back bone amide of Y151, G155 and the side chain of K122 (Figure 4.43 **B**).

Interestingly, the $|F_O F_C|$ electron density of the dataset collected at the synchrotron from the same crystal shows that the pyruvyl moiety of ADIC is separated from the ring due to a break of the vinyl ether bond (Figure 4.44). This is an indication that the PhzD-D38A crystal might have suffered from remarkable radiation damage that causes the vinyl ether bond to break. Radiation damage might also be responsible for the fact that only benzoate and pyruvate were observed from the ligand-bound structure of PhzE, but neither the substrate chorismate nor the product ADIC.

This is the first time that ADIC has been observed in a crystal structure and the result has provided critical evidence regarding the activity of PhzE. Together with biochemical and analytical analysis results, PhzE has been confirmed as an ADIC synthase.
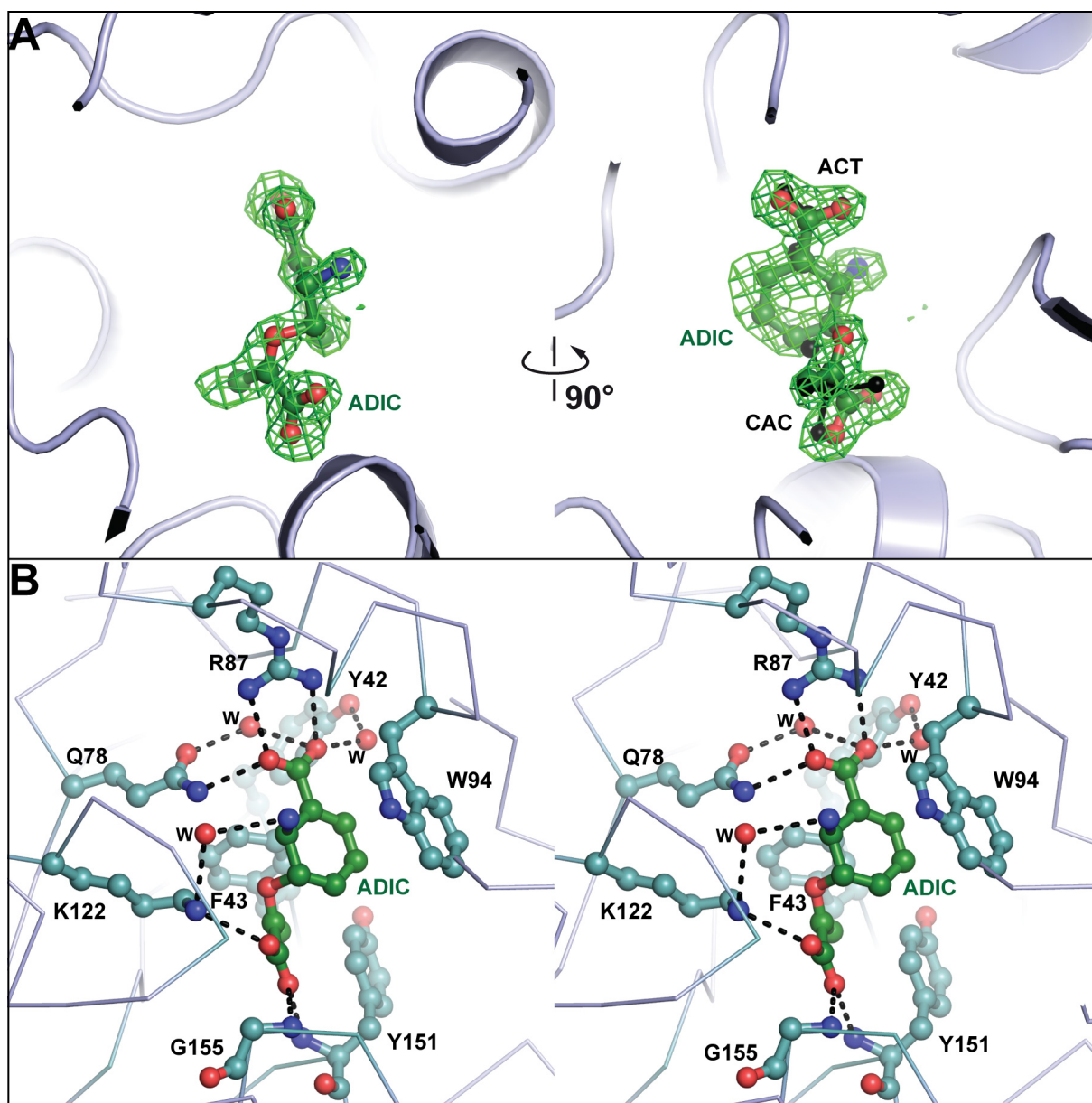
Figure 4.43 ADIC bound to the active site of D38A. (A) $|F_OF_C|$ difference electron density of ADIC is displayed at 2.5 σ. Coordinates of acetate and cacodylate (black) are taken from the native PhzD structure and superposed with ADIC. (B) Stereo-plot showing details of residues involved in ADIC coordination.
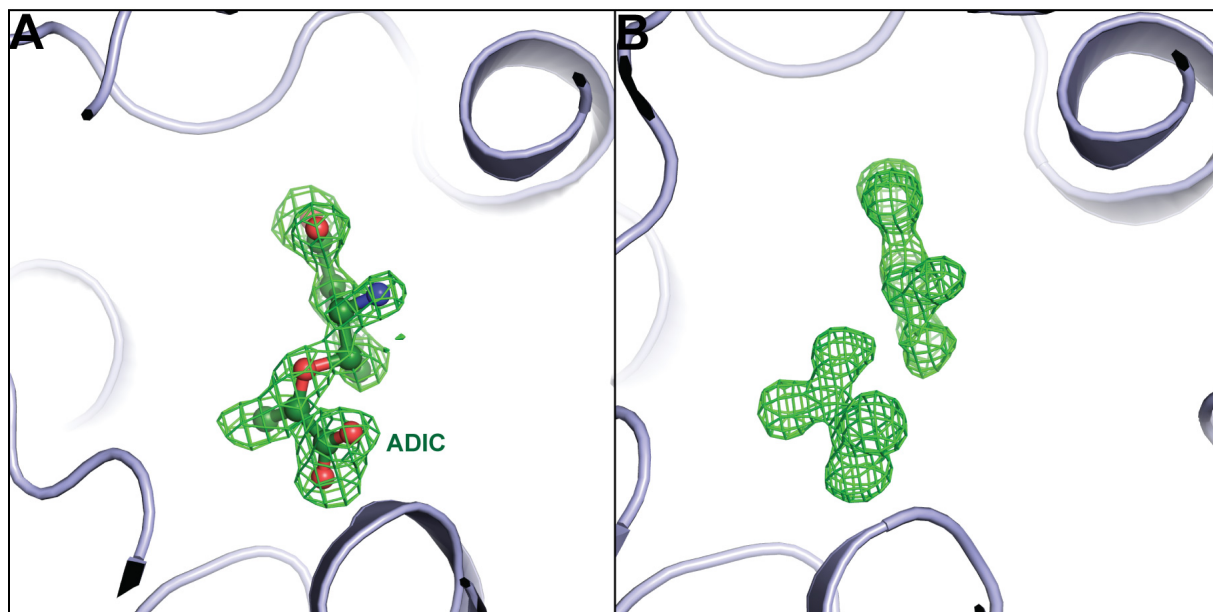
Figure 4.44: |F$_O$F$_C$| difference electron density of ADIC from (A) in-house data and (B) synchrotron data, displayed at 2.5 σ. Both datasets were collected from one same crystal.

5 **OUTLOOK**

The data presented here include two crystal structures of PhzE and provide insight into structural changes that accompany its catalytic cycle. They also give an opportunity to discuss differences that may lead to the evolution of pyruvate-eliminating and non-eliminating MST enzymes.

5.1.1 **The cause of ligand breakdown in the closed-form PhzE structure**

It is unclear why benzoate and pyruvate but not chorismate or ADIC were found in the MST active site of the closed crystal form. The breakdown of ADIC seems to be linked to the crystallization conditions since no breakdown products were observed in HPLC analysis of enzyme activity assays. X-ray radiation damage could also be a reason for the breakdown, which is supported by the observation that the vinyl ether bond of ADIC trapped at active center of PhzD-D38A was broken after the crystal has been exposed to synchrotron radiation.

Therefore, further studies concerning the cause of ligand breakdown observed in the crystal structure may focus on two aspects. The first aspect regards the crystallization conditions for growing ligand-bound PhzE crystals. In order to test if the chorismate/ADIC breakdown is caused by reagents in the crystallization conditions, crystallization trials could be carried out under the same condition but in the absence of reducing agent (in this case, 5 mM β-mercaptoethanol), as well as search for new possible crystallization conditions for ligand-bound PhzE. The second aspect is the X-ray radiation damage. Since all data of ligand-bound PhzE were collected at the synchrontron beamline, and because the radiation damage to the crystal at the synchrontron beamline is much higher than that from an in-house beamline, it would be necessary to examine if ligand breakdown retains in the data collected from an in-house beamline. However, since most of the PhzE crystals gave relatively weak diffractions in-house, difficulties remains regarding the production of high-quality ligand-bound PhzE crystals that could be used for in-house data collection. It also has to be mentioned that

due to the quality of the crystals and a relatively large cell axis in length (a = 259.7 Å, b = 94.5 Å and c = 53.5 Å, α = β = γ = 90º), data collection in-house would require longer exposure time of the cystal to the X-ray beam, from which the accumulative radiation damage to the crystal may also be distinct.

5.1.2 Understanding the differences between PhzE and AS

Although extensive mutagenetic experiments on PhzE have been performed in this study, identification of the exact residues that determine the catalytic difference between ADIC synthase and AS failed. Therefore, differences of the key residues are probably not the key factors that determine their functions. PhzE might employ a different substrate release mechanism, which avoids pyruvate elimination of ADIC by providing a different release path or shortening the ADIC residence time in the active center. In order to answer these questions, it may be necessary to extend mutagenetic approach to study residues beyond the first coordination sphere of ADIC. In addition, it would be interesting to compare the binding affinity of ADIC to AS and PhzE, which might be able to provide hints explaining their catalytic difference.

To obtain AS for the binding assay, three well-studied enzymes AS from *Serratia marcescens* (PDB ID: 1I7Q, 1I7S) (Spraggon et al., 2001), *Salmonella typhimurium* (PDB ID: 1I1Q) (Morollo & Eck, 2001), and *Sulfolobus solfataricus* (PDB ID: 1QDL) (Knochel et al. 1999) could be used for recombinant protein production. Enzymatically synthesized ADIC could be purified as described in this study (see chapter 3.6.10). Interaction between ADIC and the proteins will be measured first with ITC, and will be further extended to crystallographic experiments according to the crystallization conditions reported for all three anthranilate synthases to be used.

Apart from ADIC itself, analogues of chorismate/ADIC can be investigated for their interactions with PhzE and AS in order to reveal more details regarding ligand-binding mechanisms. Three synthetic analogues of chorismate/ADIC are proposed (Figure 5.1). Both A1 and A2 are analogues of chorismate and are supposed to be inert to the amination at C2 position. A1 reflects the very early stage of the reaction and A2 reflects the later on stage where the cleavage of C4 hydroxyl group has not been processed

128

(see Figure 4.18). A3 is an analogue of ADIC that has a stable C3 moiety against cleavage. Studying the binding properties of these analogues is expected to provide detais in understanding the reaction mechanisms of the MST enzyme family.
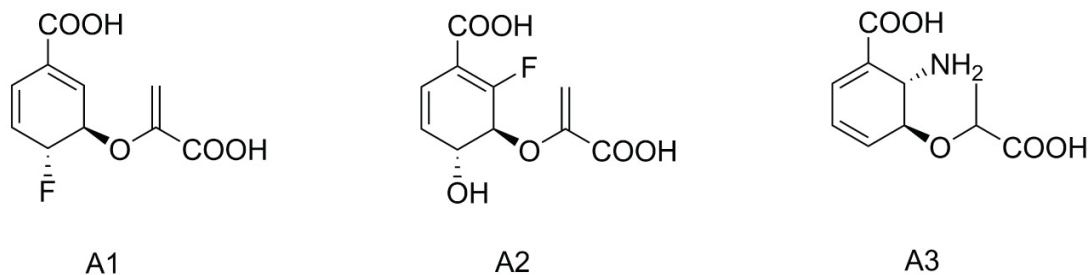

Figure 5.1: Synthetic analogues of chorismate (A1 and A2) and ADIC (A3).

5.1.3 Ligand binding of the MST: sequential or simultaneous?

Another interesting observation was the entropy-driven binding of chorismate to PhzE in the presence of Mg^{2+} by ITC measurement. Since chorismate alone doesn't show affinity to PhzE when Mg^{2+} is excluded, and because no Mg^{2+} was observed in the open PhzE structure although it has been included in the crystallization condition, it is therefore suggested that chorismate and Mg^{2+} bind simultaneously to the MST domain of PhzE. However, Mg^{2+} alone has not been tested for its binding affinity to PhzE by ITC. To further elucidate the sequential binding, in which chorismate/Mg^{2+} bind first to the MST domain that enables glutamine binding to the GATase1 domain, it would be necessary to conduct ITC measurements to confirm the binding of Mg^{2+} and glutamine to PhzE in the absence of chorismate, respectively.

5.1.4 Inhibition of PhzE by divalent transition-metal ions

Apart from the novel quaternary structure of PhzE, another intriguing result presented in this study was the inhibition of PhzE by Zn^{2+}, Mn^{2+} and Ni^{2+}. The observation of a member of the MST enzyme family being inhibited by transition metal cations is unprecedented. Together with crystallographic evidence that Zn^{2+} was trapped at the active center of the GATase1 domain, it is suggested that the reduced overall activity of PhzE by di-valent metal cations is due to the inhibition of GATase1. Clearly, this aspect will have to be further investigated since the concentration of the

metal cation applied in the preliminary inhibition assay (1 mM) was much higher than that found under *in vivo* conditions. In this case, a full inhibition assay with various amounts of divalent metal would be necessary and whether other metal ions could affect PhzE activity is also a question to be answered.

5.1.5 **Further investigations regarding PhzD**

The reaction product DHHA observed at the active site of PhzD was a result from soaking the crystals with the substrate ADIC, indicating that PhzD is still catalytically active in the crystal. However, the C2 amine group is ambiguous on the electron density map (Figure 4.42 **A**). To improve this, further soaking experiments should be carried out with respect to different soaking conditions (the concentration of ADIC and the length of incubation). Data with better quality to display the C2 amine of DHHA is expected.

The vinyl ether bond of ADIC was broken as observed in the synchrotron dataset but not the in-house one from the same crystal indicates that synchrotron radiation damage are very likely responsible for the observation of only break down product of ADIC (benzoate and pyruvate) in the closed PhzE structure. This emphasized the importance of analyzing in-house data of ligand-bound PhzE crystals.

6 SUMMARY (ZUSAMMENFASSUNG)

Summary

A relatively large number of bacteria strains produce and secrete a group of nitrogen-containing aromatic metabolites named phenazines. In early studies, it was believed that due to its redox-activity, phenazines are solely used by their producers as redox-active antibiotics in microbial competitiveness. However, it has been recently recognized that these compounds have diverse physiological functions. They can act as signalling molecules and respiratory pigments under anoxic conditions e.g. in the deeper anoxic layers of biofilms, controlling size and shape of bacteria colonies. For example, in the phenazine deficient *Pseudomonas* colonies, bacteria tend to spread out to get more direct contact with oxygen. Since the biosynthsis of phenazines is almost exlusively limited to eubacteria, enzymes involved in this process are potentially attractive targets for pharmaceutical intervention.

Chorismate is the precursor for a large number of primary and secondary metabolites in bacteria and plants including e.g. the aromatic amino acids, ubiquinone, folate (vitamin B_9), menaquinone (vitamin K), the siderophores enterobactin, pyochelin and phenazines. PhzE utilizes chorismate and glutamine to synthesize 2-amino-2-deoxyisochorismate (ADIC) in the first step of phenazine biosynthesis, followed by the step in which ADIC is converted to *trans*-2,3-dihydro-3-hydroxyanthranilic acid (DHHA) by an isochorismate pyruvate hydrolase PhzD.

In this work, the crystal structure of both PhzE and PhzD in their ligand-free and ligands/substrates-bound form were determined. While focused on investigating the biological function and potential regulation mechanism of PhzE by a combination of biochemical and biophysical techniques, novel results regarding structure and substrate-binding mechanism of PhzD are also presented. The new data on PhzE have extended the spectrum of structurally characterized chorismate utilizing enzymes to ADIC synthase. Furthermore, the structural and functional study of these two proteins

provides new insight into the initial steps of phenazine biosynthesis which may guide future investigation in searching for potential regulatory mechanisms of this pathway.

Structure and function of PhzE

PhzE utilizes chorismate and glutamine to synthesize 2-amino-2-desoxyisochorismate (ADIC) in the first step of phenazine biosynthesis. PhzE's monomer contains both a chorismate-converting menaquinone, siderophore, tryptophan biosynthesis (MST) and a type 1 glutamine amidotransferase (GATase1) domain connected by a 45-residue linker. In this study, the crystal structure of PhzE from *Burkholderia lata* 383 in a ligand-free open and ligand-bound closed conformation at 2.9 and 2.1 Å resolution are presented respectively.

PhzE arranges in an intertwined dimer such that the GATase1 domain of one chain provides NH_3 to the MST domain of the other. This quaternary structure was confirmed by small angle x-ray scattering. Binding of chorismic acid, which was found converted to benzoate and pyruvate in the MST active centers of the closed form, leads to structural rearrangements that establish an ammonia transport channel approx. 25 Å in length within each of the two MST/GATase1 functional units of the dimer. The assignment of PhzE as an ADIC synthase was confirmed by mass-spectrometric analysis of the product, which was also visualized at 1.9 Å resolution by trapping in crystals of inactive mutants (D38A/D38N) of PhzD, an isochorismatase that catalyzes the subsequent step in phenazine biosynthesis.

Michaelis–Menten kinetic parameters of PhzE regarding chorismate have been determined, indicating comparable activity of PhzE to other members of the MST enzyme family. Mutagenesis experiments were performed to investigate residues either involved in maintaining the structural/functional integrity of PhzE, or those that differ between PhzE and anthranilate synthase (AS) within the first coordination sphere of the substrate chorismate. Unlike in some of the related anthranilate synthases, no allosteric inhibition was observed in PhzE. This can be attributed to a tryptophan residue of the protein blocking the potential regulatory site. Additional electron density in the GATase1

active center was identified as zinc and it was demonstrated that Zn^{2+}, Mn^{2+} and Ni^{2+} reduce the activity of PhzE.

Crystal structures of PhzD and implications of substrate-binding mechanism

PhzD is an isochorismatase which catalyzes the second step in phenazine biosynthesis, hydrolyzing the vinyl ether group of ADIC to produce *trans*-2,3-dihydro-3-hydroxyanthranilic acid (DHHA) and pyruvate. It is composed of a single, 209 amino-acid peptide chain which folds into an α/β structure that resemble members of the α/β-hydrolase family.

In this study, a thourough crystallographic study has been conducted. Crystal structures of native PhzD in its ligand-free form and in complex with cacodylate and acetate have been determined, as well as of PhzD in complex with its reaction product DHHA. Both of the inactive mutants PhzD-D38A and PhzD-D38N in complex with ADIC were crystallized too. This is the first time that both the native substrate and product of PhzD have been observed in crystal structures. The results not only confirmed the ADIC synthase function of PhzE, but also provided insight into substrate binding mechanism of PhzD. A number of residues involved in substrate coordination and reaction catalysis have been identified. More interestingly, the loop region P79 – R87 in the ligand-free structure is of high flexibility and completely invisible on the electron density map. Structural comparison between ligand-bound and ligand-free structures suggests that this loop may act as a gatekeeper of the active site, controlling the entrance/exit of the substrate.

Zusammenfassung

Eine relativ große Zahl von Bakterienstämmen produziert und sezerniert Derivate von stickstoffhaltigen aromatischen Metaboliten der Phenazingruppe. Nach den Ergebnissen früherer Untersuchungen wurde angenommen, dass Phenazin-produzenten diese Verbindungen lediglich als redoxaktive Antibiotika zur Steigerung ihrer Kompetitionsfähigkeit einsetzen. Neuere Studien zeigen jedoch, dass Phenazine noch andere physiologische Funktionen besitzen. So können sie als Signalmoleküle wirken und unter anoxischen Bedingungen, z.B. in den tieferen Schichten von Biofilmen, als respiratorische Pigmente eingesetzt werden, wodurch die Größe und Form von Bakterienkolonien beeinflusst wird. So tendieren die Bakterien in phenazindefizienten *Pseudomonas*-Kolonien dazu, sich stärker auszubreiten, um direkten Kontakt mit Luftsauerstoff zu erhalten. Weil die Phenazinbiosynthese fast ausschließlich auf Eubakterien beschränkt ist, sind die in diesen Prozess involvierten Enzyme potentiell attraktive Angriffspunkte für die pharmakologische Intervention.

Chorismat ist das Vorläufermolekül für eine große Zahl von primären und sekundären Metaboliten in Bakterien und Pflanzen, z.B. für die aromatischen Aminosäuren, für Ubichinon, Folat (Vitamin B_9), Menachinon (Vitamin K), die Siderophore Enterobactin und Pyochelin sowie für die Phenazine. PhzE verwendet Chorismat und Glutamin zur Erzeugung von 2-Amino-2-desoxyisochorismat (ADIC) im ersten Schritt der Phenazinbiosynthese. Im folgenden Schritt wird ADIC dann von der Isochorismat-Pyruvat-Lyase PhzD in *trans*-2,3-Dihydro-3-hydroxyanthranilat (DHHA) umgewandelt.

In dieser Arbeit wurden die Kristallstrukturen von PhzE und PhzD sowohl in der freien als auch in der liganden- bzw. subtratgebundenen Form bestimmt. Während der Fokus in der Untersuchung der biologischen Funktion und von potentiellen regulatorischen Mechanismen von PhzE mittels einer Kombination aus biochemischen und biophysikalischen Methoden bestand, wurden auch neue Einblicke in den Substratbindungsmechanismus von PhzD gewonnen. Die erhaltenen Daten erweitern das Spektrum der strukturell charakterisierten chorismat-verwendenden Enzyme um

134

ADIC-Synthasen. Außerdem vermitteln die strukturellen und funktionellen Untersuchungen an diesen beiden Proteinen neue Einblicke in die ersten Schritte der Phenazinbiosynthese und könnten so zukünftige Untersuchungen an möglichen Regulationsmechanismen dieses Weges leiten.

Struktur und Funktion von PhzE

PhzE verwendet Chorismat und Glutamin, um im ersten Schritt der Phenazinbiosynthese 2-Amino-2-desoxyisochorismat (ADIC) zu erzeugen. Das Monomer von PhzE enthält sowohl eine chorismat-umwandelnde Menachinon-Siderophor-Tryptophan-Biosynthese- (MST) als auch eine Typ-1-Glutamin-Amidotransferase-Domäne (GATase 1), welche durch einen 45 Aminosäure langen Linker miteinander verbunden sind. In dieser Arbeit werden die Kristallstrukturen von PhzE aus *Burkholderia lata* 383 in einer ligandenfreien offenen und in einer ligandengebundenen geschlossenen Konformation bei 2.9 bzw. 2.1 Å Auflösung vorgestellt.

PhzE formt ein überkreuztes Dimer, in dem die GATase-1-Domäne der einen Kette NH_3 für die MST-Domäne der anderen Kette erzeugt. Diese Quartärstruktur wurde auch durch Kleinwinkel-Röntgenstreuungsexperimente bestätigt. Die Bindung von Chorismat, das in der Kristallstruktur in Benzoat und Pyruvat umgewandelt vorgefunden wurde, führt zu strukturellen Umlagerungen unter Ausbildung eines Ammoniak-Transportkanals von 25 Å Länge innerhalb jedes der beiden funktionalen MST/GATase1-Paare des Dimers. Die funktionelle Einordnung von PhzE als ADIC-Synthase wurde durch massenspektrometrische Analyse des Produkts bestätigt. Dieses konnte auch durch Abfangen in Kristallen von inaktiven Mutanten (D38A/D38N) von PhzD, einer Isochorismatase, die den nachgelagerten Schritt in der Phenazinbiosynthese katalysiert, bei 1.9 Å sichtbar gemacht werden.

Enzymkinetische Michaelis-Menten-Parameter für den Umsatz von Chorismat durch PhzE wurden bestimmt und zeigen, dass PhzE ein ähnliches Aktivitätsniveau wie andere Mitglieder der MST-Familie besitzt. Mutageneseexperimente wurden durchgeführt, um einerseits Aminosäuren, die an der Aufrechterhaltung der Stabilität bzw. Funktion von PhzE beteiligt sind, zu identifizieren und um anderseits Unterschiede

in der ersten Substratkoordinationssphäre von Chorismat zwischen PhzE und Anthanilatsynthasen (AS) zu untersuchen. Im Gegensatz zu den verwandten Anthranilatsynthasen wurde bei PhzE keine allosterische Inhibition festgestellt. Dies ist darauf zurückzuführen, dass ein Tryptophan die potentielle regulatorische Bindungs-stelle blockiert. Zusätzliche Elektronendichte im aktiven Zentrum der GATase1 wurde als Zink identifiziert und es wurde gezeigt, dass Zn^{2+}, Mn^{2+} und Ni^{2+} die Aktivität von PhzE verringern.

Kristallstrukturen von PhzD und Implikationen für den Substratbindungs-mechanismus

PhzD ist eine Isochorismatase, die den zweiten Schritt der Phenazinbiosynthese katalysiert, indem sie die Vinylethergruppe von ADIC unter Erzeugung von *trans*-2,3-Dihydro-3-hydroxyanthranilat (DHHA) und Pyruvat hydrolysiert. Das Enzym besteht aus einer einzigen Kette mit 209 Aminosäuren, die zu einer α/β-Struktur mit Ähnlichkeit zu Mitgliedern der α/β-Hydrolasefamilie faltet.

Im Rahmen dieser Arbeit wurde eine umfassende kristallografische Untersuchung durchgeführt. Kristallstrukturen von nativem PhzD in seiner ligandenfreien Form und im Komplex mit Cacodylat und Acetat sowie mit seinem Produkt DHHA wurden bestimmt. Die beiden inaktiven Mutanten PhzD-D38A und PhzD-D38N im Komplex mit ADIC wurden ebenfalls kristallisiert. Dies ist das erste Mal, dass sowohl das natürliche Substrat und Produkt von PhzD in Kristallstrukturen beobachtet wurden. Die Ergebnisse bestätigen nicht nur die ADIC-Synthasefunktion von PhzE, sondern liefern auch Einblick in die Substratbindemechanismen von PhzD. Eine Gruppe von Aminosäuren, die in die Koordination des Substrats und in die Reaktionskatalyse verwickelt sind, wurde identifiziert. Interessanterweise ist die Schleifenregion P79 – R87 in der ligandenfreien Struktur sehr flexibel und in der Elektronendichtekarte vollständig unsichtbar. Der Strukturvergleich zwischen ligandenbesetzter und freier Struktur zeigt an, dass diese Schleife als ein Torwächter fungiert, der den Zu- und Abgang des Substrats kontrolliert.

7 APPENDICES

7.1 Symbols and abbreviations

7.1.1 Symbols

Amino acids

Amino acids	One letter code
Alanine	A
Cysteine	C
Aspartic acid	D
Glutamic acid	E
Phenylalanine	F
glycine	G
Histidine	H
Isoleucine	I
Lysine	K
Leucine	L
Methionine	M
Asparagine	N
Proline	P
Glutamine	Q
Arginine	R
Serine	S
Threonine	T
Valine	V
Tryptophan	W
Tyrosine	Y

Units and scales

Symbol		**Symbol**	
Å	angstrom	Da	dalton
ε	molar extinction coefficient	kDa	kilodalton
Ω	Standard unit of electrical resistance	h	hour
º	degree (angle)	min	minute
ºC	degree centigrade	s	second
g	gravitational constant	m	Meter (length)
rpm	rounds per minute	cm	centimeter
L	liter	µm	micrometer
mL	milliliter	nm	nanometer
µL	microliter	V	voltage
M	molar	kV	kilovolt
mM	millimolar	µF	microfarad (capacitance)
µM	micromolar		

7.1.2 **Abbreviations**

3HBA	3-hydroxybenzoate
β-me	beta-mercaptoethanol
ACT	acetate
ADC	4-amino-4-deoxychorismate
ADCS	4-amino-4-deoxychorismate synthase
ADIC	2-amino-2-deoxyisochorismate,
ADICS	2-amino-2-deoxyisochorismate synthase
AS	anthranilate synthase
ATP	adenosine-5'-triphosphate
BSA	bovine serum albumin
CAC	cacodylate

138

cAMP	cyclic adenosine monophosphate
CH	chorismate hydrolase
CL	chorismate lyase
CM	chorismate mutase
DAHP	3-deoxy-D-arabino-heptulosonate-7-phosphate
DCDC	4,5-dihydroxycyclohexa-1,5-dienecarboxylic acid
DHHA	*trans*-2,3-dihydro-3-hydroxyanthranilic acid
DNA	deoxyribonucleic acid
E4P	erythrose-4-phosphate
EDTA	ethylenediaminetetraacetic acid
EPSPS	5-enolpyruvylshikimate-3-phosphate synthase
ESI-MS	Electrospray Ionisation Mass Spectrometry
GATase1	glutamine amidotransferase type 1
GSH	glutathione
GSSG	glutathione disulfide
GR	glutathione reductase
HPLC	high-performance liquid chromatography
HS	3-hydroxybenzoate synthase
IC	isochorismate
ICS	isochorismate synthase
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	Luria Broth
MAD	Multi-wavelength Anomalous Dispersion
MALDI-TOF	Matrix Assisted Laser Desorption Ionization – Time of Flight
MM	Molecular Mass
MST	Menaquinone, Siderophore and Tryptophan
Mw	Molecular weight
NADH	nicotinamide adenine dinucleotide

NADPH	nicotinamide adenine dinucleotide phosphate
Ni-NTA	nickel-nitrilotriacetic acid
OD_{600}	optical density at 600 nm
PAB	*para*-aminobenzoate
PCA	phenazine-1-carboxylic acid
PCN	phenazine-1-carboxamide
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDC	phenazine-1,6-dicarboxylic acid
PEP	phosphoenolpyruvate
PHB	*para*-hydroxybenzoate
PPP	pentose phosphate pathway
PYO	pyocyanin
RMSD	root mean square deviation
RNA	ribonucleic acid
ROS	reactive oxygen species
RP-HPLC	reverse phase high-performance liquid chromatography
RT	retention time
SAXS	Small Angel X-ray Scattering
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
Se-Met	selenomethionine
SOD	superoxide dismutase
SS	salicylate synthase
TB	Terrific Broth
TEV	tobacco etch virus
TFA	trifluoroacetic acid
TLC	thin layer chromatography

7.2 *In-vivo* production and purification of chorismate

Large amounts of chorismate have been used throughout this study. Since the commercially available chorismate is typically impure and rather expensive, chorsimate used here was obtained via *in vivo* synthesis and a rapid purification method in lab based on previous report (Grisostomi et al., 1997).

7.2.1 *In vivo* synthesis of chorismate

1 L growth medium A (Table 7.1) comtaining 0.16% (w/v) glucose was inoculated with 10 mL overnight culture of *E.coli* KA12 strain (chorismate accumulating, see chapter 3.1.3).

	Composite	Amount
Growth medium A (1 L)		
	Casamino acids	2 g
	Yeast extract	2 g
	L- tryptophan	41 mg
	50× Vogel & Bonner salts	20 mL
50× Vogel & Bonner salts (1 L)	H_2O (at 45 °C)	670 mL
	$Mg_2SO_4 \cdot 7H_2O$	10 g
	Citric acid monohydrate	100 g
	K_2HPO_4	500 g
	$NaH(NH_4)PO4 \cdot 4H_2O$	175 g

Table 7.1: Composition of growth medium A and Vogel & Bonner salts.

The culture was then incubated at 30 °C (220 rpm) till OD_{600} reached 1.9 to 2.1. The cells were centrifuged for 20 min at 2000g (4 °C), and resuspended in 1 L nonsterile

accumulation medium B (Table 7.2). The suspension was incubated at 30 ºC overnight (> 15 h, 220 rpm) to allow chorismate accumulation.

Composite	Amount
Na_2HPO_4	12.8 g
KH_2PO_4	1.36 g
Glucose monohydrate	19.8 g
NH_4Cl	2.7 g
$MgCl_2 \bullet 6H_2O$	20.3 mg
L-trypotophan	2 mg

Table 7.2: Composition of accumulation medium B.

The supernatant was then collected by centrifuging the cells for 20 min at 2000*g* (4 ºC), adjusted to pH 9.0 using 10 M NaOH, flash-frozen with liquid nitrogen and stored at -80 ºC for further purification.

7.2.2 **Purification of chorismate**

7.2.2.1 *Ion-exchange chromatography*

The supernatant from the accumulation culture of *E.coli* was loaded on an ion-exchange column (BioRad Dowex 1×8, 200 – 400 mesh) coupled to a Waters purification system (see chapter 6.1.7). After loading, the column was first washed with 100 mL H_2O (flowrate 10 mL/min). Chorismate was then eluted with 120 mL (flowrate 2 mL/min) 1 M NH_4Cl (pH 8.5), and was collected in 60× 2 mL fractions. The chromatogram of elution was recorded at 274 nm and 225 nm (Figure 7.1) and the absorption at 274 nm of selected fractions at 1:200 dilutions in H_2O was measured with a spectrophotometer (Table 7.3).

According to the spectrum and measured absorption, chorismate-containing fractions (5 to 27) were pooled and acidified with 25% HCl to a final pH of 1.5, and extracted with dichloromethane (3 × 75 mL) to remove phenylpyruvate. The aqueous phase containing chorismate was then extracted with ethylacetate (4 × 50 mL) and the

combined ethylacetate extracts were washed with saturated NaCl (Brine) solution (Ethylacetate extracts:Brine = 1:1 – 1:2) and then dried over Na_2SO_4. Removal of the solvent using a rotary evaporator at room temperature yield a oily yellow product, which was stored at -80 ºC for further purification.
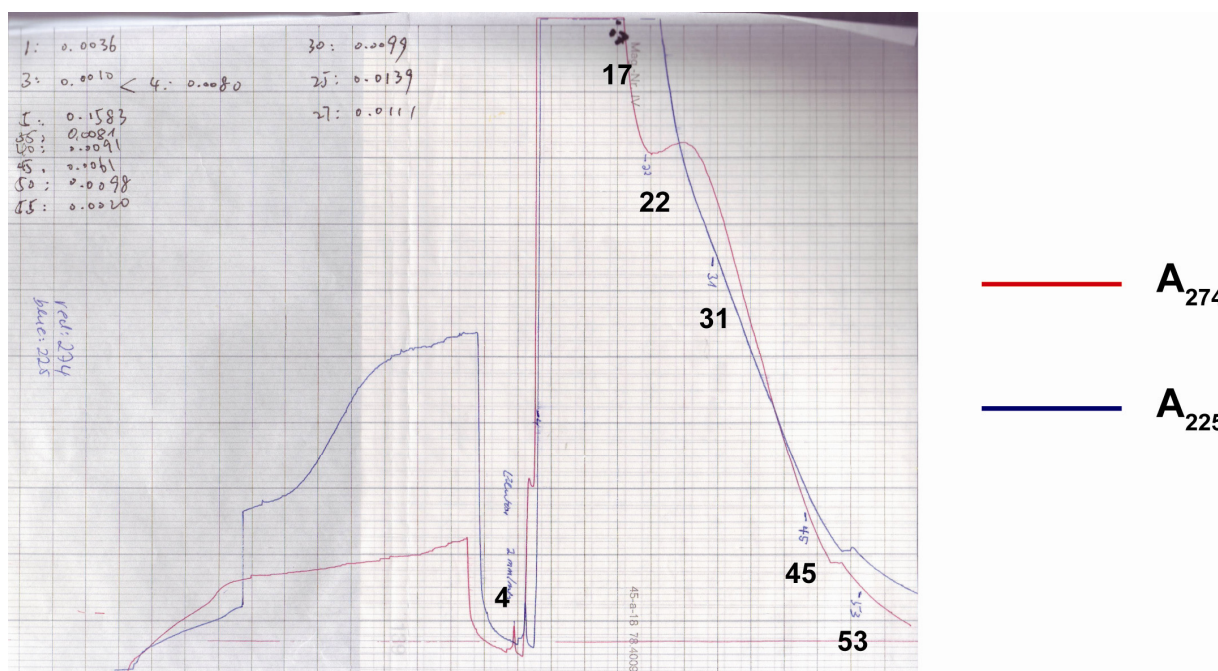

Figure 7.1: Chromatogram of the ion-exchange purification of chorismate. Fraction numbers are marked. Absorptions were measured at 274 nm and 225 nm and are shown in red and blue curves, respectively.

Fraction	A_{274} (1:200 dilutions)
1	0.0036
3	0.0010
4	0.0080
5	0.1583
25	0.0139
27	0.0111

30	0.0099
35	0.0081
40	0.0091
45	0.0061
50	0.0098
55	0.0020

Table 7.3: A_{274} of fractions eluted from ion-exchange column.

7.2.2.2 *Reverse phase flash chromatography*

Crude chorismate product was further purified by a single step reverse phase flash chromatography. The crude material was dissolved with 5 mL buffer (10 mM ammonium acetate, pH 6.8), loaded on a C18 reverse phase column (4.5 × 25 cm bed size) packed with Dowex(R) 1×8 octadecyl-functionalized silica gel (Sigma-Alderich, Deisenhofen DE), and was eluted with the same buffer under pressure. 60× 10 mL fraction were collected by hand and incubated on ice. Fractions were checked by thin layer chromatography on C18 reverse phase silica gel plates and those containing chorismate (20 to 52) were pooled (Figure 7.2).
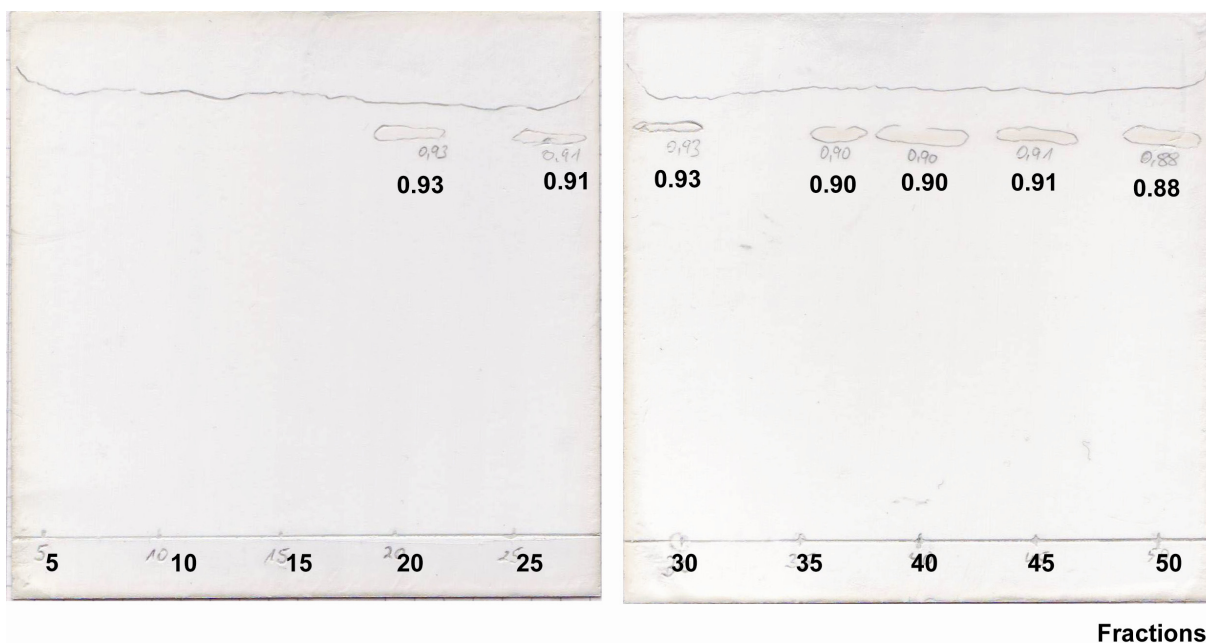
Fractions

Figure 7.2: TLC of fractions from flash chromatography. 10 µL products from each fraction were spotted on the plate. Fractions containing chorismate were measured for their R_f values. (R_f = 0.88 – 0.93, reference R_f = 0.83 for chorismate in 1 M ammonium acetate).

The pooled solution was then lyophilized to remove ammonium acetate and the resulting product of chorismate in powder (161 mg) was collected and stored at -80 ºC.

7.2.3 Quality control of self-produced chorismate by RP-HPLC

The quality of self-prepared chorismate was checked by comparing the HPLC-spectrum to that of commercially available product. Both self-prepared and commercial chorismate were dissolved with 50 mM Tris-HCl buffer (pH 7.5) to a final chorismate concentration of 1 mM. 10 µL of each samples was injected and checked by RP-HPLC as described in chapter 3.4.4 (Figure 7.3).
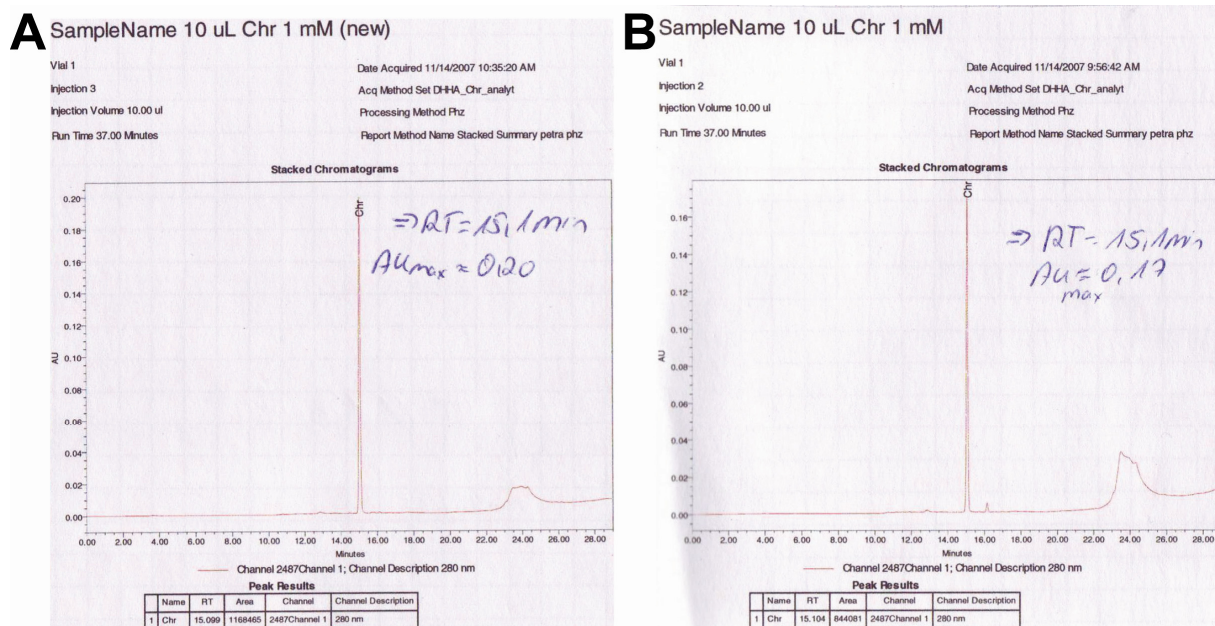
Figure 7.3: RP-HPLC of chorismate from different sources. (A) Self-prepared. (B) Commercial product.

The spectra show that while the self-prepared chorismate has exactly the same retention time (15.1 min) as that of the commercial product, the purity is even higher. Therefore, the self-prepared chorismate was used throughout this study instead of the commercially available product.

7.3 **Introduction to Protein crystallography**

7.3.1 **Viewing microscopic objects**

When viewing an object, a lens is used to collect light diffracted by the object placed just beyond the focus of the lens. A reverse-image is reconstructed beyond the focus of the lens on the opposite side (Figure 7.4). In case of a simply lens, the relationship between the image position and the object position can be difined using equation (1):

$$(OF)(IF') = (FL)(F'L) \tag{1}$$

Because *FL* and *F'L* are constants for a fixed lens (although not necessarily), there is an inverse proportion between distances *OF* and *IF'*. This could be explained as if an object is put closely beyond the focus *F* of a lens, an inversed magnified image would

be generated at a considerable distance beyond *F'*, enabling a convenient viewing for observers.
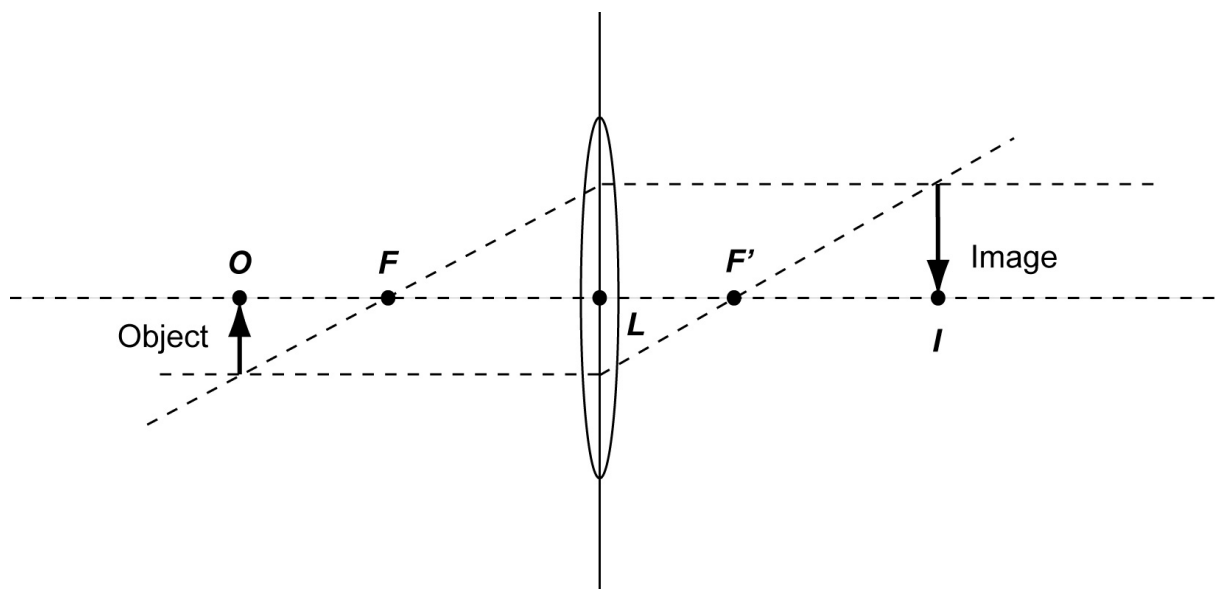

Figure 7.4: Viewing object by a simple lens. *L*: the lens, *F* and *F'*: position of the focus, *O*: position of the object, *I*: position of the image.

However, the size of the object that can be examined from a lens (termed as resolution) is limited by two indispensable factors: the wavelength of the light and the property of the lens. As Ernst Abbe first described in 1873, the resolution of a microscope is defined by the wavelength of the light (λ), refractive index of the lens medium (n), and the aperture half angle (α), as shown in equation (2):

$$R = \frac{\lambda}{2n \sin \alpha} \qquad (2)$$

As determined by equation (2), the wavelength of the light used for observation must not be larger than two times the scale of the object. In protein crystallography, researchers are expecting to examine individual atoms in protein molecules, in which bond atoms are only about 1.5 Å (0.15 nm) apart. Therefore, visible light that has wavelengths of 400-700 nm cannot be used to observed details of a protein structure. The fact that bond-length of 1.5 Å falls typically in the range of X-rays (0.1-100 Å) makes X-rays an ideal light source to reveal details of bio-macromolecular structures. The most commonly used source of X-rays for protein crystallography is 1.54 Å Kα-Cu

X-ray, which is emitted when an L-shell electron of a Copper atom replaces a displaced K-shell electron. Besides, modern synchrotron radiation generated by particle acceleration provides fine-tunable X-rays at the wavelength around 1 Å with much greater magnitude and collimation than those generated with X-ray tubes, making it possible to examine sub-angstrom details of objects.

Although the problem with wavelength could be solved using X-ray instead of visible light, other key factors of viewing molecular structures, which are decided by the property of the lens and the objects, remain unsolved. First, it is impossible to focus X-rays with a physical lens. Therefore, computer has to be introduced to simulate the lens and calculate the image of the object using measured intensities and directions of the diffracted X-rays. Second, a single molecule is not strong enough in diffracting X-rays. In order to solve this problem, a protein crystal that is composed of well-ordered, identically-oriented arrays of protein molecules is used. Since the diffracted beams from those identical molecules in the crystal interfere, the intensity of diffractions is enhanced and become detectable.

Therefore, to conduct protein crystallography experiments, it is important for researchers to generate good-quality protein crystals and to collect and interpret diffraction data of the crystals. The basic concepts will be discussed briefly in the following chapters.

7.3.2 **Growing protein crystals**

The most commonly used method for growing protein crystals is called vapor diffusion method. This is done by mixing the purified protein solution with prepared reservoir solution containing appropriate amount of buffer and precipitants in a drop, this drop of mixture is then equilibrated against the reservoir solution in an air-tight system, usually a 24-well plate sealed with silicon gel and cover slips. For example, in the hanging drop method, protein droplets are spotted and hanging on a covers slip when sealed (Figure 7.5 **B**). Due to the vapor diffusion, water is transferred gradually from the drops that contain lower concentration of precipitants, to the reservoir solution that contain higher concentration of precipitants. During this process, the concentration of

the protein and the precipitants in the drop will increase, and hopefully to the supersaturated phase that nucleation (small crystals) may form. Because of nucleation, the concentration of protein will drop slightly to reach a metastable zone, in which nucleation will stop but the protein crystals are able to stay growing (Figure 7.5 **A**).
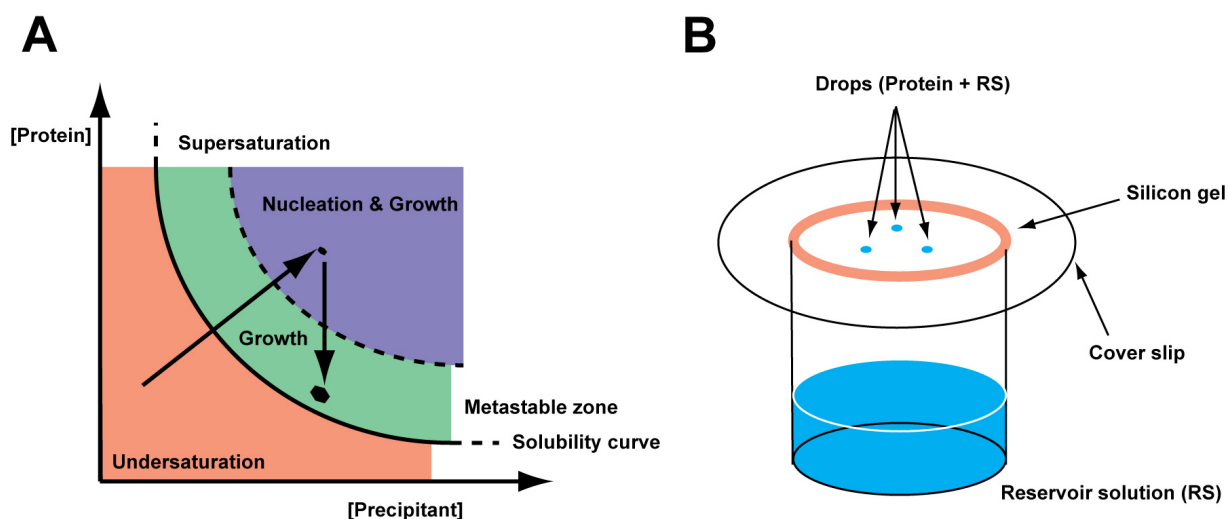

Figure 7.5: (A) Phase diagram of crystallization controlled by precipitant and vapor diffusion. (B) Sketch of crystallization set-up by the hanging-drop vapor diffusion method.

Because the exact mechanism of crystal formation is not yet clear, crystallographers usually need to test a number of precipitants under various conditions (concentration, pH, temperature, etc.) in order to obtain high-quality crystals. Therefore, at the initial stage of crystallization trials, robot-assissted high-throughput screenings are usually applied to identify the proper condition for crystallizing a protein.

7.3.3 **Collecting diffraction data and generating electron density**

7.3.3.1 *The real space and the reciprocal space*

If the shape of the protein molecules was reduced into a spot in the space, the crystal of a protein can be reduced to sets of equivalent, parallel planes of spots. X-rays shinned on the crystal are scattered and the scattering waves can interfere constructively when certain prerequisite of the crystal lattice is satisfied (Figure 7.6). The relationship between the constructive interference of scattering waves and the lattice plane of the crystal is defined as the Bragg's law (equation 4).
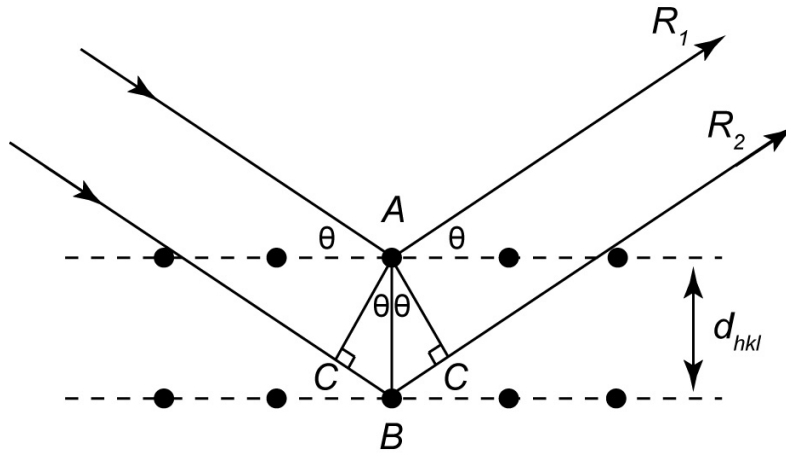
Figure 7.6: Scattering of X-rays by real space crystal lattice.

According to figure 7.6, the difference of traveling distance between R2 and R1 equals two times the length of *BC*, and can be calculated as follows:

$$2BC = 2AB \sin \theta = 2d_{hkl} \sin \theta \tag{3}$$

If the additional distance traveled by R2 is equal to an intergral number of wavelengths (nλ), the interference of the diffracted waves is constructive.

$$2d_{hkl} \sin \theta = n\lambda \tag{4}$$

Parameters of the real space lattice (unit cells) have to be derived from the interferred reflections that can be detected. In this case, in order to simplify calculations, a system called reciprocal space is introduced. In the lattice of reciprocal space, the points are actually locations of all the Bragg reflections from the real space scattering. It is named reciprocal space because the distances between the new lattice points are reciprocal to that of the real space lattice points. Bragg's law is also applied in this case and can be extended to the three-dimensional situation, and with the knowledge of unit cell type and the reciprocal lattice, the real space lattice parameters can be calculated.

7.3.3.2 *Reflections and electron density*

The mathematical relationship between an object (in this case the electron distribution in the crystal) and its diffraction pattern (reflections recorded) can be precisely described by Fourier transform. Crystals are three-dimensional repetition of

small unit cells, therefore a reflection is described by a sum of structure factors, crystallography uses the Fourier transform to convert the structure factors to the desired electron density equation ρ(x, y, z).

Any complicate wave can be described as the sum of a series of simple waves. The sum is called a *Fourier sum* and each wave in the sum is named a *Fourier term*, as shown in equation (5):

$$f(x) = \sum_{h=0}^{n} F_h \cos 2\pi \left(hx + \alpha_h\right) \tag{5}$$

Given a basic waveform $[\cos 2\pi \left(hx\right) + i \sin 2\pi(hx)]$, the general Fourier sum could be transformed as:

$$f(x) = \sum_{h=0}^{n} F_h \left[\cos 2\pi \left(hx\right) + i \sin 2\pi(hx)\right] \tag{6}$$

Because $\cos \theta + i \sin \theta = e^{i\theta}$, and since in this case $\theta = 2\pi(hx)$, the Fourier sum of equation (6) becomes:

$$f(x) = \sum_{h} F_h e^{2\pi i(hx)} \tag{7}$$

In the case of three-dimensional waves, three variables *h*, *k*, and *l* are needed to specify frequencies in each of the x-, y- and z-axes. A Fourier sum for the wave $f(x, y, z)$ can therefore be written as follows:

$$f(x, y, z) = \sum_{h} \sum_{k} \sum_{l} F_{hkl} e^{2\pi i(hx+ky+lz)} \tag{8}$$

Fourier transform can then be applied (to periodic functions of any dimensions), and for any function $f(x, y, z)$, a function $F\left(h, k, l\right)$ exists, called the Fourier transform of $f(x, y, z)$. $f(x, y, z)$ and $F\left(h, k, l\right)$ can be therefore described as:

$$f(x, y, z) = \int_h \int_k \int_l F(h, k, l)e^{-2\pi i(hx+ky+lz)} \, dhdkdl \qquad (9)$$

$$F(h, k, l) = \int_x \int_y \int_z f(x, y, z)e^{2\pi i(hx+ky+lz)} \, dxdydz \qquad (10)$$

h, k, and l have reciprocal units to that of x, y and z, and are exactly the variables represented by the reciprocal lattice indices. F_{hkl} is the Fourier transform of the electron density equation ρ(x, y, z) on the set of real-space lattice planes (*hkl*), and since F_{hkl} represent a set of discrete reflections of the diffraction pattern, the Fourier transform of it is a triple sum rather than a triple integral:

$$\rho(x, y, z) = \frac{1}{V} \sum_h \sum_k \sum_l F_{hkl} e^{-2\pi i(hx+ky+lz)} \qquad (11)$$

where V is the volume of the unit cell.

7.3.4 **Obtaining phases**

By constructing a Fourier sum using the structure factors F_{hkl}, it is possible to calculate ρ(x, y, z) with equation (11). However, since each structure factor F_{hkl} is a recorded reflection of diffracted ray, and being a wave function, all three factors frequency, amplitude and phase have to be specified for each F_{hkl}. Since the frequencies are the indices of the lattice planes that produce reflection *hkl*, and the amplitude is proportional to the square root of the measured intensity of reflection *hkl,* the only information needed to compute ρ(x, y, z) is the phase.

There are three commonly applied methods to obtain phases: Isomorphous replacement, anomalous scattering and molecular replacement. Isomorphous replacement allows addition of atoms to identical sites of the proteins in all unit cells of a crystal. And the added atoms (usually heave metal atoms) contribute to a slight perturbation in the diffraction pattern, which can be used to obtain phases. The most commonly used technique is to soak the protein crystals in heavy metal solutions, for example ionic complexes of Hg, Pt, Au, etc. Since certain amino acid residues interact

readily with the heavy metal complexes, a specific modification of the protein in a crystal could be expected.

Anomalous scattering takes advantage of the heavy atom's property to absorb X-rays at specific wavelength. As a result, in a protein crystal containing heavy-atom derivatives, the diffractions do not obey the Friedel's law and the reflections hkl and $\overline{hkl}$ are not equal in intensity. This inequality is termed anomalous dispersion. Introducing seleno-labeled methionine as a substitution of methionine in the protein is the most applied technique in anomalous scattering experiments, which allows addition of Se atoms without altering the protein structure.

In both of the methods described above, a powerful tool is used to determine the coordinates of heavy atoms, named the *Patterson function* (equation 12).

$$P(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}|^2 e^{-2\pi i(hu+kv+lw)} \tag{12}$$

Equation (12) shows that the Patterson function is propotional to the square of F_{hkl}, it can be calculated directly from reflection datasets without any information of phases. To obtain the Patterson function only for the heavy atoms, a difference Patterson function is applied. The amplitude differences are $(\Delta F)^2 = (|F_{PH}| - |F_P|)^2$, resulting in the difference Patterson map as:

$$\Delta P(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l (|F_{PH}| - |F_P|)^2 e^{-2\pi i(hu+kv+lw)} \tag{13}$$

where F_{PH} and F_P are intensities from derivative data sets and native datasets, repectively. Given the difference Patterson function, computer softwares can determine vectors between heavy atoms, shown as peaks in the Patterson map. From the peaks, positions of heavy atoms can be calculated.

Another method to solve phase problem is by molecular replacement. In molecular replacement, phases are taken from structure factors of a related protein whose structure is known, to estimate the initial phases of the desired structure. When the

phasing model and the target structure are isomophous (only small differences, such as a new ligand, etc.), the phases from the model can be directly used to compute ρ(x, y, z) together with the intensities from the desired protein, shown in equation (14):

$$\rho(x,y,z) = \frac{1}{V}\sum_h\sum_k\sum_l \left|F_{hkl}^{target}\right|e^{-2\pi i\left(hx+ky+lz-\alpha_{hkl}^{\prime model}\right)} \tag{14}$$

where $\left|F_{hkl}^{target}\right|$ are native intensity amplitudes of the new protein, and $\alpha_{hkl}^{\prime model}$ are phases from the model.

In case the phasing model is nonisomorphous to the new protein, equation (10) is used by computer softwares, trying to put the model into the target unit cell. A theoretical set of structure factors and intensities is calculated and compared with that of the experimental set. At this stage, the proper orientation and positions of the model to be placed in the target unit cell have to be tested. A successful solution should lead to significant better correlations between the experimental intensities and the calculated ones, in which one particular orientation and position of the model shoud be determined. Then the election densities can be calculated with the experimental intensities and the theoretical phases, assuming that the target protein lies in the same manner in the unit cell as the search model does.

7.3.5 **Building and refining models**

After the electron density is calculated from the structure factor amplitudes and the phases, it can be displayed by computer softwares (O, Coot, etc.). Based on the electron density and the knowledge of primary & secondary building blocks of proteins, a model containing specific atom positions and tempareture factors (B-factors) can be built to interpret the density map.

However, the initial electron density map often contains many errors due to the errors of initial phases. Rounds of structural refinement are therefore carried out in order to correct these errors. The refinement can be divided into two parts: the real space refinement and the reciprocal space refinement.

In the real space refinement, the model is corrected according to the electron density map as well as the common rules in the protein structure. For example, all naturally occurring amino acids are L-amino acids, and normally the main-chain peptide bonds of a folded protein are trans-peptide bonds (trans:cis ≈ 1000:1).

In the reciprocal space, a set of structure factor amplitudes $|F_c|$ is calculated from the refined real space model and the phases using the Fourier transform in equation (10), and are compared with the experimental structure factor amplitudes $|F_o|$ by computer softwares. These softwares update the parameters of the atoms in the model in order to minimize the difference between the model and the experimental data. After calculation in reciprocal space is done, two maps are usually generated for the next round of real space refinement, the F_o-F_c map (15) and the $2F_o$-F_c map (16).

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l (|F_o| - |F_c|)e^{-2\pi i(hx+ky+lz-\alpha'_{calc})} \tag{15}$$

$$\rho(x,y,z) = \frac{1}{V}\sum_h \sum_k \sum_l (2|F_o| - |F_c|)e^{-2\pi i(hx+ky+lz-\alpha'_{calc})} \tag{16}$$

The F_o-F_c map excludes the influence of current model and emphasizes the errors by comparing current model with original data. The $2F_o$-F_c map, on the other hand, contains information from both the model and the experimental data. Usually, building of most parts of the model in real space is guided by $2F_o$-F_c map, and errors are checked according to F_o-F_c map.

It usually requires several rounds of refinement in real space and reciprocal space. One key indicator of the model quality is R-factor. R-factor describes the discrepancy between the model and the experimental data. To avoid introducing bias by over-manipulating phases during the refinement, 5% of the data is usually left out and kept the same through out the refinement process. The R-factor of this 5% data is called R_{free}, which is the one of the most important indicators of model quality. Other indicators of model quality include bond length/angel deviations from the ideal value, the Ramachandran plot, etc.

7.4 **Principle of Small Angle X-ray Scattering (SAXS)**

Small Angle X-ray Scattering (SAXS) is a foundamental method for structure analysis of both organic and inorganic materials. In biological research, SAXS provides possibility to study structure of macromolecules in solution. During the experiments, X-rays (with typical wavelength at around 0.15 nm) are scattered elastically by monodisperse non-interacting particles in the solution, and recorded at very low angles (0.1–10º). The scattered intensity *I(s)* is a function of magnitudes of scattering vector *q*. It is known that $q = 4\pi \sin\theta/\lambda$, where θ is half of the angle between the incident and scattered X-ray beam, and λ is the wavelength of the incident X-ray (Figure 7.7).
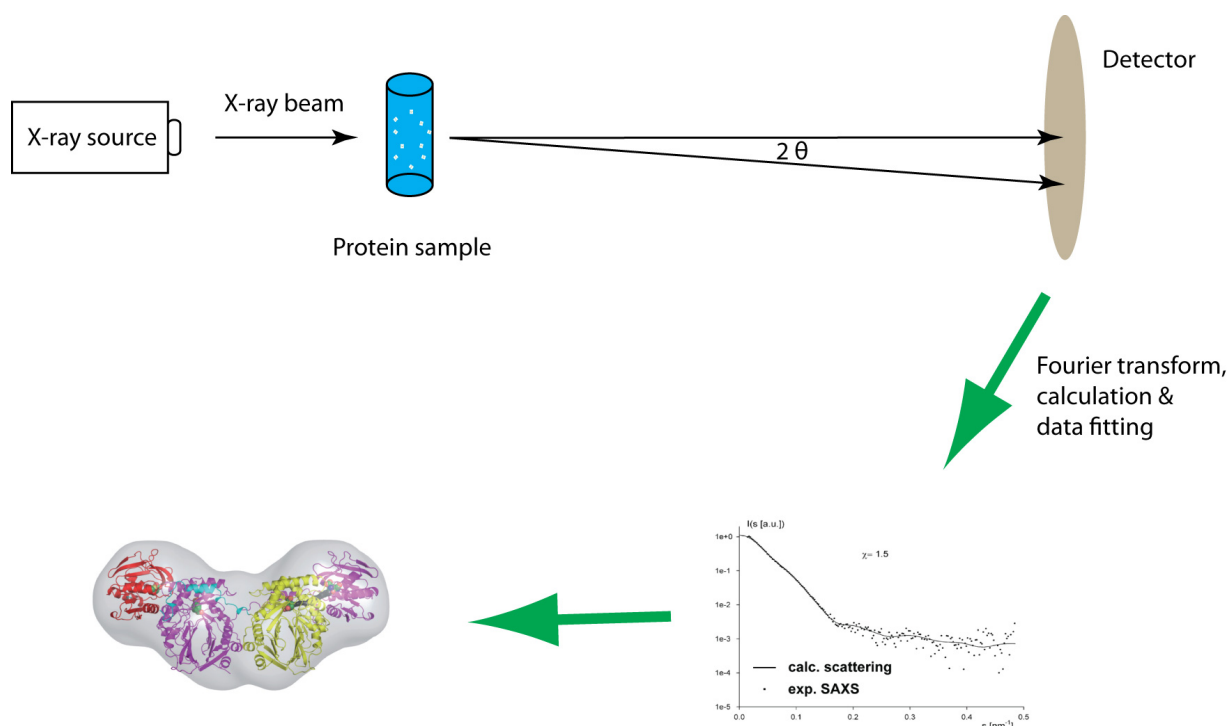
Figure 7.7: Sketch of protein SAXS measurements.

In the solution of monodisperse particles (for example homogeneous protein solution), one strategie to obtain the structural information is to measure the scattering pattern of solutions containing different concentrations of particles, from which the intensity pattern for a single particle can be estimated. This procedure is essential for eliminating concentration effect, which is indicated as a small shoulder in the intensity patterns. The isotropic intensity distribution recorded is propotional to the average

scattering from single particles in random orientations. The scattering intensity is described as:

$$I(q) = P(q)S(q) \tag{17}$$

$I(q)$ is the intensity magnitude of scattering factor q, $P(q)$ and $S(q)$ are the form factor and the structure factor, respectively. Guinier law applies at small q values in the beginning part of the scattering curve, and when the particle concentration is extrapolated from low to infinite dilution, the structure factor $S(q)$ equals to 1. The intensity at small q values is then depending on the gyradius of the particle.

The first step in processing SAXS curves usually is to perform a Fourier transform, and the transformed curves can be interpreted by the distance distribution function, as shown below:

$$P(r) = \frac{r^2}{2\pi^2} \int_0^\infty I(q) \frac{\sin qr}{qr} q^2 dq \tag{18}$$

The distance distribution function starts from 0 when when $r = 0$, and is related to the certain distances r within the particle.

Since SAXS experiments derive three-dimensional structural information from one-dimensional scattering curves, it usually does not imply a single solution, which means different proteins may give identical scattering curve. Reconstitution of three-dimensional structure from SAXS measurements is often conducted with the help of high-resolution X-ray crystal structures or solution NMR structures. Models are taken and different approaches are applied to search for optimum fitting of the experimental data. The quality of the SAXS model could be examined by different metholds. One of the commonly used indicators is the X-value test, as calculated in equation (19):

$$X^2 = \frac{1}{N-1} \sum_{j=1}^{N} \left[\frac{I_{exp}(s_j) - \eta I(s_j)}{\sigma(s_j)} \right]^2 \tag{19}$$

where η is a scaling factor. A typical good fitting usually have a X value between 1 and 5.

7.5 **Principle of Isothermal Titration Calorimetry (ITC)**

Isothermal Titration Calorimetry (ITC) is a thermodynamic technique that could be used to measure the heat released or absorbed during a protein ligand binding process. The measurement allows accurate determination of the binding constant (K_a), enthalpy (ΔH), and the binding stoichiometry (n). The change of entropy could be calculated from ΔH and K_a, given that:

$$\Delta G = -RT \ln K_a = \Delta H - T\Delta S \tag{20}$$

Since it provides a complete thermodynamic characterization of the molecular interaction with only one measurement, and because the measurement does not require the presence of fluorophores or chromophores, the ITC technique has become one of the most important methods for characterizing biomolecular interactions.

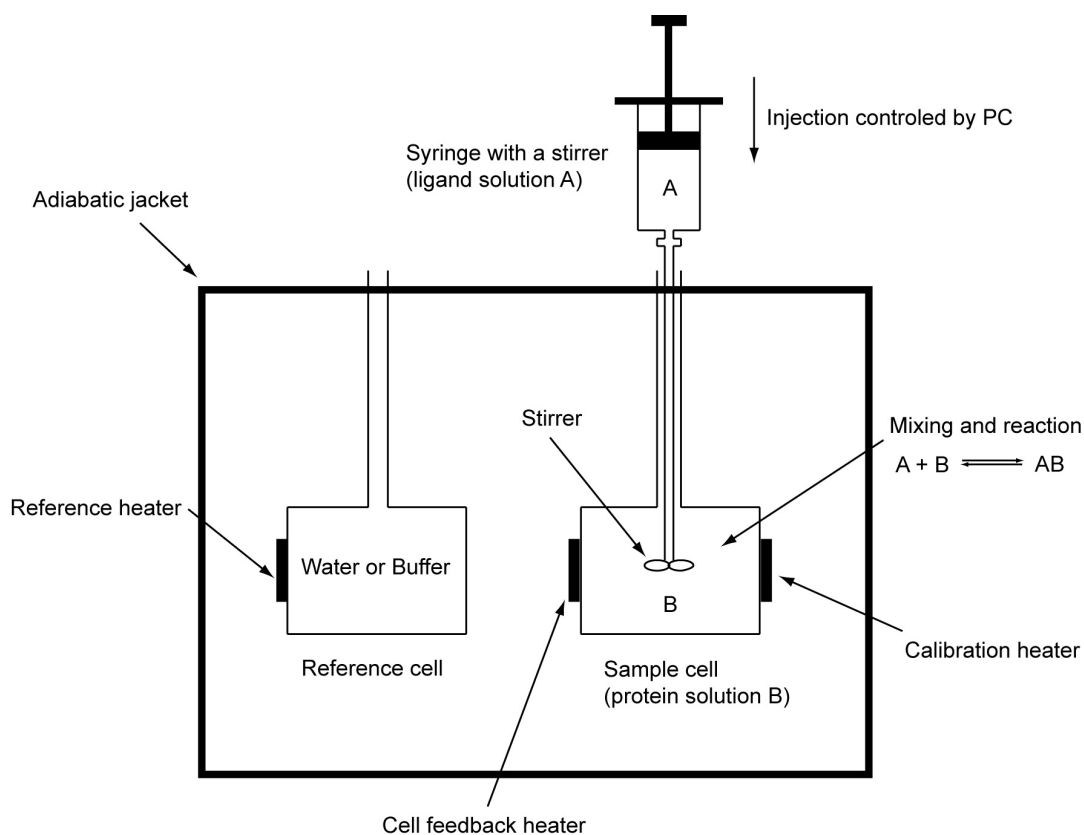A sketch of the instrumental setup is shown in Figure 7.8.

Figure 7.8: Instrumental setup of an ITC for protein-ligand interaction measurement.

At the beginning of the measurement, the sample cell is filled with protein solution (at volume V_0). When an injection of the ligand solution is made (at volumne V_{inj}) to the sample cell, an equal volume of protein solution is driven out. The concentration of the protein remains in the cell after injection i would be:

$$M_{tot,i} = nM_{tot,0}\left(1 - \frac{V_{inj}}{V_0}\right)^i \tag{21}$$

Where $M_{tot,i}$ is the concentration of protein in the cell after injection i, $M_{tot,0}$ is the original concentration of protein, n is the number of lingand binds to one protein molecule, V_0 is the volume of the cell and V_{inj} is the injection volume.

If $L_{tot,i}$ is defined as the total concentratin of ligand in the sample cell after injection i, then:

$$L_{tot,i} = L_{tot,0}\left(\frac{V_{inj}}{V_0}\right)^i \tag{22}$$

where $L_{tot,0}$ is the concentration of ligand in the syringe.

After injection i, the total concentration of protein and ligand in the sample cell is the sum of the concentration of the two molecules in their bound-form and unbound-form, respectively.

$$M_{tot,i} = M_i + ML_i = M_i(1 + K_aL_i) \tag{23}$$

$$L_{tot,i} = L_i + ML_i = L_i(1 + K_aM_i) \tag{24}$$

where M_i, L_i and ML_i are the concentration of unbound protein, unbound ligand, and protein-ligand complex after injection i, respectively. Rearranging equation (23) and (24) gives an quadratic expression (25):

$$K_aL_i{}^2 + \left[1 + K_a\left(M_{tot,i} - L_{tot,i}\right)\right]L_i - L_{tot,i} = 0 \tag{25}$$

The meaningful solution of (25) results in the concentration of unbound ligand in the sample cell after injection i:

$$L_i = \frac{-\left[1 + K_a\left(M_{tot,i} - L_{tot,i}\right)\right] + \sqrt{\left[1 + K_a\left(M_{tot,i} - L_{tot,i}\right)\right]^2 + 4K_aL_{tot,i}}}{2K_a} \qquad (26)$$

The difference in heat content in the sample cell before and after injection *i* is the heat absorbed or produced during the injection, defined as Q_i:

$$Q_i = \langle\Delta H\rangle_i M_{tot,i} V_0 - \langle\Delta H\rangle_{i-1} M_{tot,i-1}\left(V_0 - V_{inj}\right) \qquad (27)$$

where $\langle\Delta H\rangle_i$ is the cumulative change of enthalpy per mole of the protein. If the binding is a one to one binding of ligand to protein, then:

$$\langle\Delta H\rangle_i = \Delta H \frac{ML_i}{M_{tot,i}} = \Delta H \frac{ML_i}{M + ML_i} = \Delta H \frac{K_aL_i}{1 + K_aL_i} \qquad (28)$$

where ΔH is the change of enthalpy when one mole of ligand binds to one mole of protein. After combining equation (27) and (28), Q_i could be expressed as:

$$Q_i = \frac{\Delta H K_a L_i M_{tot,i} V_0}{1 + K_aL_i} - \frac{\Delta H K_a L_{i-1} M_{tot,i-1}\left(V_0 - V_{inj}\right)}{1 + K_aL_{i-1}} \qquad (29)$$

Since $M_{tot,i}$ and L_i can be calculated from equation (23) and (26), respectively, and because Q_i is the parameter measured with the ITC instrument, (29) is a quadratic equation of K_a. Therefore, ΔH, n and K_a can be derived by fitting the ITC data to this equation.

8 REFERENCES

Abken, H.J. et al., 1998. Isolation and characterization of methanophenazine and function of phenazines in membrane-bound electron transport of Methanosarcina mazei Gö1. *J. Bacteriol.*, 180(8), pp.2027-2032.

Abrahams, J.P. & Leslie, A.G.W., 1996. Methods used in the structure determination of bovine mitochondrial F-1 ATPase. *Acta Crystallogr. D Biol. Crystallogr.*, 52, pp.30-42.

Adams, Paul D et al., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2), pp.213-221.

Afonine, P.V., Grosse-Kunstleve, R.W. & Adams, P.D., 2005. *CCP4 Newsl.*, 42(contribution 8).

Agrawal, H. et al., 2007. Ligand based virtual screening and biological evaluation of inhibitors of chorismate mutase (Rv1885c) from Mycobacterium tuberculosis H37Rv. *Bioorg. Med. Chem. Lett.*, 17(11), pp.3053-3058.

Ahuja, E.G. et al., 2008. PhzA/B catalyzes the formation of the tricycle in phenazine biosynthesis. *J.Am.Chem.Soc.*, 130(50), pp.17053-17061.

Alibhai, M.F. & Stallings, W.C., 2001. Closing down on glyphosate inhibition—with a new structure for drug discovery. *Proc. Natl. Acad. Sci.*, 98(6), pp.2944-2946.

Andexer, J.N. et al., 2011. Biosynthesis of the immunosuppressants FK506, FK520, and rapamycin involves a previously undescribed family of enzymes acting on chorismate. *Proc. Natl. Acad. Sci.*, 108(12), pp.4776 -4781.

Ashenafi, M. et al., 2008. The fused TrpEG from Streptomyces venezuelae is an anthranilate synthase, not a 2-amino-4-deoxyisochorismate (ADIC) synthase. *Ethn. Dis.*, 18(2 Suppl 2), pp.S2-9-13.

Baker, N.A. et al., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci.*, 98(18), pp.10037-10041.

Bankhead, S.B. et al., 2004. Minimal changes in rhizobacterial population structure following root colonization by wild type and transgenic biocontrol strains. *FEMS Microbiol. Ecol.*, 49(2), pp.307-318.

Beifuss et al., 2000. Methanophenazine: Structure, Total Synthesis, and Function of a New Cofactor from Methanogenic Archaea. *Angew. chem. Int. Ed. Engl.*, 39(14), pp.2470-2472.

Blankenfeldt, W. et al., 2004. Structure and function of the phenazine biosynthetic protein PhzF from Pseudomonas fluorescens. *Proc.Natl.Acad.Sci.U.S.A*, 101(47), pp.16431-16436.

Bradford, M.M., 1976. Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding. *Anal. Biochem.*, 72(1-2), pp.248-254.

Britton, G., 1983. Biochemistry of Natural Pigments. In *Cambridge Univ. Press.* Cambridge, UK, p. 366 pp.

Byng, G.S., Eustice, D.C. & Jensen, R.A., 1979. Biosynthesis of phenazine pigments in mutant and wild-type cultures of Pseudomonas aeruginosa. *J. Bacteriol.*, 138(3), pp.846-852.

Calfee, M.W., Coleman, J.P. & Pesci, E.C., 2001. Interference with Pseudomonas quinolone signal synthesis inhibits virulence factor expression by Pseudomonas aeruginosa. *Proc. Natl. Acad. Sci.* , 98(20), pp.11633-11637.

Calhoun, D.H., Carson, M. & Jensen, R.A., 1972. The branch point metabolite for pyocyanine biosynthesis in Pseudomonas aeruginosa. *J. Gen. Microbiol.*, 72(3), pp.581-583.

Caligiuri, M.G. & Bauerle, R., 1991a. Identification of amino acid residues involved in feedback regulation of the anthranilate synthase complex from Salmonella typhimurium. Evidence for an amino-terminal regulatory site. *J. Biol. Chem.*, 266(13), pp.8328-8335.

Caligiuri, M.G. & Bauerle, R., 1991b. Subunit communication in the anthranilate synthase complex from Salmonella typhimurium. *Science*, 252(5014), pp.1845-1848.

Chang, P.C. & Blackwood, A.C., 1968. Simultaneous biosynthesis of pyocyanine, phenazine-1-carboxylic acid, and oxychloroaphine from labelled substrates by Pseudomonas aeruginosa Mac 436. *Can. J. Biochem.*, 46(8), pp.925-929.

Chaparro, C. et al., 2001. Infection with Burkholderia cepacia in Cystic Fibrosis . Outcome Following Lung Transplantation. *Am. J. Respir. Crit. Care Med.*, 163(1), pp.43-48.

Chin-A-Woeng, T.F. et al., 2000. Root colonization by phenazine-1-carboxamide-producing bacterium Pseudomonas chlororaphis PCL1391 is essential for biocontrol of tomato foot and root rot. *Mol. Plant. Microbe. Interact.*, 13(12), pp.1340-1345.

Chook, Y.M. et al., 1994. The monofunctional chorismate mutase from Bacillus subtilis. Structure determination of chorismate mutase and its complexes with a transition state analog and prephenate, and implications for the mechanism of the enzymatic reaction. *J. Mol. Biol.*, 240(5), pp.476-500.

Chook, Y.M., Ke, H. & Lipscomb, W.N., 1993. Crystal structures of the monofunctional chorismate mutase from Bacillus subtilis and its complex with a transition state analog. *Proc. Natl. Acad. Sci.* , 90(18), pp.8600-8603.

Cowtan, K., 1994. "dm": An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, 31, pp.34-38.

Cox, C.D., 1986. Role of Pyocyanin in the Acquisition of Iron from Transferrin. *Infect. Immun.*, 52(1), pp.263-270.

Crosa, J.H. & Walsh, Christopher T, 2002. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol. Mol. Biol. Rev.*, 66(2), pp.223-249.

Davis, B.D., 1951. Aromatic biosynthesis I. The role of shikimic acid. , 191, pp.315-326.

DeClue, M.S. et al., 2005. Isochorismate pyruvate lyase: a pericyclic reaction mechanism? *J. Am. Chem. Soc.*, 127(43), pp.15002-15003.

delaFortelle, E. & Bricogne, G., 1997. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Macromolecular Crystallography, Pt A*, 276, pp.472-494.

Denning, G M et al., 1998. Pseudomonas pyocyanin increases interleukin-8 expression by human airway epithelial cells. *Infect. Immun.*, 66(12), pp.5777-5784.

Dietrich, Lars E P et al., 2008. Redox-active antibiotics control gene expression and community behavior in divergent bacteria. *Science*, 321(5893), pp.1203-1206.

Diggle, S.P. et al., 2008. Pseudomonas: Model Organism, Pathogen, Cell Factory. In Weinheim: Wiley-VCH, pp. pp 167-194.

Dosselaere, F. & Vanderleyden, J., 2001a. A metabolic node in action: chorismate-utilizing enzymes in microorganisms. *Crit. Rev. Microbiol.*, 27(2), pp.75-131.

Doublié, S., 1997. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.*, 276, pp.523-530.

Drake, E.J., Nicolai, D.A. & Gulick, A.M., 2006. Structure of the EntB multidomain nonribosomal peptide synthetase and functional analysis of its interaction with the EntE adenylation domain. *Chem. Biol.*, 13(4), pp.409-419.

Du, X. et al., 2001. Crystal Structure and Mechanism of Catalysis of a Pyrazinamidase from Pyrococcus horikoshii†. *Biochemistry*, 40(47), pp.14166-14172.

Emsley, P. et al., 2010. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 4), pp.486-501.

Essar, D.W. et al., 1990. Identification and characterization of genes for a second anthranilate synthase in Pseudomonas aeruginosa: interchangeability of the two anthranilate synthases and evolutionary implications. *J. Bacteriol.*, 172(2), pp.884-900.

Finnan, S. et al., 2004. Genome diversity of Pseudomonas aeruginosa isolates from cystic fibrosis patients and the hospital environment. *J. Clin.Microbiol.*, 42(12), pp.5783-5792.

Fitzpatrick, D.A., 2009. Lines of evidence for horizontal gene transfer of a phenazine producing operon into multiple bacterial species. *J. Mol. Evol.*, 68(2), pp.171-185.

Fordos, M.J., 1859. *Recl. Trav. Soc. d'Émul. Sci. Pharm.*, 3(30).

Gaille, C., Reimmann, C. & Haas, D., 2003. Isochorismate synthase (PchA), the first and rate-limiting enzyme in salicylate biosynthesis of Pseudomonas aeruginosa. *J. Biol. Chem.*, 278(19), pp.16893-16898.

Gasteiger, E. et al., 2005. Protein Identification and Analysis Tools on the ExPASy Server. *John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press*, pp.571-607.

Giddens, S.R., Feng, Y. & Mahanty, H.K., 2002. Characterization of a novel phenazine antibiotic gene cluster in Erwinia herbicola Eh1087. *Mol. Microbiol.*, 45(3), pp.769-783.

Girard, G. & Bloemberg, Guido V, 2008. Central role of quorum sensing in regulating the production of pathogenicity factors in Pseudomonas aeruginosa. *Future Microbiol*, 3(1), pp.97-106.

Gohain, N., Thomashow, Linda S, Mavrodi, Dmitri V & Blankenfeldt, Wulf, 2006a. The purification, crystallization and preliminary structural characterization of FAD-dependent monooxygenase PhzS, a phenazine-modifying enzyme from Pseudomonas aeruginosa. *Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 62(Pt 10), pp.989-992.

Gohain, N., Thomashow, Linda S, Mavrodi, Dmitri V & Blankenfeldt, Wulf, 2006b. The purification, crystallization and preliminary structural characterization of PhzM, a phenazine-modifying methyltransferase from Pseudomonas aeruginosa. *Acta. Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, 62(Pt 9), pp.887-890.

Gouet, P., Robert, X. & Courcelle, E., 2003. ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. *Nucleic Acid Res.*, 31(13), pp.3320 -3323.

Gould, S.J. & Eisenberg, R.L., 1991. The origin of the C-2 hydroxyl in the isochorismate synthase reaction. *Tetrahedron*, 47(31), pp.5979-5990.

Green, J.M. & Nichols, B.P., 1991. p-Aminobenzoate biosynthesis in Escherichia coli. Purification of aminodeoxychorismate lyase and cloning of pabC. *J. Biol. Chem.*, 266(20), pp.12971-12975.

Grisostomi, C. et al., 1997. Efficient in Vivo Synthesis and Rapid Purification of Chorismic Acid Using an Engineered Escherichia coli Strain,. *Bioorg. Chem.*, 25(5-6), pp.297-305.

Guiner, A. & Fournet, G., 1955. *Small angle scattering of X-ray*, New York: Wiley.

Gurusiddaiah, S. et al., 1986. Characterization of an antibiotic produced by a strain of Pseudomonas fluorescens inhibitory to Gaeumannomyces graminis var. tritici and Pythium spp. *Antimicrob. Agents. Chemother.*, 29(3), pp.488-495.

Haagen, Y. et al., 2006. A gene cluster for prenylated naphthoquinone and prenylated phenazine biosynthesis in Streptomyces cinnamonensis DSM 1042. *Chembiochem*, 7(12), pp.2016-2027.

Hanson, J.R. ed., 2003. The classes of natural product and their isolation. In *Natural Products*. Cambridge: Royal Society of Chemistry, pp. 1-34.

Harrison, A.J. et al., 2006. The structure of MbtI from Mycobacterium tuberculosis, the first enzyme in the biosynthesis of the siderophore mycobactin, reveals it to be a salicylate synthase. *J. Bacteriol.*, 188(17), pp.6081-6091.

Hassan, H.M. & Fridovich, I., 1980. Mechanism of the antibiotic action pyocyanine. *J. Bacteriol.*, 141(1), pp.156-163.

Hayashi, Y. et al., 2008. Cloning of the gene cluster responsible for the biosynthesis of brasilicardin A, a unique diterpenoid. *J. Antibiot. (Tokyo)*, 61(3), pp.164-174.

He, Z. et al., 2004. Conservation of mechanism in three chorismate-utilizing enzymes. *J. Am. Chem. Soc.*, 126(8), pp.2378-2385.

Hernandez, M.E., Kappler, A. & Newman, D.K., 2004. Phenazines and other redox-active antibiotics promote microbial mineral reduction. *Appl. Environ. Microbiol.*, 70(2), pp.921-928.

Herrmann, K.M. & Weaver, L.M., 1999. THE SHIKIMATE PATHWAY. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.*, 50, pp.473-503.

Hollstein, U. & Van Gemert, R.J., 1971. Interaction of phenazines with polydeoxyribonucleotides. *Biochemistry*, 10(3), pp.497-504.

Holm, L. & Rosenstrom, P., 2010. Dali server: conservation mapping in 3D. *Nucleic Acid Res.*, 38, p.W545-W549.

Huang, X., Holden, H M & Raushel, F M, 2001. Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.*, 70, pp.149-180.

Ingledew, W.M. & Campbell, J.J., 1969. Evaluation of shikimic acid as a precursor of pyocyanine. *Can. J. Microbiol.*, 15(6), pp.535-541.

Jones, T.A. et al., 1991. Improved Methods for Building Protein Models in Electron-Density Maps and the Location of Errors in These Models. *Acta. Crystallogr. A*, 47, pp.110-119.

Kabsch, W., 2010. XDS. *Acta Crystallogr. D Biol. Crystallogr.*, 66(Pt 2), pp.125-132.

Kaneda, T., Obata, H. & Tokumoto, M., 1993. Aromatization of 4-oxocyclohexanecarboxylic acid to 4-hydroxybenzoic acid by two distinctive desaturases from Corynebacterium cyclohexanicum. Properties of two desaturases. *Eur. J. Biochem.*, 218(3), pp.997-1003.

Kanthakumar, K. et al., 1993. Mechanisms of action of Pseudomonas aeruginosa pyocyanin on human ciliary beat in vitro. *Infect. Immun.*, 61(7), pp.2848-2853.

Kerbarh, Olivier et al., 2006. Crystal structures of Yersinia enterocolitica salicylate synthase and its complex with the reaction products salicylate and pyruvate. *J. Mol. Biol.*, 357(2), pp.524-534.

Knochel, T. et al., 1999. The crystal structure of anthranilate synthase from Sulfolobus solfataricus: Functional implications. *Proc. Natl. Acad. Sci.* , 96(17), pp.9479-9484.

Kolappan, S. et al., 2007. Lysine 190 is the catalytic base in MenF, the menaquinone-specific isochorismate synthase from Escherichia coli: implications for an enzyme family. *Biochemistry*, 46(4), pp.946-953.

Konarev, P.V. et al., 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.*, 36, pp.1277-1282.

Larkin, M.A. et al., 2007. Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)*, 23(21), pp.2947-2948.

Lau, G.W. et al., 2004. Pseudomonas aeruginosa pyocyanin is critical for lung infection in mice. *Infect. Immun.*, 72(7), pp.4275-4278.

Lauredo, I.T. et al., 1998. Mechanism of pyocyanin- and 1-hydroxyphenazine-induced lung neutrophilia in sheep airways. *J. Appl. Physiol.*, 85(6), pp.2298-2304.

Laursen, J.B. & Nielsen, J., 2004. Phenazine natural products: Biosynthesis, synthetic analogues, and biological activity. *Chem. Rev.*, 104(3), pp.1663-1685.

166

Levitch, M.E. & Rietz, P., 1966. The isolation and characterization of 2-hydroxyphenazine from Pseudomonas aureofaciens. *Biochemistry*, 5(2), pp.689-692.

Levitch, M.E. & Stadtman, E.R., 1964. STUDY OF THE BIOSYNTHESIS OF PHENAZINE-1-CARBOXYLIC ACID. *Arch. Biochem. Biophys.*, 106, pp.194-199.

Longley, R.P. et al., 1972. Branchpoint of Pyocyanine Biosynthesis. *Can. J. Microbiol.*, 18(9), pp.1357-&.

Loscher, R. & Heide, L., 1994. Biosynthesis of p-Hydroxybenzoate from p-Coumarate and p-Coumaroyl-Coenzyme A in Cell-Free Extracts of Lithospermum erythrorhizon Cell Cultures. *Plant Physiol.*, 106(1), pp.271-279.

Mattei, P., Kast, P & Hilvert, D, 1999. Bacillus subtilis chorismate mutase is partially diffusion-controlled. *Eur. J. Biochem.*, 261(1), pp.25-32.

Mavrodi, D.V., Blankenfeldt, W. & Thomashow, L.S., 2006. Phenazine compounds in fluorescent Pseudomonas spp. Biosynthesis and regulation. *Annu. Rev. Phytophathol.*, 44, pp.417-445.

Mavrodi, D.V. et al., 2001. Functional analysis of genes for biosynthesis of pyocyanin and phenazine-1-carboxamide from Pseudomonas aeruginosa PAO1. *J. Bacteriol.*, 183(21), pp.6454-6465.

Mavrodi, D.V. et al., 1998. A seven-gene locus for synthesis is of phenazine-1-carboxylic acid by Pseudomonas fluorescens 2-79. *J. Bacteriol.*, 180(9), pp.2541-2548.

Mavrodi, Dmitri V et al., 2010. Diversity and evolution of the phenazine biosynthesis pathway. *Appl. Environ. Microbiol.*, 76(3), pp.866-879.

McCoy, A.J. et al., 2007. Phaser crystallographic software. *J. Appl. Crystallogr.*, 40(Pt 4), pp.658-674.

McDonald, M. et al., 2001. Phenazine biosynthesis in Pseudomonas fluorescens: Branchpoint from the primary shikimate biosynthetic pathway and role of phenazine-1,6-dicarboxylic acid. *J. Am. Chem. Soc.*, 123(38), pp.9459-9460.

Mentel, Matthias, Ahuja, Ekta G, Mavrodi, Dmitri V, Breinbauer, Rolf, Thomashow, Linda S & Blankenfeldt, Wulf, 2009a. Of two make one: the biosynthesis of phenazines. *Chembiochem*, 10(14), pp.2295-2304.

Mentel, Matthias, Ahuja, Ekta G, Mavrodi, Dmitri V, Breinbauer, Rolf, Thomashow, Linda S & Blankenfeldt, Wulf, 2009b. Of two make one: the biosynthesis of phenazines. *Chembiochem*, 10(14), pp.2295-2304.

Millican, R.C., 1962. Biosynthesis of pyocyanine. Incorporation of [14C]shikimic acid. *Biochim. Biophys. Acta.*, 57(1), pp.407-409.

Morollo, A.A. & Bauerle, R., 1993. Characterization of Composite Aminodeoxyisochorismate Synthase and Aminodeoxyisochorismate Lyase Activities of Anthranilate Synthase. *Proc. Natl. Acad. Sci.* , 90(21), pp.9983-9987.

Morollo, A.A. & Eck, M.J., 2001. Structure of the cooperative allosteric anthranilate synthase from Salmonella typhimurium. *Nat. Struct. Biol.*, 8(3), pp.243-247.

Mouilleron, S. & Golinelli-Pimpaneau, B., 2007. Conformational changes in ammonia-channeling glutamine amidotransferases. *Curr. Opin. Struct. Biol.*, 17(6), pp.653-664.

Muller, M., 2002. Pyocyanin induces oxidative stress in human endothelial cells and modulates the glutathione redox cycle. *Free. Radic. Biol. Med.*, 33(11), pp.1527-1533.

Murshudov, G.N., Vagin, A.A. & Dodson, E.J., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.*, 53, pp.240-255.

Nakai, K. & Kanehisa, M., 1991. Expert system for predicting protein localization sites in gram-negative bacteria. *Proteins*, 11(2), pp.95-110.

Newman, D.J. & Cragg, G.M., 2007. Natural Products as Sources of New Drugs over the Last 25 Years. *J. Nat. Prod.*, 70(3), pp.461-477.

O'Malley, Y.Q. et al., 2004. Pseudomonas aeruginosa pyocyanin directly oxidizes glutathione and decreases its levels in airway epithelial cells. *Am. J. Physiol. Lung Cell Mol. Physiol.*, 287(1), pp.L94-103.

Parsons, J.F. et al., 2003. Structure and mechanism of Pseudomonas aeruginosa PhzD, an isochorismatase from the phenazine biosynthetic pathway. *Biochemistry*, 42(19), pp.5684-5693.

Parsons, J.F. et al., 2004. Structure of the phenazine biosynthesis enzyme PhzG. *Acta Crystallogr. D Biol. Crystallogr.*, 60, pp.2110-2113.

Parsons, J.F. et al., 2004. Structure and function of the phenazine biosynthesis protein PhzF from Pseudomonas fluorescens 2-79. *Biochemistry*, 43(39), pp.12427-12435.

Parsons, James F et al., 2007. Structural and functional analysis of the pyocyanin biosynthetic protein PhzM from Pseudomonas aeruginosa. *Biochemistry*, 46(7), pp.1821-1828.

Parsons, James F et al., 2002. Structure of Escherichia coli aminodeoxychorismate synthase: architectural conservation and diversity in chorismate-utilizing enzymes. *Biochemistry*, 41(7), pp.2198-2208.

Parsons, James F., Shi, K.M. & Ladner, Jane E., 2008. Structure of isochorismate synthase in complex with magnesium. *Acta Crystallographica Section D Biological Crystallography*, 64(5), pp.607-610.

Pesci, E.C. et al., 1999. Quinolone signaling in the cell-to-cell communication system of Pseudomonas aeruginosa. *Proc. Natl. Acad. Sci.*, 96(20), pp.11229-11234.

Petrek, M. et al., 2006. CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics*, 7, p.316.

Pierson, L.S. et al., 1995. Molecular analysis of genes encoding phenazine biosynthesis in the biological control bacterium. Pseudomonas aureofaciens 30-84. *FEMS Microbiol. Lett.*, 134(2-3), pp.299-307.

Pierson, L.S. & Thomashow, L S, 1992. Cloning and heterologous expression of the phenazine biosynthetic locus from Pseudomonas aureofaciens 30-84. *Mol. Plant. Microbe. Interact.*, 5(4), pp.330-339.

Podojil, M. & Gerber, N.N., 1967. The biosynthesis of 1,6-phenazinediol 5,10-dioxide (Iodinin) by Brevibacterium iodinum. *Biochemistry*, 6(9), pp.2701-2705.

Policastro, P.P. et al., 1984. trans-6-Amino-5-[(1-carboxyethenyl)oxy]-1,3-cyclohexadiene-1-carboxylic acid: an intermediate in the biosynthesis of anthranilate from chorismate. *J. Am. Chem. Soc.*, 106(8), pp.2443-2444.

Price-Whelan, A., Dietrich, L.E.P. & Newman, D.K., 2006. Rethinking "secondary" metabolism: physiological roles for phenazine antibiotics. *Nat. Chem. Biol.*, 2(2), pp.71-78.

Price-Whelan, Alexa, Dietrich, Lars E P & Newman, Dianne K, 2007. Pyocyanin alters redox homeostasis and carbon flux through central metabolic pathways in Pseudomonas aeruginosa PA14. *J. Bacteriol.*, 189(17), pp.6372-6381.

Raushel, Frank M, Thoden, J.B. & Holden, Hazel M, 2003. Enzymes with molecular tunnels. *Acc. Chem. Res.*, 36(7), pp.539-548.

Roessle, M.W. et al., 2007. Upgrade of the small-angle X-ray scattering beamline X33 at the European Molecular Biology Laboratory, Hamburg. *J Appl Crystallogr*, 40, p.s190-s194.

Round, A.R. et al., 2008. Automated sample-changing robot for solution scattering experiments at the EMBL Hamburg SAXS station X33. *J. Appl. Crystallogr.*, 41(5), pp.913-917.

Ruan, X. et al., 1997. A second type-I PKS gene cluster isolated from Streptomyces hygroscopicus ATCC 29253, a rapamycin-producing strain. *Gene*, 203(1), pp.1-9.

Saiki, R.K. et al., 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732), pp.1350-1354.

Saleh, O. et al., 2009. Aromatic prenylation in phenazine biosynthesis: dihydrophenazine-1-carboxylate dimethylallyltransferase from Streptomyces anulatus. *J. Biol. Chem.*, 284(21), pp.14439-14447.

Schüttelkopf, A.W. & van Aalten, D.M.F., 2004. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. D Biol. Crystallogr.*, 60(Pt 8), pp.1355-1363.

Sheldrick, G.M., 2010. <title> Experimental phasing with *SHELXC / D / E* : combining chain tracing with density modification. *Acta Crystallographica Section D Biological Crystallography*, 66(4), pp.479-485.

Siebert, M., Severin, K. & Heide, L, 1994. Formation of 4-hydroxybenzoate in Escherichia coli: characterization of the ubiC gene and its encoded enzyme chorismate pyruvate-lyase. *Microbiology*, 140 (Pt 4), pp.897-904.

Smirnov, V. & Kiprianova, E., 1990. Bacteria of Pseudomonas genus. In Naukova Dumka, p. 264 pp.

Smith, N. et al., 2006. Structural analysis of ligand binding and catalysis in chorismate lyase. *Arch. Biochem. Biophys.,* 445(1), pp.72-80.

Spraggon, G. et al., 2001. The structures of anthranilate synthase of Serratia marcescens crystallized in the presence of (i) its substrates, chorismate and glutamine, and a product, glutamate, and (ii) its end-product inhibitor, L-tryptophan. *Proc.Natl.Acad.Sci.*, 98(11), pp.6021-6026.

Sridharan, S. et al., 2010. Crystal Structure of Escherichia coli Enterobactin-specific Isochorismate Synthase (EntC) Bound to its Reaction Product Isochorismate: Implications for the Enzyme Mechanism and Differential Activity of Chorismate-utilizing Enzymes. *J.Mol.Biol.*, 397(1), pp.290-300.

Stewart, A.J. et al., 2001. Antitumor activity of XR5944, a novel and potent topoisomerase poison. *Anti-Cancer Drugs*, 12(4), pp.359-367.

Stover, C.K. et al., 2000. Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen. *Nature*, 406(6799), pp.959-964.

Stover, C. et al., 2000. Crystallization and 1.1-A diffraction of chorismate lyase from Escherichia coli. *J. Struct. Biol.*, 129(1), pp.96-99.

Svergun, D.I., 1999. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, 76(6), pp.2879-2886.

Svergun, D.I., 1993. A direct indirect method of small-angle scattering data treatment. *J. Appl. Crystallogr.*, 26(2), pp.258-267.

Thomashow, L S & Weller, D M, 1988. Role of a phenazine antibiotic from Pseudomonas fluorescens in biological control of Gaeumannomyces graminis var. tritici. *J. Bacteriol.*, 170(8), pp.3499-3508.

Turner, J.M. & Messenger, A.J., 1986. Occurrence, Biochemistry and Physiology of Phenazine Pigment Production. *Adv. Microb. Physiol.*, 27, pp.211-275.

Vagin, A. & Teplyakov, A., 1997. MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Cryst.*, 30, pp.1022-1025.

Van Lanen, S.G., Lin, S. & Shen, B., 2008. Biosynthesis of the enediyne antitumor antibiotic C-1027 involves a new branching point in chorismate metabolism. *Proc. Natl. Acad. Sci.* , 105(2), pp.494-499.

Villavicencio, R.T., 1998. The history of blue pus. *J. Am. Coll. Surg.*, 187(2), pp.212-216.

Webby, C.J. et al., 2005. The structure of 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase from Mycobacterium tuberculosis reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *J. Mol. Biol.*, 354(4), pp.927-939.

Wilson, R. et al., 1988. Measurement of Pseudomonas-Aeruginosa Phenazine Pigments in Sputum and Assessment of Their Contribution to Sputum Sol Toxicity for Respiratory Epithelium. *Infect. Immun.*, 56(9), pp.2515-2517.

Wriggers, W., 2010. Using Situs for the integration of multi-resolution structures. *Biophys. Rev.*, 2(1), pp.21-27.

Ye, Q.Z., Liu, J. & Walsh, C T, 1990. p-Aminobenzoate synthesis in Escherichia coli: purification and characterization of PabB as aminodeoxychorismate synthase and enzyme X as aminodeoxychorismate lyase. *Proc. Natl. Acad. Sci.* , 87(23), pp.9391-9395.

Zalkin, H., 1993. The amidotransferases. *Adv. Enzymol. Relat. Areas Mol. Biol.*, 66, pp.203-309.

Zwahlen, J. et al., 2007. Structure and mechanism of MbtI, the salicylate synthase from Mycobacterium tuberculosis. *Biochemistry*, 46(4), pp.954-964.

ACKNOWLEDGEMENT

as colleagues from **the whole MPI Dortmund** for their contribution to a fantastic atmosphere.

Last but not least, I could not imagine achieving my Ph.D. without consistent love and support from my Mom and Dad, who are the pillar of my life and are truly beloved.

CURRICULUM VITAE

Name: Qiang Li	Date of birth: 30.06.1982	
Place of birth: Beijing	Nationality: Chinese	
Email: eddie.qali@gmail.com		

EDUCATION

03/2007 – summer 2011	Ph.D. candidate in Chemical Biology (expected Aug. 2011). Max Planck Institute of Molecular Physiology, Dortmund, Germany
09/2005 – 02/2007	Master student of International Max Planck Research School in Chemical Biology, Dortmund, Germany
09/2000 – 07/2004	Bachelor of Science in Biological Sciences & Biotechnology Tsinghua University, Beijing, China

RESEARCH EXPERIENCE

03/2007 – 11/2010	Structural and functional insights into initial steps of phenazine biosynthesis. **(Ph.D. thesis)**
03/2010 – 09/2010	Crystal structure of phenazine biosynthesis protein PhzD in complex with (i) its substrate ADIC or (ii) its product DHHA.
09/2005 – 03/2007	Cloning, expression, purification and functional studies of two phenazine biosynthesis proteins, PhzC and PhzE.
10/2003 – 06/2004	Study of the effect of conotoxin on synchronized spontaneous calcium spikes in cultured hippocampal networks using calcium imaging. **(Bachelor thesis)**

PUBLICATIONS

Li, Q-A., Mavrodi, D.V., Thomashow, L.S., Roessle, M. & Blankenfeldt, W. Ligand binding induces an ammonia channel in 2-amino-2-deoxyisochorismate (ADIC) synthase PhzE. *J. Biol. Chem.* **286**, 18213-18221 (2011).

Wang X, Xie LP, **Li Q-A**, Zhang RQ, Zhou XW & Huang PT. Effect of O-superfamily conotoxin SO3 on synchronized spontaneous calcium spikes in cultured hippocampal networks. *Cell Biol. Toxicol.* **24**, 11-17 (2008).