Structural and functional analysis of virulence factors of *Xanthomonas oryzae* pv. oryzae, the bacterial leaf blight pathogen of rice

Thesis submitted to

Jawaharlal Nehru University

for the degree of

Doctor of Philosophy

G. Aparna

Centre for Cellular and Molecular Biology (CCMB),

Hyderabad-500007, India.

2009

CERTIFICATE

The research work presented in this thesis has been carried out at the Centre for Cellular and Molecular Biology, Hyderabad, India. This work is original and has not been submitted in part or in full for any other degree or diploma to any other university or institute.

Dr. Ramesh V. Sonti
(Thesis Supervisor)

Dr. Rajan Sankaranarayanan
(Thesis Supervisor)

G. Aparna

ACKNOWLEDGEMENTS

This work is an outcome of priceless advice & good wishes from several people & I take this opportunity to thank all of them. First & foremost, I extend my heartfelt gratitude to my thesis supervisors ***Dr. Ramesh Sonti*** *&* ***Dr. Rajan Sankaranarayanan****, whose unconditional support, fine guidance & mentoring gave me an excellent platform to grow as a researcher. Their collective experience & diverse perspective on various issues helped me learn '2X'! I thank* ***Dr. Lalji Singh****, Director, CCMB for providing me an opportunity to pursue research at this esteemed state-of-the-art institute. I also thank CSIR, India for financial support.*

I will always treasure Dr. Rajeshwari Ramanan's tutoring during the early part of my Ph. D. and her infectious enthusiasm. Interesting ideas & technical help from all my lab members (Sujatha, Vishnupriya, Hitendra, Prabhu, Amit, Rukmini, Rajakumara, Shweta, Gopaljee, Zahid, Alok, Aravind, Aneesh, Madhan, Tanweer, Anil, Asfarul, Dipanwita, Madhumita, Shobha, Biswajit, Shambavi, Sadeem, Ravi, Ashish, Shanti, Venu, Sushil & Lavanya) at various phases of the work is gratefully appreciated. I sincerely thank Avradip for assisting me with many experiments and innumerable protein purifications. Invaluable suggestions from Drs. T. Ramakrishnan, Amirul Islam, M. V. Jagannadham, Nandini Rangaraj, Shashi Singh, Mandar Deshmukh & Will Stanley helped me troubleshoot various technical problems. I learnt a lot about chromatography, spectrophotometry & antibodies from Venkat, Lora, Asha, Dwarakanath & I thank them for their help. I also gratefully acknowledge Joshiji for the timely provision of fine chemicals. Working at CCMB, any part of it, is a pleasure & more so, if it is the east wing second floor or the basement and I laud all the members of these floors. I thank Sahasransu for his help with primer extension and for some very innovative ideas that have been incorporated as experiments in this work. I take this opportunity to warmly acknowledge all my friends for their constant encouragement.

My parents and brother are pivotal to all my endeavours and no acknowledgement can suffice for their thought and love. I dedicate my thesis to them...

To My Parents

Who Live My Beliefs & Cherish My Dreams

TABLE OF CONTENTS

SYNOPSIS

Coexistence with plant hosts imposes a selection pressure on the phytopathogens, leading to the evolution of a repertoire of specialized features that provide advantage in attachment, entry and growth inside the plant. Comprehensive studies on various plant-pathogen interactions can provide insights into the various adaptive features in pathogen proteins and consequently, the process of plant pathogenesis as a whole. The aim of this study is to gain a better understanding of the virulence functions of *Xanthomonas oryzae* pv. oryzae (Xoo), the causative agent of bacterial leaf blight of rice. Towards this aim, structural and functional aspects of two key virulence factors of Xoo were studied.

Chapter 1 provides an overview of the factors defining pathogen evolution and the host-pathogen coevolution with emphasis on plant pathogens. The chapter outlines our current knowledge on the virulence factors of Xoo and their role in rice pathogenesis. The need for structural characterization of these virulence factors for a better understanding of their *in planta* functions is emphasized.

Chapter 2 is an in depth description of the crystallization and structure solution of Xoo esterase LipA, a secreted enzyme, which plays an important role in the degradation of rice cell wall during pathogenesis. Preparation of heavy atom derivatives of LipA crystals and the use of these crystals to obtain the high-resolution crystal structure of LipA using a method known as Multiple Isomorphous Replacement is explained. An overview of the general structural features of LipA is also provided.

Chapter 3 is a comprehensive study of a ligand-bound form of LipA structure and the insights it provides into the function of this enzyme. LipA structure reveals an all-helical ligand-binding domain as a unique functional attachment to the canonical hydrolase scaffold, with a

glycoside recognition pocket. Disruption of this pocket and its constituents, the distinct sugar and acyl chain binding regions, causes loss of LipA function and concomitant loss of virulence and elicitation of host defense responses. LipA-like substrate recognition is a remarkable example of the evolution of new functions around existing scaffolds to promote plant pathogenesis.

Chapter 4 details the *in silico* investigations on Xanthomonas adhesin-like protein A (XadA), one of the virulence factors of Xoo that promotes Xoo entry into rice leaves through the hydathodal pores. This analysis led to the detection of interesting sequence features of XadA and the way in which these might dictate the structure and function of XadA. The chapter also describes the identification of four distinct domains in XadA, each equivalent to the autotransporter adhesins of animal pathogenic bacteria. Homology models of each of the domains and the most probable model of the whole XadA molecule are shown.

Chapter 5 offers a discussion on the regulation of XadA expression. Antibodies raised against one of the domains of XadA have been used to assess the regulatory control of XadA. Surprisingly, XadA is expressed in a glutamate-dependent manner under post-translational regulation. Glutamate is a major component of the exudates of the rice hydathodal pores. It appears that Xoo is using the presence of glutamate as a indicator to perceive that it is in the proximity of the hydathodal pore for the expression of XadA.

Overall, this thesis work provides the structural characterization of a unique mode of substrate recognition by a rice cell wall degrading esterase and the evidence for a distinct glutamate-mediated regulation of the expression of a Xanthomonas adhesin-like protein of Xoo.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

°C:	degree Celsius
A:	Adenine
BOG:	β-octyl glucoside
bp:	Base pair
BSA:	Bovine Serum Albumin
C:	Cytosine
EDTA:	Ethylene diamine tetracetic acid
G:	Guanine
HPLC:	High Performance Liquid Chromatography
hr:	Hour
IPTG:	Isopropyl β-thiogalactoside
kDa:	Kilo dalton
mg:	milligram
min:	minute
ml:	millilitre
ng:	nanogram
Ni-NTA:	Nickel-Nitriloacetic acid
nm:	Nanometre
nM:	Nanomolar
O.D:	Optical density
PAGE:	Poly acrylamide gel electrophoresis
PCR:	Polymerase chain reaction
PMSF:	Phenyl methyl sulfonyl chloride
rpm:	Rotation per minute
SDS:	Sodium dodecyl phosphate
T:	Thymidine
Xac:	*Xanthomonas axonopodis* pv. citri
Xav:	*Xanthomonas axonopodis* pv. vesicatoria
Xcc:	*Xanthomonas campestris* pv. campestris
Xoc:	*Xanthomonas oryzae* pv. oryzicola
Xoo:	*Xanthomonas oryzae* pv. oryzae
μg:	microgram
μl:	microlitre

LIST OF PUBLICATIONS

ırna, G., Chatterjee, A., Sonti, R. V., Sankaranarayanan, R. **(2009)** A cell wall-degrading rase of Xanthomonas oryzae requires a unique substrate recognition module for ıogenesis on rice. **Plant Cell**. 21: 1860-1873.

:berg, S. L., Sommer, D. D., Schatz, M. C., Phillippy, A. M., Rabinowicz, P. D., Tsuge, Furutani, A., Ochiai, H., Delcher, A. L., Kelley, D., Madupu, R., Puiu, D., Radune, D., mway, M., Trapnell, C., **Aparna, G.**, Jha, G., Pandey, A., Patil, P. B., Ishihara, H., 'er, D. F., Szurek, B., Verdier, V., Koebnik, R., Dow, J. M., Ryan, R. P., Hirata, H., yumu, S., Won Lee, S., Seo, Y. S., Sriariyanum, M., Ronald, P. C., Sonti, R. V., Van 's, M. A., Leach, J. E., White, F. F., Bogdanove, A. J. **(2008)** Genome sequence and rapid ution of the rice pathogen *Xanthomonas oryzae* pv. oryzae PXO99A. **BMC Genomics**. 9:

ırna, G., Chatterjee, A., Jha, G., Sonti, R. V., Sankaranarayanan, R. **(2007)** Crystallization preliminary crystallographic studies of LipA, a secretory lipase/esterase from *thomonas oryzae* pv. oryzae. **Acta Crystallogr. Sect. F** Struct. Biol. Cryst. Commun. 63: 710.

Chapter One

Introduction

CHAPTER ONE

1.1 Pathosystems and host-pathogen coevolution

The word 'pathogenesis' is derived from the Greek words '*pathos*' and '*genesis*', which literally mean disease and creation, respectively. It is a collective term used for the series of events leading to a functional abnormality in the body of a living organism. Although certain disease conditions can be attributed to agents like toxic chemicals, allergens, radiations, genetic predispositions etc., many pathologies are associated with pathogenic microbial agents. While the pathogens utilize the host organism as a source of nutrients, the parasitic multiplication of these microbes is detrimental to the host, leading to clinical symptoms that can be fatal. Both animals and plants are susceptible to fungal, bacterial, protozoan and viral pathogenesis, in addition to invasive attacks by nematodes, mites and insects. The pathogen, its host and all the environmental conditions affecting the infection are the three essential components of an integrated ecological system called pathosystem.

Post infection, the behaviour of host and pathogen is treated as a single system and thus, the central theme of a pathosystem is the interaction between the populations of pathogen and its host (Robinson, 1986). The close co-existence and constant conflict between the divergent interests of the pathogen and the host exerts strong evolutionary pressure on the two species. This selective process is often referred to as coevolution, a phenomenon analogous to an arms race where weapons are amassed as a counter-measure against the weapons of the competing partner. A hugely accepted Red Queen hypothesis states that "For an evolutionary system, continuing development is needed in order to maintain its fitness relative to the system it is co-evolving with" (Van Valen, 1973). The fitness or reproductive success of the pathogen depends on a better survival rate after infection, especially by evasion of host immune system. Pathogenic mechanisms that modulate host immunity have the most profound effect on the pathogen fitness (Frank & Schmid-Hempel, 2008). The host fitness, in turn, is increased by either escaping invasion by developing resistance or destruction of the resident pathogens. Therefore, coevolution involves a large fitness cost and selects for stronger virulence in pathogens or better resistance in hosts. This reciprocal and competitive selection pressure leads to an 'extended phenotype' arising from not only the respective genotypes but also the genotype of the partner (Dawkins, 1982). Interestingly, in addition to the phenotypes, traits

like virulence, disease transmission, extent of disease and relapse are under the dual control of both host and the pathogen genomes (Lambrechts et al., 2006). From this concept, stems the gene-for-gene model, which in its simplest manifestation, speculates that for each gene conferring resistance to the host, there is a corresponding 'avirulence' gene in the pathogen (Flor, 1956). This model has been proved right repeatedly in a number of plant-pathogen interactions (Flor, 1971). The crosstalk between resistance-conferring gene products of the plant and certain proteins of the pathogen, called avirulence proteins, finally determines the outcome of infection, thereby playing an important role in pathosystem dynamics.

1.2 Understanding pathosystems through protein evolution

Pathogens as well as the host are under selection pressure to increase fitness and this leads to a set of physiological changes known as adaptations. Neo-Darwinian theory explains that when a repertoire of genetic variations is subjected to natural selection, the traits best adapted to a particular environment arise. Variations at the genetic level have a direct consequence on the resultant proteins. When selection pressure is the host-pathogen coexistence, marked changes in proteomes of both partners ensue. The changes include evolution of new specialization in the existing proteins as well as new types of proteins. Comparing protein sequences and three-dimensional structures, usually, gives a good estimate of the evolutionary timescales as well as phylogenetic relationships, even between different pathogen and host species. Acquisition of large chunks of genetic material from external sources like phages is a common mode of introducing distinct traits and has been exploited heavily by pathogens to acquire novel pathogenic features in a commendably short time frame (Wilson et al., 1977). Nevertheless, subtle changes in protein sequences can also lead to large phenotypic variations (Thornton et al., 1999). Interestingly, analysis of several protein structures shows that the total number of protein folds is not equivalent to the total number of proteins, suggesting that several proteins can possess the same fold (Chothia, 1993). Predominance of a certain fold has been correlated with its stability and the capacity to adapt to multiple functions with minimum modification (Goldstein, 2008). Selection pressure may push a particular protein scaffold to adapt a different function in a pathosystem and thereby, evolve newer functions. 'Designability' of proteins is much simpler when they are arranged in modules or domains. Rearrangements of domains and different permutations and combinations of domains are more ways of introducing novelty into protein functions (Deeds & Shakhnovich, 2006; Moore

et al., 2008). Pathosystems seem to have taken up a multipronged approach to achieve variability and thus, manipulate host and pathogen fitness.

1.3 Bacterial-plant pathosystems

Plant pathosystem is a specific term used when the host is a plant. A large number of plant pathosystems are bacterial and more than 200 species of bacterial phytopathogens have been identified until now (Schuster & Coyne, 1974). While animal and plant pathogens share several features of pathogenesis such as the ability to secrete proteins into host cells and modulate host responses, each group has evolved numerous specializations to gain advantage in specialized niches and thus, enhanced fitness. Diverse plant hosts can be infected by pathovars belonging to the same bacterial species, implicating immense specificity in plant-bacterial interactions. Bacterial phytopathogens have an epiphytic stage in their lifecycles where they attach to the plant host and increase their population sizes without causing disease (Blakeman, 1991). Existence of large number of bacteria in close proximity to the host increases the chances of successful entry (Montesinos et al., 2002). The onset of pathogenic stage occurs only when the environmental conditions are favourable. The bacteria enter the host and try to access plant nutrients. It is important to note that bacterial phytopathogens colonize the intercellular spaces and never gain entry into the plant cells.

While it may be true that being sessile, plants cannot evade pathogenic attack, they have developed ways of detecting and circumventing the same. Bacterial infections are not merely the pathogen 'attach-invade-prevail' series of events. Plants do sense the presence of the bacteria but the detection is not due to a circulating immune system like in the animal bodies. Every plant cell can either detect the damage caused by bacterial enzymes to the plant cell wall or perceive certain bacterial markers using cell membrane receptors. Downstream signalling pathways in plants can lead to short-term defense responses and long-term disease resistance (Staskawicz et al., 2001). The ability to overcome these host responses is an extremely important virulence attribute and plays an important role in shaping disease epidemics (Sacristan & Garcia-Arenal, 2008).

The socio-economic effects of the bacterial infections of important crops have been immense and pose a continuing threat to the world food security today. Bacterial blights, cankers, wilts, scabs and soft rots have taken epidemic proportions across the world, leading to great

agricultural losses and famines. Although fungi surpass bacteria in the number and severity of plant diseases, the latter are dreaded because of lack of effective control and post-infective treatment measures. Large-scale field application of bactericides and antibiotics is not effective because they have deleterious effects on the host plant and there are concerns about exposure of non-target animal pathogenic bacteria to these chemicals, with a consequential reduction in the efficacy of these drugs for treating human infections. Development of disease resistant plant varieties has proved beneficial for combating bacterial infections (Burkholder, 1948), but this is time-consuming and the pathogen can evolve to breakdown host resistance genes. Development of more proficient prognostic and diagnostic remedies of bacterial plant disease requires a thorough understanding of the biology of these pathosystems. Information regarding symptomatic identification of the disease, etiology, epidemiology and the genetics of the pathosystem is essential for management of the particular disease. Moreover, it is imperative to use advanced biotechnological techniques and new approaches to appreciate the host-pathogen interactions and the evolution of pathogenicity, facilitating the study of complex plant systems and fast-evolving bacterial pathogens and aiding in the search for therapeutic targets against plant diseases. Permanent solutions to bacterial plant diseases would be a boon for developing countries like India that are heavily dependent on agriculture for employment and revenues.

1.4 Bacterial blight of rice

About one-quarter of the total cropped area in India is occupied by rice cultivation, which contributes to approximately 43% of the total food grain production. Bacterial Blight (BB), a serious disease of rice has been reported to claim upto 50% of the total yield. Although BB is found in tropical and temperate areas alike, the disease intensity and yield loss is more in areas with higher humidity and temperatures ranging from 28-34°C (Mew, 1989).

The characteristic feature, which also renders its name to the disease, is the leaf blight. Typical blight symptoms are water-soaked or yellow lesions that appear on leaf blades of adult plants and turn yellow to white as the disease advances. Severely infected leaves tend to dry up quickly. Milky yellow droplets on leaf tips and young lesions are noticeable early in the morning, indicating bacterial ooze. A more severe manifestation of the disease is the seedling wilt or Kresek wherein entire seedlings turn pale yellow and wilt, indicating a systemic infection (Reddy, 1984). The disease is more damaging when Kresek precedes

blight. Widespread chlorosis, resulting in pale yellow mature leaves, is also very common during blight infection (Fig.1.1c; Mizukami & Wakimoto, 1969).

The causative agent of BB was identified as a bacterium by Takaishi in 1909 using isolates from the dewdrop-like ooze on infected leaves. Researchers like Bokura (1911) and Ishiyama (1922) studied the bacterium in detail and characterised its morphology, physiology and pathogenicity. The organism was initially named as *Pseudomonas oryzae*, renamed as *Xanthomonas campestris* pv. oryzae but finally classified as a new species *Xanthomonas oryzae* pv. oryzae in 1990 (Swings et al. 1990). The US Department of Agriculture recently listed this species as a potential bio-terror agent (Nature 455, 30 October 2008).

1.5 The BB pathogen *Xanthomonas oryzae* pv. oryzae and its relatives

Xanthomonas oryzae pv. oryzae, abbreviated as Xoo, is a Gram-negative capsulated motile rod-shaped bacterium, with a long monotrichous polar flagellum. Xoo colonies have a yellow and mucoid appearance (Fig. 1.1a; Bradbury, 1970a). The bacterium gains entry into the rice leaves by way of small openings on the leaf-tips called hydathodes as well as wounds on the roots or leaves (Fig.1.1b). Xoo is a vascular pathogen and once inside the vascular system, primarily the xylem, it multiplies and spreads to smaller veins. The bacterial ooze from the lesions and dewdrops are the major mode of transmission of Xoo in the fields because of splash dispersal by means of wind and rain. Xoo can only move short distances in infected crops (Mizukami & Wakimoto, 1969). Although seed transmission is not considered a particularly major way of Xoo transmission, it is the only mode of long-distance dispersal of the pathogen and is frequent enough to present a quarantine risk (Kauffman & Reddy, 1975).

Xoo belongs to a group of gram-negative, rod-shaped bacteria with monotrichous polar flagellum that constitute the genus Xanthomonas (Dowson, 1939). Most members of this genus are plant-associated and are known to cause close to 400 serious plant diseases (Chan & Goodwin, 1999). According to a classification scheme based on 16s rRNA sequence analysis, this genus belongs to phylum Proteobacteria subclass gamma. A significant feature of genus Xanthomonas is the yellow colour of the mucoid colonies due to the presence of a particular brominated aryl polyene pigment called xanthomonadin (Starr and Stephens 1964) and a mucoid appearance due to production of copious amounts of extracellular polysaccharides.

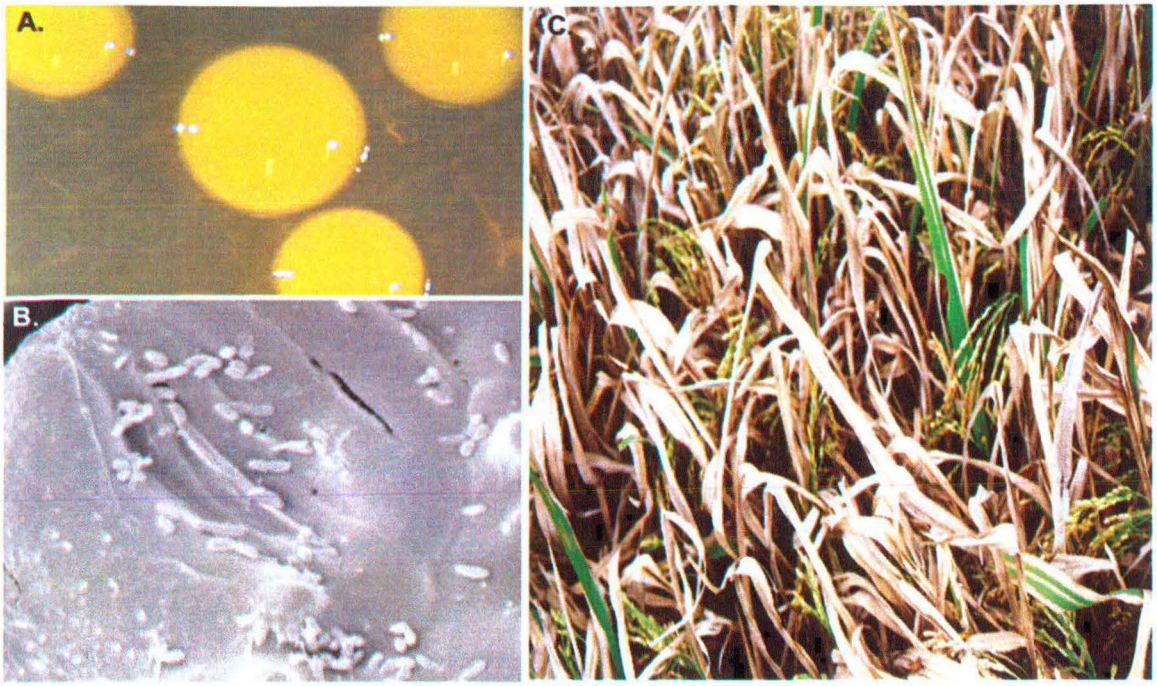


Figure 1.1. Xoo-rice pathosystem. A. Bright yellow, mucoid colonies of Xoo growing on a peptone-sucrose agar plate (Photo courtesy: Ramesh.V. Sonti, CCMB, India); **B.** Scanning electron microscopic image of Xoo cells (seen as large white spots) surrounding rice hydathodes (seen as a slit in the centre) (Photo courtesy: Amit Das & Ramesh.V. Sonti, CCMB, India); **C.** Advanced bacterial blight symptoms indicated by the yellow lesions on rice leaves (Photo courtesy: www.knowledgebank.irri.org/ricebreedingcourse).

The mode of entry in Xanthomonas pathogens is through natural openings in the host or wounds. The species that enter through hydathodes, like Xoo, colonize the vasculature while those specialized in entry through the stomata lay claim to the intercellular spaces called the mesophyll. Although there is a high level of phenotypic similarity in the members of the genus, phytopathogenic diversity is remarkable. Discussed below in some details are the phenotypic and pathogenic characteristics of certain highly pathogenic Xanthomonads that have been extensively studied and for whom, the whole genome sequence information is available:

a. *Xanthomonas oryzae* pv. oryzicola (Xoc) also infects rice but is non-vascular and colonizes the mesophyll tissue causing a disease called bacterial leaf streak (Fang et al. 1957). The characteristic difference between the two subspecies is that while Xoo enters rice through the hydathodal openings, Xoc enters rice leaves through stomata. Xoc is a mesophyllic

pathogen and causes disease symptoms that are very distinct from BB in the early stages. Greenish, water-soaked interveinal streaks of various lengths appear, spread, turn yellowish-orange to brown and eventually coalesce. In the advanced stages, bacterial leaf streak symptoms resemble BB symptoms caused by Xoo (Swings et al. 1990).

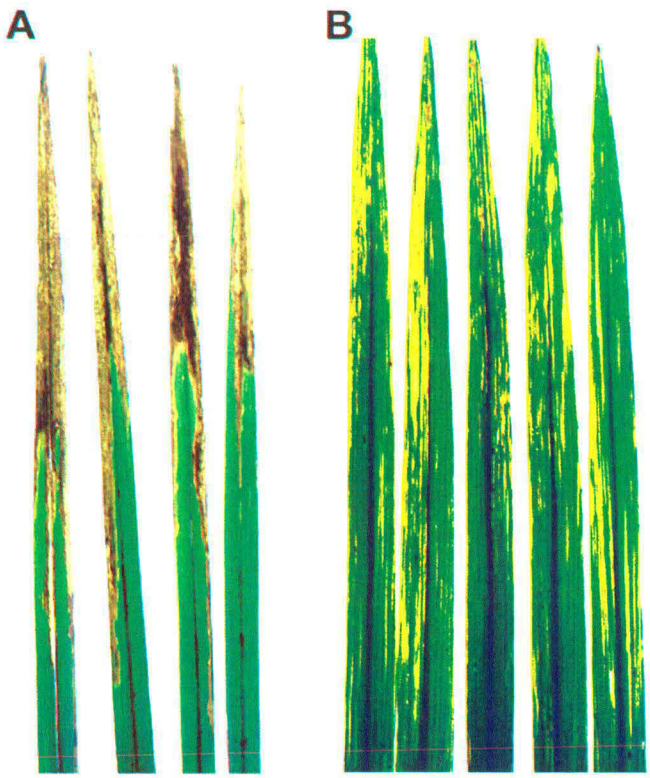


Figure 1.2. Symptoms of the diseases caused by Xanthomonas pathogens on rice. A. Bacterial Blight symptoms caused by *Xanthomonas oryzae* pv. oryzae (Xoo) on rice; **B.** Bacterial Leaf Streak symptoms caused by *Xanthomonas oryzae* pv. oryzicola (Xoc) on rice. Note that the Xoo lesions trace the leaf xylem since Xoo is a xylem pathogen while the Xoc lesions are restricted to the interveinal regions as Xoc infects the intercellular spaces between the rice parenchyma (Jha & Sonti, 2009).

b. *Xanthomonas axonopodis*: Atleast two members of this species cause important plant diseases. The first member is *Xanthomonas axonopodis* pv. citri (Xac), which is the causal agent of citrus canker, a serious disease that affects most commercial citrus production worldwide (Fig.1.3c). The other member is *Xanthomonas axonopodis* pv. vesicatoria (Xav) that causes bacterial spot of pepper and tomato (Fig.1.3d). Xav is vascular while Xac is a mesophyllic pathogen.

c. *Xanthomonas campestris*: Members of this Xanthomonas species cause very diverse plant diseases. A key member, however, is *Xanthomonas campestris* pv. campestris (Xcc), that causes black rot of crucifers such as cabbage, cauliflower etc. (Fig1.3a). Xcc also infects the model plant *Arabidopsis thaliana*, making it a scientifically very lucrative model pathogen. *Xanthomonas campestris* pv. armoraceae (Xca) is the causative agent of bacterial leaf spot of crucifers and its genome was sequenced very recently (Fig1.3b). Xcc is vascular while Xca is mesophyllic.

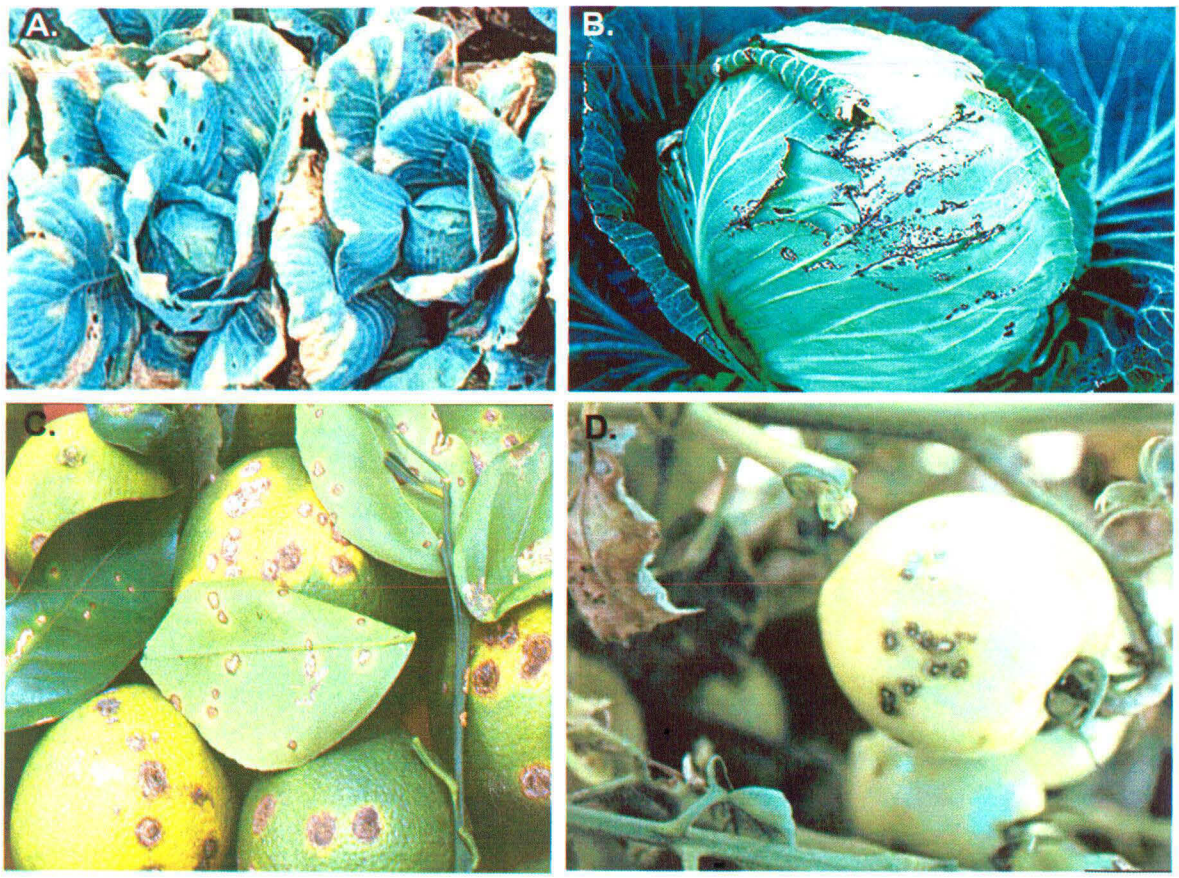


Figure 1.3. Symptoms of various plant diseases caused by genus Xanthomonas. A. The V-shaped lesions of Black Rot of Crucifers caused by *Xanthomonas campestris* pv. campestris on cabbage (Photo courtesy: Mohammad Babadoost, University of Illinois, USA); **B.** Bacterial Leaf Spot disease caused by *Xanthomonas campestris* pv. armoraceae on cabbage (Photo courtesy: Lowell L. Black, AVRDC, Taiwan); **C.** Citrus Canker symptoms on grapefruit stem, leaves and fruits, caused by *Xanthomonas axonopodis* pv. citrii (Photo courtesy: Timothy Schubert, Florida Dept. of Agriculture and Consumer Services, USA); D. Bacterial Spot on tomato plants, caused by *Xanthomonas campestris* pv. vesicatoria (Photo courtesy: David B. Langston, University of Georgia, USA).

d. *Xylella fastidiosa*: This species constitutes a group of fastidious, Gram-negative, xylem-limited bacteria causing diseases in economically important plants. The mode of transmission of *X. fastidiosa* is through xylem-feeding leafhoppers and other insect vectors. It is a major threat to grape, almond, citrus, peach and coffee crops. The symptom of *X. fastidiosa* infection is a scorched variegated chlorosis of leaves, dieback of various plant parts and stunting, leading to fruit-less or seed-less plants. *X. fastidiosa* is closely related to the genus Xanthomonas and both groups are placed in the same family, Xanthomonadaceae.

1.6 Comparison of Xanthomonas genomes

It is evident from the above discussion that there exists a vast diversity in genus Xanthomonas, with several hundred strains infecting an equally large number of host plants. Several genomic fingerprinting techniques like Amplified Fragment Length Polymorphisms (AFLP) and DNA-DNA hybridization have been employed to understand the relationship among the various Xanthomonas species (Rademaker et al. 2000). Recent surge in whole genome sequencing efforts and availability of the information in public domain has given a new dimension to comparative genomic analyses. At present, whole genome sequences are available for not only the three Xanthomonas species *Xanthomonas oryzae*, *Xanthomonas axonopodis* and *Xanthomonas campestris* but for several pathovars within each species (da Silva et al., 2002, Lee et al., 2005, Qian et al., 2005, Thieme et al., 2005, Salzberg et al., 2008 and Vorholter et al., 2008). In addition, the genome sequences of several strains of *Xylella fastidiosa* and other phytopathogens like *Pseudomonas syringae* pv. phaseolicola, *Erwinia carotovora*, *Ralstonia solanacearum* and *Agrobacterium tumefaciens* are also available. Such resources have facilitated a gene-by-gene comparison among the phytopathogens and attempts to understand these changes in the context of the evolution of plant-microbe interactions.

Intergeneric comparison of the phytopathogen genomes, in general, has revealed that certain classes of virulence determinants required for pathogen attachment to the host, plant cell wall degradation for obtaining host nutrients and suppressing host immunity are common themes in plant pathogenesis. A very interesting feature that comes across from genomic studies is the extent of variability in the bacterial genomes, brought about by horizontal gene transfer events. Mobile genetic elements like plasmids, phages and transposons carry genes coding for several virulence determinants from one genome to another and affect 'innovation and

variation' (Pallen & Wren, 2007). Genome sequencing has exposed a large number of transposable elements in almost every pathogen, explaining the frequent recombination events. Other phenomena observed commonly in certain phytopathogenic bacterial genomes (such as *X. fastidiosa*) are genome decay and extreme genome shrinkage caused by niche restrictions and development of a very specialized gene repertoire specific to plant hosts (Setubal et al., 2005).

Apart from the common pathogenesis related features, finer genome comparison studies within genus Xanthomonas have led to identification of several pathovar and species-specific genes, which have further led to interesting hypotheses explaining host and tissue (vascular or mesophyllic) specificity. Strain-specific differences have been reported in nitrate assimilation and peptide absorption, suggesting adaptation to different host plants. Similarly, differences in the repertoire of plant cell wall degrading enzyme arrays may also be indicating host-dependent divergent evolution. Host specificity may depend on a combination of differences in not only the afore-mentioned genes but also gene clusters involved in exopolysaccharide production, protein secretory pathways and various other regulatory proteins responsible for host-dependent protein expression (da Silva et al., 2002, Lu et al. 2008). However, subtle differences in certain coding regions as well as regulatory sequences may define tissue specificity (Lu et al., 2008). It is interesting to note that specificity arises not due to some major genomic rearrangements but due to more fine-tuned and directed changes in the Xanthomonas genomes.

Extensive genome sequencing has led to generation of a plethora of sequence information that needs to be catalogued and analysed. We now know about a number of putative ORFs and hypothetical genes for which function(s) need to be assigned. Large stretches of regulatory and non-coding sequences has been unearthed and experimental probing of functions is required. Genome sequence information is the seed for structural genomics, a high-throughput structure determination effort to capture snapshots of pathosystem biology.

1.7 Dissecting Xoo-rice interactions: Attack-counterattack strategies

BB is the outcome of a multitude of events ranging from Xoo entry and invasion to rice defense responses and finally the suppression of these responses to cause disease symptoms. These events are neither temporally mutually exclusive to each other, nor independent of the

host response. At every step, there are multiple regulatory controls operating at the pathosystem level. However, for simplicity in description, these events will be discussed individually. Since many aspects of Xoo infection are yet to be discovered, parallels would be drawn from other bacterial phytopathogens.

1.7.1 Xoo entry into rice

One of the most important fitness determinants in the life cycle of a phytopathogen is the ability to colonize the host. The first step in such colonization is a non-specific reversible chance attachment to both host and non-host surfaces, which is predominantly electrostatic in nature (Romantschuk, 1992). Most pathogenic bacteria can form a matrix of exopolysaccharides (EPS) on the leaf surface to anchor themselves, escape desiccation and UV radiations. Bacterial behaviour in this biofilm is different and density dependent, suggesting that the bacterial colonies have traits more directed towards infecting the host (Beattie and Lindow, 1999). Xoo has high foliar growth potential, successfully overcoming detachment threats from environmental conditions like rain, wind, temperature and establishing itself near the entry portal (Blakeman, 1991). Studies on genus Xanthomonas show that these phytopathogens are capable of both, tolerating environmental stress and avoiding it by colonizing interiors of the leaves, much better than the non-pathogens (Beattie and Lindow, 1995).

Specific adherence of Xoo to rice is not yet fully understood. However, sequence homology and knockout studies implicate various cell surface and cell wall-anchored molecules in mediating this process. In general, bacterial adhesion involves a 'dock-lock' mechanism involving bacterial surface-located proteins called adhesins and their cognate ligands on the host. Presence of multiple adhesins reiterates that adhesion is a crucial process that has to be achieved in a foolproof manner. Redundancy in adhesin proteins could also indicate a requirement to bind to different host receptors or cell types. Bacterial adhesins can be broadly categorised into fimbrial and afimbrial types. Fimbrial adhesins are wiry or rod-shaped hair-like appendages. The Type I fimbriae (referred to as fimbriae) have been shown to mediate attachment to host cells or surfaces while the Type IV fimbriae (also called pili) are implicated in twitching motility of the bacteria involving adhesion to solid substratum (Smyth et al., 2006). Motility seems to be an important virulence attribute of Xanthomonas and helps

escape environmental stress (Esseaberg, 1979). Xoo genome shows the presence of gene clusters encoding for both fimbriae and pili.

Xanthomonas adhesin-like proteins A and B (XadA and XadB; Ray et al., 2002 and Das et al., 2009) and a homolog of Yersinia autotransporter protein H (YapH; Das et al., 2009) perform afimbrial adhesin-like functions in Xoo. Although the role of each of these adhesins and respective ligands are not known, it appears that these proteins act in conjunction with each other and can convert initial passive attachment into a (host) specific adhesion. XadA and XadB appear to be important in leaf colonization and entry while the Type IV pili are involved in promoting growth within the rice xylem vessels. YapH appears to play a role in both stages of the infection process (Das et al., 2009).

1.7.2 Detection of breach and rice pre-alert

A very interesting observation regarding plant and microbe evolutionary time scales is that early land plant evolution was facilitated by interaction with fungal symbionts. This suggests that plants have had a long-standing exposure to various kinds of microbes and this close association must have led to evolution of pathogenesis (Chisholm et al., 2006). Moreover, considering how common the exposure to phytopathogens is, the frequency of occurrence of plant diseases is much less, indicating evolution of various defense mechanisms by plants (Abramovitch et al. 2006). These defense strategies have been identified to be not only *in planta* but also present on the aerial surfaces (Shepherd and Wagner, 2007). The first line of rice defense is a physical barrier called cuticle, composed of an insoluble cuticular membrane impregnated by and covered with soluble waxes that are mixtures of C16 to C36 hydrocarbons and polyester epoxide polymers (Bianchi, 1979). Rice leaves contain stalked multi-cellular projections called trichomes that might secrete certain anti-microbial terpenoids and phenylpropanoids reported to be present in rice hulls. Cationic peroxidases found in the rice guttation fluid, the xylem exudate coming out of the hydathodal openings and looking like dew drops, are also implicated as being anti-microbial (Young et al. 2007). Together, the rice leaf surface is a complex biological barrier that can be crossed only by proficient phytopathogens, unlike the nonspecific non-pathogenic populations.

Phytopathogens and certain non-pathogenic plant-associated bacteria like rhizobia cross the surface barriers efficiently when the environmental conditions are conducive. Detection of

common pathogen- or microbe-associated molecular patterns (PAMPs/MAMPs) that are not found in host cells serves as the trigger for plant innate immune system (Boller, 1995). Some of the bacterial PAMPs recognized by the plants include lipopolysaccharides (LPS) and flagellin protein among others and are essential for its fitness, survival and virulence. When PAMPs are perceived by cell surface-located pattern-recognition receptors (PRRs) of plants, massive coordinated reprogramming in the gene expression occurs, triggering convergent defense responses. Plants, in general, show a surge in cytoplasmic Ca^{2+} levels, production of reactive oxygen species (ROS), nitric oxide (NO) and certain antibiotic-like compounds called phytoalexins, deposition of callose and lignin, polymeric phenolics that reinforce the already rigid cell wall to prevent bacterial entry (He et al., 2007). As in animal systems, activation of MAP kinases mediates all the responses downstream of PAMP-PRR interactions (Nurnberger et al. 2004). The PAMP-Triggered Immunity (PTI) is a deterrent for non-pathogens only, while pathogens like Xoo have evolved efficient ways of suppressing these defense responses (Chisholm et al. 2006).

1.7.3 Suppression of rice innate immune responses by Xoo

The ability of pathogenic bacteria to use proteinaceous apparatuses called secretion systems to pump proteins into either the extracellular milieu or directly the eukaryotic host cells is an important virulence attribute (Chisholm et al., 2006). The secretion systems can be divided into seven groups based on the construction of the system itself and its mode of protein secretion across inner and outer membranes (Saier, 2006). The Type III secretion system (T3S) is chiefly responsible for the secretion of 'effectors' that suppress the PAMP-triggered immune responses (Alfano and Collmer, 2004, Grant et al., 2006). The T3S effectors suppress innate immune as well as pathogen-induced specific defense responses by inhibiting eukaryotic cellular functions by mimicking the structures of crucial host enzymes, targeting host proteins for degradation, affecting host intracellular transport and reprogramming host transcription (Jones and Dangl, 2006). Upon disruption of T3S, Xoo is unable to cause disease and is restricted after elicitation of basal defense responses from rice (Zhu et al., 2000, Jha et al., 2007). Interestingly, priming with $T3S^-$ strain can induce resistance to further Xoo infections, suggesting that in the absence of T3S effectors, localised innate immune responses of rice are efficient in preventing pathogen multiplication and spread. Several Xoo effectors have been identified and are yet to be characterized functionally.

1.7.4 Active invasion of rice tissues by Xoo

The third factor that has a major influence on the fitness of a bacterial phytopathogen is invasiveness, the ability to enter host tissues by force. The Type II secretion system (T2S) is involved in secreting the enzymes responsible for plant cell wall degradation, the pathogenic approach to invade plants (Cianciotto et al., 2005). Xoo genome shows an abundance of T2S-secreted enzymes like carbohydrate hydrolases, proteases, esterases etc. that can function as rice cell wall degrading enzymes (Jha et al., 2005). Several of these enzymes have been characterized and shown to have important virulence functions in Xoo pathogenesis (Jha et al., 2007, Rajeshwari et al., 2005, Hu et al., 2007). These enzymes serve to disrupt the intricate meshwork of cellulose, hemicelluloses, other complex polysaccharides and phenolics that constitutes the tough and rigid cell wall, perhaps for facilitating the movement of Xoo through sieve elements around the xylem vessels. Disruption of T2S gene cluster causes severe virulence deficiency in Xoo (Ray et al. 2000). Mutations in genes coding for the cell wall degrading enzymes such as a lipase/esterase (LipA), xylanase (XynB) or cellulase (ClsA) of Xoo cause partial loss of virulence, while *lipA xynB* and *lipA clsA* double mutants have much more severe effects on virulence (Rajeshwari et al. 2005; Jha et al. 2007). A cellobiosidase (CbsA) deficient mutant of Xoo is also severely virulence deficient (Jha et al. 2007). It is evident that while some enzymes may have a more important role than the others may, a concurrence in secretion and action is essential for Xoo virulence. Interestingly, preliminary sequence analysis shows that there are interesting differences between these enzymes and the corresponding characterized homologs. Detailed structural analyses and biochemical studies with the various combinations of enzymes will uncover the specific mechanistic roles and *in planta* substrates of these enzymes.

A $T2S^-$ $T3S^-$ double mutant of Xoo is severely virulence deficient and compromised in elicitation of any defense response from rice. This suggests that the basal defense responses induced by a $T3S^-$ mutant are mostly due to proteins that are secreted through the T2S. Consistent with this possibility, purified T2S secreted proteins induce basal defense responses when infiltrated into the rice tissues. These defense responses are induced by soluble elicitors that are released by the action of these cell wall degrading enzymes on rice cell walls (Jha et al., 2007). Induction of plant defense responses by oligosaccharide derivatives, released by the action of cell wall degrading enzymes, was first reported by Darvill and Albersheim (1984).

1.7.5 Specific defense responses of rice against Xoo

Innate immune responses are the simplest form of non-self recognition by plants. As the intensity of pathogenic invasion increases, plants have evolved a systematic retort called hypersensitive responses (HR), a spree of programmed cell death (PCD) leading to pathogen limitation in the immediate vicinity of infection and activation of defense genes (Goodman and Novacky, 1994). Detection of cell wall degradation products, mostly oligosaccharide derivatives, produced by bacterial enzyme action along with the other PAMPs elicits HR (Darvill and Albersheim, 1984). This form of plant resistance is termed Effector-Triggered Immunity (ETI), a faster and stronger version of PTI.

The trigger for ETI is cognate recognition of specific avr proteins by the plant resistance or '*R*' gene product (De Wit, 2007). R proteins often have nucleotide binding domains and contain leucine-rich repeats (NB-LRR type) (Dangl and Jones, 2001). NB-LRR activation results in a network of cross talks between various signalling pathways deployed to curb pathogen attack. Several resistance loci of rice act against different Xoo strains (Nino-Liu et al., 2006). Resistance is a result of interplay between salicylic acid (SA), a local and systemic signal, jasmonic acid (JA) and ethylene accumulation. SA-dependent signalling activates expression of certain antimicrobial proteins that provide resistance against diverse microbial pathogens. The JA and ethylene signalling pathways function in defense against general wounding or insect attack. Prevention of the accumulation of plant hormone Indole-3-acetic acid (IAA) also induces resistance in rice against Xoo infection. This induced resistance seems to be independent of JA or SA defense signalling pathways (Ding et al., 2008).

1.7.6 Disease

Disease is an outcome when a pathogen infects a susceptible host, resultant of a 'compatible interaction' and occurs when the host suffers sufficient damage to perturb homeostasis (Casadevall & Pirofski, 1999). Resistant hosts or avirulent pathogens can effect only 'incompatible' interactions and thus, no disease. Studies demonstrate that Xoo cells can be successfully enter into resistant rice cultivars but lose their shape and integrity soon (Horino, 1981). Interestingly, xylem vessel wall thickening occurs within 48 hr of inoculation with a Xoo strain that gives rise to an incompatible reaction and not with a strain that gives a compatible interaction (Hilaire et al. 2001). In general, pathogens have several ways of

overcoming ETI. Firstly, effector genes are often associated with mobile genetic elements and are commonly observed as present or absent (polymorphic) across different strains. This helps in rapid gain and, when necessary, rapid loss of the effectors. It is also seen in some cases that the host might act as pathogen-induced modified self, and fail to interact with the effectors. ETI can also be overcome through evolution of pathogen effectors that suppress SA signalling by up-regulating JA pathway (De Wit, 2007).

Microbial evolution in response to ETI causes some NB-LRR gene loci to evolve very rapidly. In pathogen populations, the frequency of an effector gene is enhanced by its ability to promote virulence, and reduced by host recognition. Hence, natural selection should maintain effector function in the absence of recognition. Thus, disease ensues when the pathogen wins the evolutionary conflict while resistance occurs when R genes suppress all the pathogenic effectors (Jones and Dangl, 2006). Several studies are being undertaken to identify and characterize the T3S effectors of Xoo (Furutani et al., 2009) and it would be interesting to understand the exact way(s) in which Xoo can overcome rice defense responses.

1.8 Unresolved questions and prospects: Utilizing protein structures

Thorough genetic analysis has lead to the identification of a large repertoire of gene products involved in the Xoo-rice interaction. However, our understanding of the molecular mechanisms involved at each step is far from complete. Modes of action of adhesins, T2S and T3S effectors, respective host targets and events downstream of the ligand-receptor interactions are still major challenges of this field, owing to the complexity of the plant system and rapid evolution of bacteria. Although there are common themes throughout biology, every individual phytopathosystem has evolved its own nuances and specializations. In molecular terms, whereas major protein fold families are common to all kingdoms of life, subtle adaptations lead to various domain architectures that are specific to individual genomes, and contribute to the diverse phenotypes observed (Lee et al., 2003). Identifying specific mediators for each phytopathogen is the key to designing therapeutic agents for the resultant diseases, similar to the animal pathosystems.

While genomics has huge potential of detecting characterized domains and motifs in newly sequenced proteins and thus, annotate and assign function, it is also true that there are several uncategorized domains, which are yet to be allocated biological roles. *Ab initio*,

characterization of a given protein is still a difficult problem. However, powerful tools such as X-ray crystallography, nuclear magnetic resonance (NMR) and cryo-electron microscopy allow detailed investigations, providing insights into protein structure, interactions with other proteins or small molecules and therefore, the function. Structures of several adhesins from other Gram-negative animal pathogenic bacteria have already been solved and these provide useful insights and structural templates for several other highly conserved adhesins and pili. Structures of important elements of protein secretory systems and their secreted proteins have been solved from *Pseudomonas syringae*, *Erwinia carotovora* and *Agrobacterium tumeifaciens* (Remaut & Waksman 2004). Genome sequences and proteome analysis of Xanthomonas reveals a large number of T2S enzymes and T3S effectors that possess limited sequence similarity to characterized proteins, emphasizing on the need for structural investigations.

Elucidating structure-function relationships of various Xanthomonas proteins would unravel molecular details of Xoo-rice interactions and prove to be advantageous in more than one ways. BB is still rampant because it lacks predictability and spreads uncontrollably upon initiation. Prevention of the disease would arise from the knowledge of the proteins involved in the arrest of Xoo growth in both, genetically engineered and naturally occurring resistant rice cultivars. Studying Xoo T3S effectors that are required for overcoming rice responses would help identify therapeutic targets. Similarly, T2S effector structures can provide insights into novel enzyme functions, evolved for specific action on plant substrates. Protein structures can also provide important information on the phylogeny of the organisms (Deeds & Shakhnovich, 2006).

1.9 Scope of the current study

This thesis work deals with the structural and functional studies on two important virulence determinants of Xoo. The two major events of Xoo-rice interplay considered are the attachment and entry into the hydathodal openings of rice leaves and degradation of the rice cell wall during invasion of rice tissues. The *in silico* analysis of the Xanthomonas adhesin-like protein A (XadA) helped understand the domain architecture of this adhesin, based on its homology to adhesins from animal pathogenic bacteria. The striking observation that XadA is expressed only in the presence of glutamate in the culture medium, the dissection of molecular details of XadA regulation and the implications of XadA regulation on our

understanding of the entry of Xoo into rice has been examined. Crystal structure of the secretory esterase LipA revealed a unique mode of substrate recognition enclosed in a ligand-binding domain, present as a functional attachment to the canonical catalytic hydrolase domain. The mechanistic mode of action, essential role of the ligand-binding domain in Xoo pathogenesis and evolutionary consequences as derived from the structure are presented. Overall, the work aims to utilize X-ray crystallography, biophysical, biochemical, genetic and bioinformatics tools to find a few important missing links in our understanding of the Xoo-rice pathosystem.

Chapter Two

Crystal structure determination of a rice cell wall degrading esterase from *Xanthomonas oryzae* pv. oryzae

CHAPTER TWO

2.1 Abstract

Xoo secretome contains an array of hydrolyzing enzymes to degrade the rice cell walls. LipA is a secretory virulence factor of Xoo, implicated in degradation of rice cell walls and concomitant elicitation of innate immune responses such as callose deposition and programmed cell death. In order to understand the basis of LipA function in the context of Xoo pathogenesis, LipA structure was solved by Multiple Isomorphous Replacement method, using five heavy atom derivatives of the plate-like crystals of LipA. This chapter deals with the structure determination of LipA and the features of the wild-type structure. The final structure was refined to R/R_{free} values of 16.2/18.8% at 1.86Å resolution. LipA is an α/β hydrolase-fold protein with a 9-stranded central mixed β-sheet surrounded by α helices in a typical hydrolase topology and a canonical catalytic Ser-His-Asp triad. The distinctive feature of LipA structure is the presence of a 108 amino acid domain present as an insertion in the catalytic domain, corresponding topologically to the lid-domain insertion position of triacylglycerol lipases. This insertion domain consists of seven α-helices with one 3_{10} helix and shows no structural homologs in structure databases. Consequently, LipA shares only a distant sequence and structural homology with other characterized lipases and esterases.

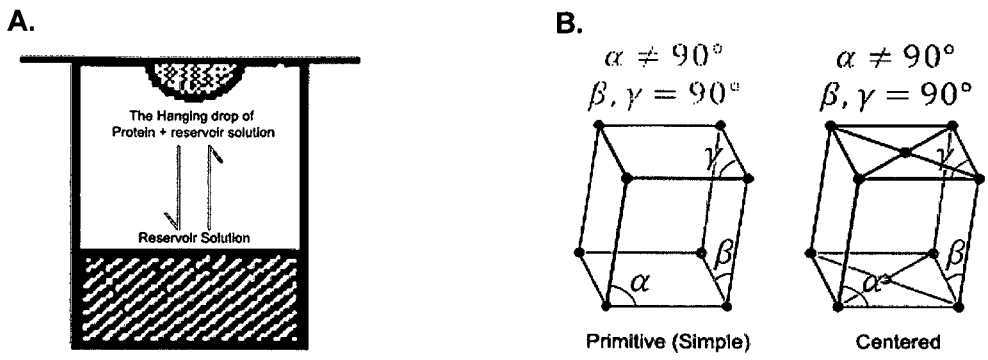
2.2 Introduction

The chemically complex plant cell wall acts as a major barrier, restricting pathogen movement and access to the plant resources. The ability to secrete specific hydrolytic enzymes to sever different components of the plant cell wall is a major virulence attribute of phytopathogenic fungi and bacteria (Albersheim et al., 1969). While cellulases, xylanases, polygalacturonases and pectate lyases degrade the main polysaccharide constituents of the plant cell walls, enzymes like pectin esterases cleave the ester crosslinks between the polysaccharide fibrils and loosen the cell walls (Esquerre-Tugaye et al., 2000). Cell wall damage not only provides a point of entry to the pathogen but also acts as a mark of invasion that the plant can sense. Degradation products that are released by the action of cell wall degrading enzymes on plant cell walls can induce potent innate immune responses of plants (Darvill and Albersheim, 1984; Ryan and Farmer, 1991). These defense responses include the synthesis of anti-microbial compounds such as phytoalexins, strengthening of plant cell walls through deposition of callose as well as programmed cell death reactions. Thus, it appears that the cell wall degrading enzymes serve dual functions. They are required for virulence but their activity also induces host defense responses. Successful microbial plant pathogens are able to overcome this difficulty because they have the capability to suppress these innate immune responses (Jha et al., 2005).

Xoo employs a battery of enzymes that includes a lipase/esterase (LipA), cellulose (ClsA), xylanase (XynB) and cellobiosidase (CbsA) to degrade rice cell walls (Rajeshwari et al. 2005; Jha et al. 2007). Mutations in the genes coding the individual secreted proteins such as LipA and ClsA cause partial loss of virulence, while a double mutant in the *lipA* and *clsA* genes results in a more severe loss of virulence, suggesting a redundancy in the functioning of these proteins. Treatment of rice leaves and roots with LipA, ClsA and CbsA proteins induces rice defense responses such as callose deposition and programmed cell death. The actual elicitors of the defense responses appear to be soluble cell wall degradation products that are released following the action of these enzymes on rice cell walls (Jha et al., 2007). This is in sharp contrast to the previously described fungal xylanases, which have been shown to function as direct elicitors of defense responses through recognition of a particular sequence motif in the xylanase by a specific plant receptor (Ron & Avni, 2004).

LipA is a 42 KDa protein with tributyrin and TWEEN-20 degrading activity, a general indicator of lipase/esterase function (Rajeshwari et al., 2005). LipA homologs are present in several gram-negative bacteria, including all xanthomonads whose genomes have been sequenced. The LipA homologs are poorly characterized and categorized as conserved hypothetical proteins or putative 'secretory' lipases, indicating that these proteins have a lipase/esterase activity different from feruloyl esterases, pectin esterases and cutinases reported to act on plant substrates (Bauer et al., 2006, Vorwerk et al., 2004). In general, very little is known about the structure-function relationships of secreted enzymes leading to plant cell wall degradation and pathogenesis. The role of Xoo secreted enzymes in degradation of rice cell wall is implied by the genetic studies; however, the functional intricacy of each of these enzymes is largely uncharacterized. In an effort to understand the role of a lipase/esterase activity in Xoo pathogenesis of rice, we obtained the high-resolution structural information on LipA using X-ray crystallography.

X-ray crystallography is a method of determining the arrangement of atoms within a crystal from the X-rays scattered by the crystal, producing a three-dimensional view of the electrons density within the crystal. From this electron density, information regarding the mean positions of the atoms, the interatomic chemical bonds, bond-lengths and bond-angles can be obtained. The foremost requirement for this study is crystallization of the given protein. A $\geq$ 95% pure protein of interest, when subjected to a controlled precipitation event, starts to aggregate into clusters, forming nuclei for crystallization (McPherson, 1999). The rate of nucleation and growth of the crystal is driven by the super-saturation of protein in the solution, which in turn is dependent on the protein concentration, solution ionic strength, temperature, pH etc. (Fig. 2.1a). The protein atoms pack in a 3-dimensionally defined and periodic manner as a lattice. The crystal lattice symmetry is defined in terms of the angles (α, β and γ) and proportions of the three axes (a, b and c) of the smallest repeat unit of the crystal or the 'unit cell'. The simplest and the least symmetric triclinic crystal system has $\alpha, \beta, \gamma \neq 90^{\circ}$ and $a \neq b \neq c$. Other crystal systems in ascending order of symmetry are monoclinic (to which LipA crystals belong; Fig. 4.2.1b), orthorhombic, rhombohedral, tetragonal and hexagonal. The crystal systems provide the matrix points but there can be several other symmetry operations such as the lattice centring and screw axes, which together define the 'space group' of the crystal (Rhodes, 2000).



ure 2.1. Schematic of a hanging drop vapour diffusion setup. A. The reservoir solution contains recipitant, buffer and certain additives like salts and mixed in a 1:1 ratio with the protein in the ging drop. **B.** Diagrammatic representation of the two types of monoclinic crystal systems, nitive and a centered cell. (Adapted from Rhodes, 1993)

·ays are generated by a cathode ray tube, filtered to produce monochromatic radiation, limated to concentrate and directed towards a well-ordered protein crystal fixed on a ıiometer and flash-frozen in a cryo-stream based on liquid-N_2. The crystal, in the presence a cryo-protectant, is slowly rotated in a stream of X-rays using the goniometer and the ultant scattered or diffracted X-ray beams are allowed to hit a detector and the set of spots 'reflections' emanating at each angle of rotation are collected as individual 'frames'. The stal is the 'real space' while the diffracted reflections represent the 'reciprocal space'. ery reflection is a result of constructive interference of the diffracted X-rays from every m of the crystal. The condition of a constructive interference is stated by the famous ıgg's law: $n\lambda = 2d\sin\theta$, where n is an integer number of the order of planes in the crystal ice, λ is the wavelength of X-rays, d is the spacing between the planes in the atomic lattice l θ is the angle between the incident ray and the scattering planes. Although several emitted ves may result from constructive interference, only waves that cross the 'Ewald sphere' of ius $1/\lambda$ passing through the origin of the reciprocal lattice and the centre concurrent with centre of the crystal, produce reflections (Drenth, 1999). The reciprocal lattice intercepts he Ewald sphere are called 'Miller indices' h, k, l (Fig. 2.2).

ɜ crystal properties, such as space group or the quality of packing ('mosaicity'), determine data collection parameters. The detector records and transfers the two-dimensional images individual frames to a computer. 'Indexing' of the data set involves allocating a 'Laue up', also called point group or crystal class, which is one of the distinct combinations of

the point operations of rotation about an axis, reflection in a plane, inversion about a centre, or sequential rotation and inversion. The addition of translational changes will yield 230 possible combinations that are referred to as 'space groups'. Moreover, indexing also assigns unit cell dimensions to the unit cell and taking into account the variations inherent in each of the multiple images. The reflections are measured as 'intensities', proportional to the squares of 'structure factor amplitudes'. The next step is to 'merge' and scale the multiple files into a single file containing information of all intensities. The data quality is judged at this step by assessing how well the different frames can be merged (R_{merge}), an estimate of the errors during scaling and the completeness of data.

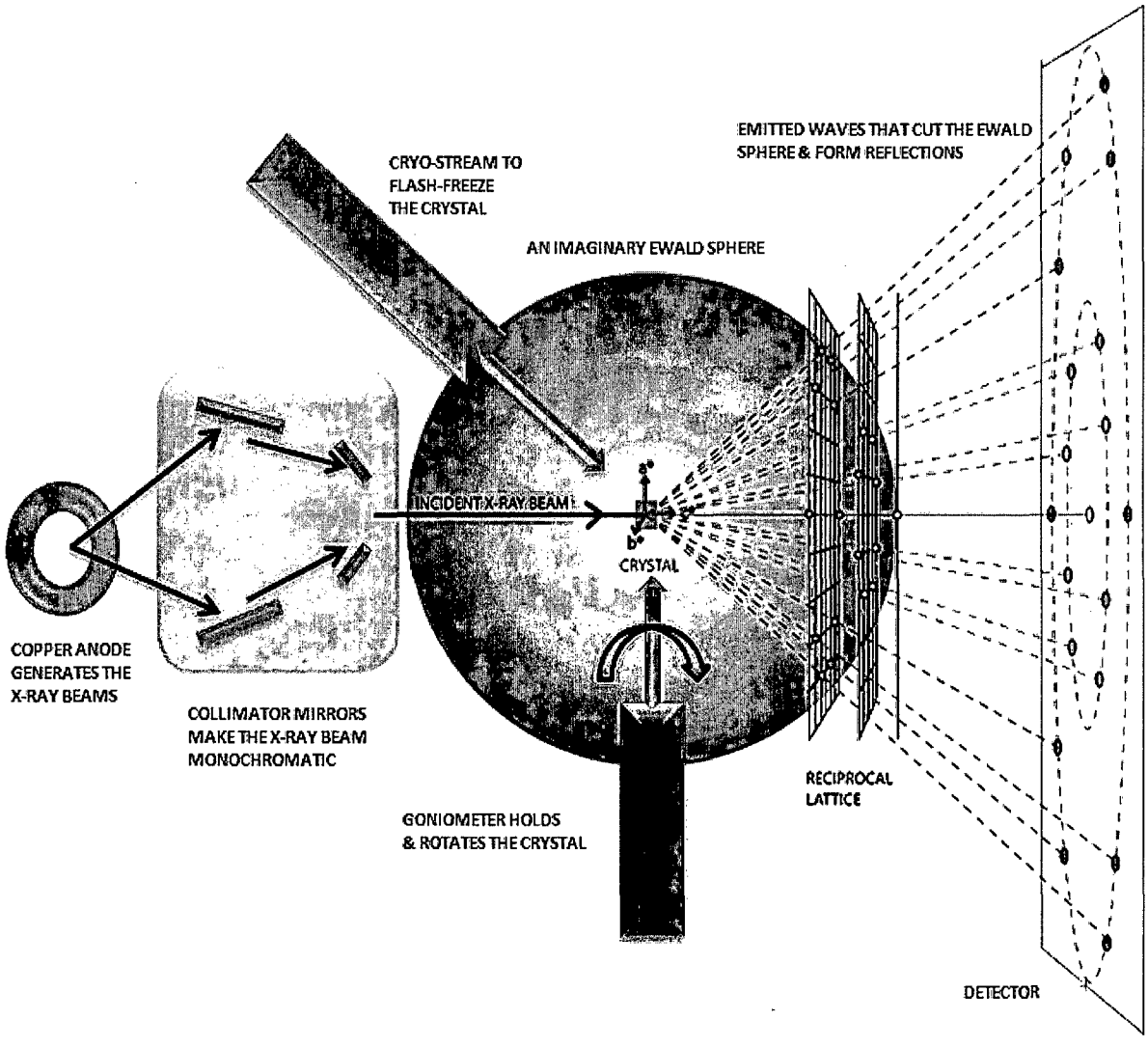


Figure 2.2. A simplified schematic of the process of in-house X-ray diffraction and data collection. (Modified from cryo-crystallography and data collection tutorials at www.doe-mbi.ucla.edu/~sawaya/m230d/Data/data.htm)

The structure factor is a complex number containing information relating to both the amplitude and phase of a wave. However, during data collection, the only information obtained in the form of intensities is the amplitude of the waves and not the phase. This loss of phase is termed 'phase problem' and the main aim of structure solution is to estimate the correct phases in order to reconstruct the model of the protein molecule (Taylor, 2003). Initial phase estimates can be obtained in a variety of ways:

a. **Molecular replacement**: A homologous structure with > 20% identity can generally be used as a search model in molecular replacement to determine the orientation and position of the molecules within the unit cell.

b. **Anomalous X-ray scattering** (Multi/single-wavelength Anomalous Dispersion phasing; MAD/SAD): A heavy atom that strongly absorbs radiation, when incorporated into the protein and exposed to X-rays at a wavelength past the absorption edge of the anomalous scatterer, causes changes in the scattering of the protein (anomalous dispersion). By recording full sets of reflections at three different wavelengths (far below, far above and in the middle of the absorption edge) one can solve for the substructure of the anomalously diffracting atoms that alter the amplitude and phase of the emitted waves and thus, the structure of the whole molecule. Incorporation of selenium as the anomalous scatterer at Methionine positions in the protein by expressing the protein in a methionine auxotroph strain in a media rich in Seleno-methionine is the method of choice (Ealick, 2000).

c. **Multiple Isomorphous Replacement** (MIR): Soaking the crystals in an electron dense metal solution can incorporate these heavy atoms into the protein molecule without much change in the crystal lattice, isomorphously. As in MAD phasing, the changes in the scattering amplitudes can be used to yield the phases.

d. ***Ab initio* phasing (direct methods)**: This is usually the method of choice for small molecules (<1000 non-hydrogen atoms), and has been used successfully to solve the phase problems for small proteins. If the resolution of the data is better than 1.4Å, direct methods can be used to obtain phase information, by exploiting known phase relationships between certain groups of reflections. (Blundell and Johnson, 1976)

The phase information obtained from any of the above methods and the structure factor amplitudes are used to construct a 3-dimensional 'electron density map', placing the atoms in 3-D space to best match the electron density and this is the initial model. This model can be used to refine the phases, leading to an improved model in an iterative manner. Given a model of some atomic positions, these positions and their respective B-factors (representing the thermal motion of the atom) can be refined to fit the observed diffraction data, ideally yielding a better set of phases. A new model can then be fit to the new electron density map and a further round of 'refinement' is carried out. Two factors, 'R and R_{free}', measure the agreement between the diffraction data and the final model and thus, signify the quality of the model. Ramachandran plot to estimate the total number of residues with allowed phi-psi angles and calculation of r.m.s.deviations in bond lengths and bond angles are parameters used for the assessment of the structure.

Structure of LipA was solved using 1.86Å resolution X-ray diffraction data collected in-house with the Multiple Isomorphous Replacement method of phasing. Attempts to solve LipA structure using molecular replacement failed due to the lack of high sequence homology with characterized proteins. This chapter deals with the purification of LipA from Xoo supernatants, crystallization, preparation of heavy atom derivatives, structure solution using MIR, model building and structure assessment along with a description of the wild-type LipA structure.

2.3 Materials & Methods

2.3.1 Overexpression of LipA

Xoo strains were grown at 301K in peptone-sucrose (PS) medium (Tsuchiya et al., 1982). The antibiotics used in this study were rifampicin at 50mg l^{-1} and spectinomycin at 50mg l^{-1}. Xoo strain BXO2001 has an insertion mutation in the lipA gene (Rajeshwari et al., 2005). This mutation abolishes secreted lipase/esterase activity. Introduction of the *lipA* gene into BXO2001 in the broad host-range vector pHM1 (to createBXO2008) not only restores the lipase activity but also results in the overproduction of lipase (Primers used: **LipAFP**: 5'CCGGAATTCGCACCCAGCGCGCAGGCAGCACCC3';**LipARP**:5'CGGGGTACCCCA ACAGCTGATCCCGCACCAC3' at *Kpn*I and *Hind*III sites). The overproduction of lipase by BXO2008 may be because the *lipA* gene in pHM1 is under the control of a constitutively

expressed *lac* promoter. Wild-type *X. oryzae* pv. oryzae secretes the LipA protein into the extracellular medium using the T2SS (Rajeshwari et al., 2005). The BXO2008 strain is proficient for the T2SS and it is expected that the overexpressed lipase be secreted into the medium through the same system. Approximately 1 mg LipA is secreted into the medium by 1l wild-type Xoo. The BXO2008 strain secretes ~25-30mg LipA per litre of culture.

2.3.2 Purification of LipA from Xoo culture supernatants

The BXO2008 strain was grown to saturation in 1l PS medium at 301 K. The culture was centrifuged at 12,000 r.p.m at 283K for 30min to separate the Xoo cells. The culture supernatant was subjected to 55% ammonium sulphate precipitation and the precipitate obtained was dialysed against 10mM potassium phosphate buffer (PB) at pH 6.0. The dialysed supernatant was passed through a Mono S column for cation-exchange chromatography. The peaks containing pure LipA were pooled and loaded onto a 24ml Superose-12 gel-filtration column (GE Pharmacia, USA) pre-equilibrated with 10mM PB pH 6.0. The purified protein was dialyzed against 20mM Tris-HCl pH7.5, 20mM NaCl and concentrated to 5mg ml^{-1} using a10kDa Amicon Ultra-15 filtration device (Millipore, USA). Protein concentration was determined using BCA reagent (Pierce, USA).

2.3.3 Crystallization

A set of 98 predesigned conditions that vary the precipitants and their concentrations, buffers and their pH along with salts and their concentrations, composing the Crystal Screens I and II (Hampton Research, USA) were used to screen for initial LipA crystallization conditions. The drops were set up using the hanging-drop vapour-diffusion method by mixing equal volumes (2ml) of protein solution and reservoir solution at 298K. This method involves a droplet containing purified protein, buffer and precipitant being allowed to equilibrate with a larger reservoir containing similar buffers and precipitants in higher concentrations under airtight condition (Fig. 2.1a). Initially, the droplet of protein solution contains an insufficient concentration of precipitant for crystallization, but as water vaporizes from the drop and transfers to the reservoir, the precipitant concentration increases to a level optimal for crystallization. Since the system is in equilibrium, these optimum conditions are maintained until the crystallization is complete (Rhodes, 1993). Small needles and clusters were obtained within 24h in Crystal Screen I condition Nos. 28, 37 and 41, and Crystal Screen II

conditions 35 and 38. Refining the initial conditions with PEG 400 and PEG 6000 as precipitants and 0.10M MES buffer pH 6.0-7.0 yielded well-diffracting crystals.

2.3.4 Preparation of heavy atom derivatives

In order to obtain phase information for the diffraction data collected from LipA crystals, multiple heavy atoms derivatives were generated for the reasons discussed below. In the simplest of cases, structure factor for the derivative crystal (F_{PH}) is equal to the sum of the protein structure factor (F_P) and the heavy atom structure factor (F_H). When considered as vectors, the three terms form the following triangle ('Argand diagram' in Fig.2.3):

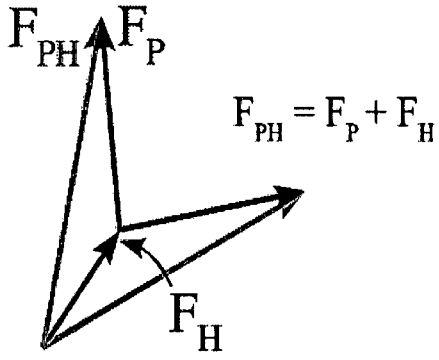


Figure 2.3. Argand diagram of the vectors representing the structure factors of native and heavy atom derivative data. (Adapted from Rhodes, 2000)

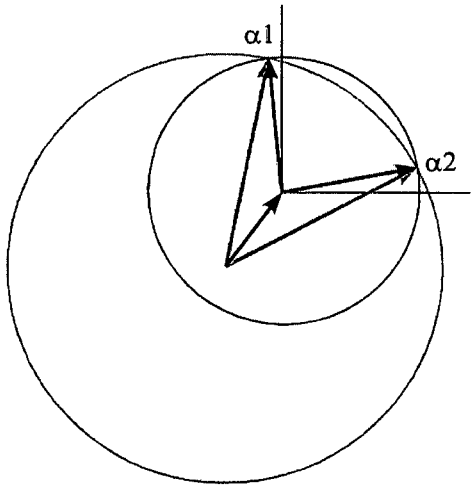


Figure 2.4. Harker construction to show the two possible phases. The bigger circle has a radius equal to the amplitude of F_P, centered at the origin and indicates all the vectors that would be obtained with all the possible phase angles for F_P. The smaller circle with radius F_{PH} represents all the possible values for F_P. $\alpha 1$ & $\alpha 2$ are two possible values for F_P. (Adapted from Howard & Brown, 2002)

However, there can be two ways of constructing this triangle as is evident in Fig.2.4. A circle with a radius equal to the amplitude of F_P centered at the origin (bigger circle) indicates all the vectors that would be obtained with all the possible phase angles for F_P. Similarly, the smaller circle is all the possible phases for F_{PH}. There would always be a two-fold phase ambiguity ($\alpha 1$ & $\alpha 2$, in this case), except when the crystal space group is centro-symmetric and has a centre of inversion. Multiple derivatives that bind to different sites on the protein molecule are prepared to resolve this phase uncertainty since there would be only one phase of choice consistent with all observations.

High throughput crystallization facility (Alchemist screen making system and Minstrel High-throughput crystal imaging system, Rigaku Corp., USA; Matrix Hydra eDrop, Thermo Fisher Scientific, USA) was used to prepare crystallization solutions and generate several LipA crystals for heavy atom derivatization. Isomorphous crystals soaked in heavy atoms (Hampton, USA) like 10mM K_2PtCl_2 for 2-15 min and 10mM Sm $(NO_3)_3$ soaks of 20-25 min were used for heavy atom substitution. Long soaks of 1hr was used for the Hg salt thimerosol to obtain substitution. Although several heavy atoms such as Ag salts, Au salts, Iodides etc. were attempted for soaking, no substitution was obtained. Therefore, only the derivatives for which successful substitutions were obtained will be discussed further.

2.3.5 Data collection and processing

An in-house MAR Research MAR345dtb image-plate detector and Cu Kα X-rays of wavelength 1.54Å generated by a Rigaku RU-H3R rotating-anode generator were used to collect diffraction data. The crystal was mounted on a nylon loop and flash-cooled in a nitrogen-gas stream at 100 K using an Oxford Cryostream system. No cryo-protectant was used for the type I crystal. 15% glycerol in mother liquor was used as a cryo-protectant for the type II crystal. Data were collected with an oscillation angle of 0.5° and an exposure time of 600 s for each image. A total of 120° and 180°of oscillation data was collected for the type I and type II crystals, respectively. The heavy atom-soaked type II crystal data was also collected similar to the wild-type type II crystal data. The software that facilitated data collection and transfer of the frames to a computer as image files was Mar345 v.1.2.8, with a constituent program AutoMar aiding in preliminary assessment of the collected frames. Indexing, scaling and merging of the data were performed using softwares DENZO and SCALEPACK, where the image files are first converted into '.x' files containing the

intensities of the recorded reflections. The intensities are proportional to the square of structure factor amplitudes (F_P for native data; F_{PH} for heavy atom data). The .x files are then one merged '.sca' file (Otwinowski & Minor, 1997).

In order to check for heavy atom substitution, ~10° of oscillation data was collected for every derivative crystal and processed. For a complete data collection, these crystals were required to satisfy two conditions. Primarily that these are isomorphous for symmetry, dimensions and contents (with the exception of heavy atom addition) when compared with the wild-type crystals. Secondly, the χ^2, an estimate of the difference between the wild-type and that the substitution is neither too high nor too low, indicating a change but not too large to cause anisomorphy. The χ^2 was derived using a program called 'scalepotderiv.com' in SCALEPACK. Only when such derivatization was obtained, a complete data was collected.

2.3.6 Calculating heavy atom positions

A reflection F(h,k,l) and its 'opposite', F(-h,-k,-l) will have the same magnitude and opposite phases. This is called Friedel's Law and the two reflections are called 'Friedel pairs' when not related by any space group symmetry operation. Symmetry-related reflections having the same intensities but opposite phases (180° apart) are called 'centric reflections'. The structure factors of centric reflections can be expressed as real numbers equal to their amplitudes (Fig.2.5a & b). In such cases, the centric reflection pairs show a much drastic change in the intensities in the presence of a heavy atom, while acentric reflections do not show large changes (Fig.2.5c).

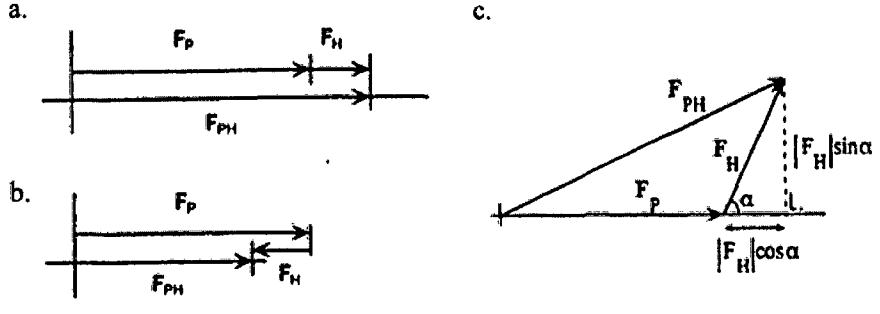


Figure 2.5. Structure factors in an isomorphous replacement method for centric reflections at **(a)** 0°; **(b)** 180° with phase reversed; **c.** for acentric reflections. (Adapted from Blow, 2002)

Therefore, a map plotted for the changes in the intensities of centric reflections can actually show the positions of heavy atoms. The variation of electron density in the unit cell or the 'Electron density' is the 'Fourier transformation' (a mathematical operation used for simplifying wave functions) of the structure factor amplitudes (h, k, l) and the phase α. 'Patterson function' is a Fourier transformation of the set of squared but not phased reflection amplitudes. This function does not produce an electron density map of the contents of the unit cell but rather a density map of the vectors between scattering objects in the cell. A 'Patterson map' is a representation of these vectors as squares of the numbers of electrons of the scattering atoms and thus, Patterson maps of crystals that contain heavy atoms are dominated by the vectors between heavy atoms. However, because the number of peaks in a Patterson map is also related to the square of the number of atoms, protein Patterson map is rarely interpretable. Nevertheless, an 'isomorphous difference Patterson' $(F_{PH} - F_P)^2$ is interpretable and helps calculate the heavy atom coordinates in reciprocal space. This information was obtained for all the derivatives of LipA individually, using the program 'Patterson' from a collection of freely available Computational Crystallography Programs suite v.4 (CCP4 suite, 1994).

The heavy atom can also scatter the X-rays anomalously, very similar to the selenium atom in MAD. In such a situation, the Friedel pairs (acentric reflections with same amplitudes and opposite phase) show different intensities. The phasing power of MIR can be increased by adding information from the anomalous scattering of the heavy atom to the centric reflections' isomorphous differences, making the technique 'MIRAS' (MIR with anomalous scattering). 'Anomalous difference Patterson maps' were also generated using the same program and the concurrence of the sites obtained from both the difference Patterson maps was confirmed for all heavy atoms.

MIRAS was used for phasing LipA structure. In order to assess the presence, position, refine the positions and level of occupancy of the heavy atoms in the derivative data sets, a program called SOLVE was used (Terwilliger & Berendzen, 1999). SOLVE solves Patterson functions for each heavy atom, calculates difference Fourier, looks at a native Fourier to see if there are distinct solvent and protein regions, and can add up partial MIR solutions with the anomalous scattering of heavy atoms to build up a complete solution (Terwilliger & Berendzen, 1999). This program was run in various combinations of the data sets and their resolution limits until the 'best possible output' was attained.

2.3.7 Improving phases

SOLVE output is the initial electron density map with a heavy 'background'. This is due to the high (30-70%) solvent content in protein crystals, which form channels through the crystal lattice. The program DM employs a 'solvent flattening method' of density improvement whereby electron density positivity criteria are imposed and the protein crystal solvent region is assumed flat with a constant value of electron density (Cowtan, 1994). DM also leads to 'phase extension' where the phases obtained for a lower resolution due to the presence anomalous/isomorphous differences for low resolution data are extended to the highest resolution of the native data.

2.3.8 Model building

An automated model-building software RESOLVE was used for tracing the density-modified electron density map (Terwilliger, 2003) to model atomic positions. However, due to incomplete phase information, the map quality was not very good and iterative rounds of chain building using the experimental SOLVE/DM map were performed. Structure visualization and a large part of manual model building were done using the software O (Jones et al., 1991). Electron density was visualized in the form of maps generated as difference density maps that show the spatial distribution of the weighted difference between the measured electron density of the experimental/observed and the calculated structure factor amplitudes ($2F_{obs}$-F_{calc} or F_{obs}-F_{calc}).

2.3.9 Structure refinement and assessment

The structure was refined using CNS (Crystallography and NMR System), using the 'Maximum likelihood' refinement method. This method works on the principles that the best model is most consistent with the observations, consistency being measured statistically by the probability that the observations would be made, given the current model. If the model is changed to make the observations more probable, the model gets better, the likelihood goes up and the probabilities have to include the effects of all sources of error, including errors in the model. However, as the model gets better, the errors get smaller and the probabilities get sharper, which also increases the likelihood. Experimentally, the CNS programs are designed to use structure factor amplitudes as raw data fitted into the generated model.

A rigid body constraint was applied to the entire protein molecule to position the molecule properly in the unit cell. Simulated annealing molecular dynamics was performed at high temperatures using iterative cycles of heating to 5000K followed by slow cooling cycles to refine individual atomic positions. Thermal disorder in the atomic positions, estimated in terms of 'B-factors', was also refined individually for each position. The water molecules associated with the protein also contribute to the electron density and were accounted for by using CNS. Iterative cycles of these refinements were performed until the error in the agreement between the model and observed diffraction data was within 10% range. This error is represented by the term R-factor ($\sum F_{obs}-F_{calc}/\sum F_{obs}$). The possibility of over-fitting the data to an extent of misfit is overcome by computing the R factor for a test set of 5% unrefined data (R_{free}). R and R_{free} were calculated for LipA structure and used as guiding factors for the refinement process. PROCHECK was used to assess the stereochemistry of the structure by calculating r.m.s. deviations in the bond lengths and bond angles and Ramachandran plot statistics. Coordinates and structure factors of the LipA structures from Type I and Type II crystals have been deposited at the PDB with codes 3H2J and 3H2G respectively.

2.3.10. Structure & sequence analysis

DALI server, freely available software in the public domain, was used for structural homology searches (Sali & Blundell, 1993). NCBI BLAST server was used for identifying sequence homologs (Altschul et al., 1997).

2.4 Results & Discussion

2.4.1. LipA crystals & X-ray diffraction data collection statistics

LipA was overexpressed in the secreted form by cultures of BXO2008 grown in peptone-sucrose medium. The protein was purified to homogeneity and crystallized at a concentration of 5mg ml^{-1} at 298 K (Fig.2.6a). Needles and clusters were obtained using Hampton Research Crystal Screens I and II, and the best conditions were used for expansion to obtain single mountable crystals of two types. Plate-shaped type I crystals (Fig.2.6b) were obtained using 48% PEG 400, 0.10 M MES pH 6.0. A crystal of this type diffracted to 1.89 Å resolution and 120° of diffraction data was collected. The unit cell parameters were found to be a = 93.1, b =

62.3, c = 66.1 Å, β = 90.8°. The crystal belonged to space group C2 and the mosaicity of the crystal was 0.98°. The R_{merge} of the data set was 5.6% at 90% completion (Table 2.1).

The rod-shaped type II crystals also belonged to space group C2, but had different unit cell parameters of a = 103.6, b = 54.6, c = 66.3 Å, β = 92.6°. The crystals were obtained from a reservoir solution containing 12% PEG 6000, 0.10 M MES pH 6.7 (Fig.2.6c). These crystals diffracted to 1.86 Å resolution (Fig.2.6d) with a mosaicity of 0.44° and 176° of diffraction data was collected. The data set was 99% complete, with an R_{merge} of 3.8% (Table 2.1).

	Type I	**Type II**
Space group	C2	C2
Unit-cell parameters (Å)	a = 93.1, b = 62.3, c = 66.1, β = 90.8	a = 103.6, b = 54.3, c = 66.3, β = 92.6
Unit-cell volume ($Å^3$)	384213.4	375639.2
Resolution (Å)	25.0–1.89 (1.96–1.89)*	25.0–1.86 (1.93–1.86)*
Observations	67204	108949
Unique reflections	27328 (2264)*	30943 (2979)*
Completeness (%)	90.1 (75.5)*	99.0 (96.2)*
Redundancy	2.5 (2.2)*	3.5 (3.2)*
R_{merge} (%)	5.6 (20.9)*	3.8 (11.6)*
I/σ(I)	18.4 (3.3)*	11.3 (3.7)*
V_M ($Å^3$ Da^{-1})	2.45	2.40
Solvent content (%)	49.7	48.5
Molecules per ASU	1	1

* Values in parentheses are for the highest resolution shell.

Table 2.1. Crystallographic statistics of the native data sets collected from the two types of LipA crystals.

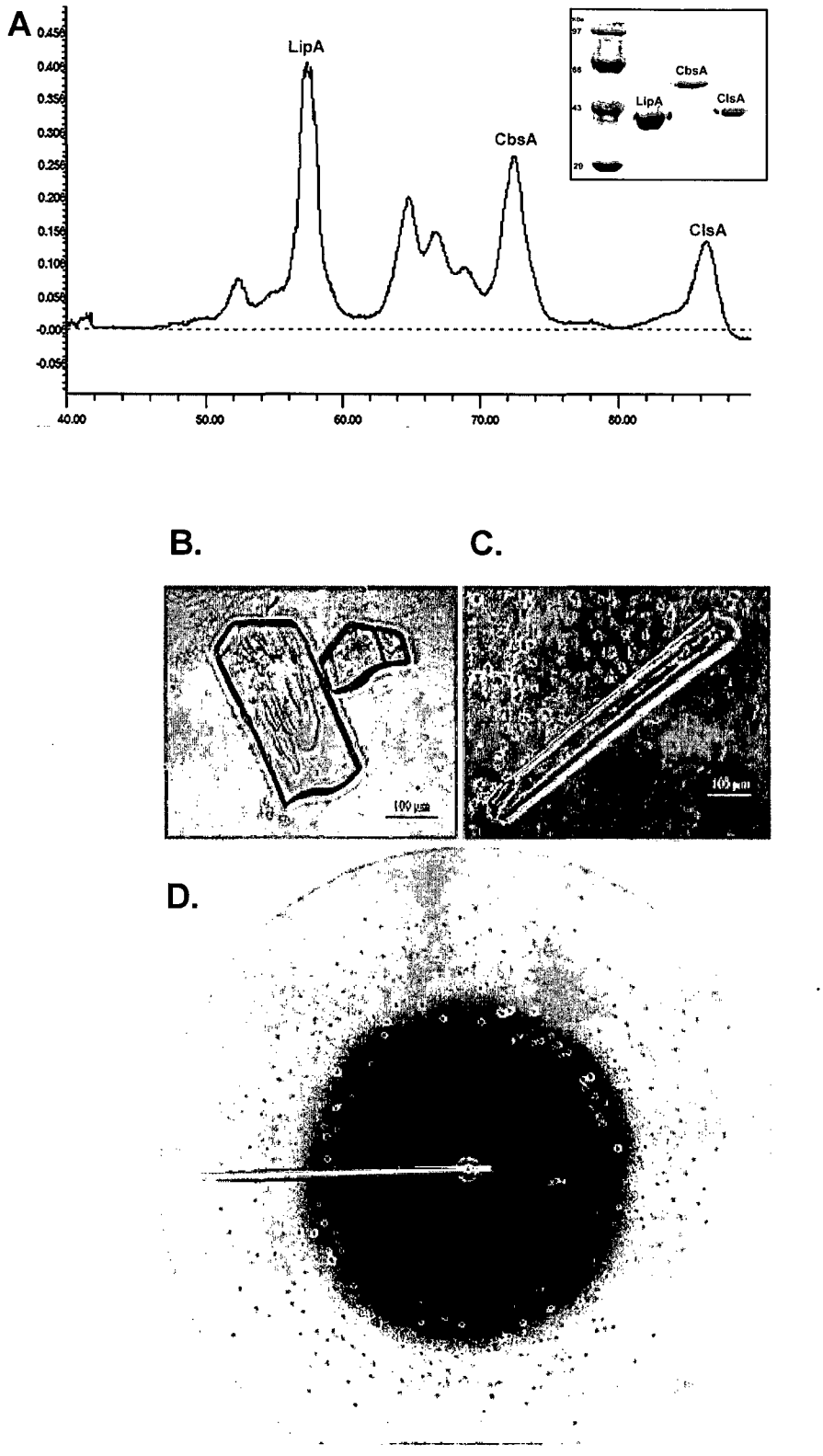


Figure 2.6. LipA purification and crystallization. **(A)** Ion-exchange chromatogram of the culture supernatant of BXO2008 LipA overexpressing stran of Xoo. ClsA: Xoo cellulase; CbsA: Xoo cellobiosidase; **Inset:** 10% SDS-PAGE showing the purity of the proteins obtained after HPLC. **(B)** Plate-shaped type I crystals. **(C)** Rod-shaped type II crystals. **(D)** A diffraction image from a type II LipA crystal. The edge corresponds to a resolution of 1.86 Å.

2.4.2. Heavy atom derivatives of LipA crystals

Molecular replacement (MR) method using the programs MOLREP (Vagin & Teplyakov, 1997) and Phaser (Read, 2001; Storoni et al., 2004) as a part of the CCP4 package (CCP4, 1994) with several structural homologues of LipA as search models did not yield any clear solution. This failure could be attributed to the low sequence identity of only around 20% and the fact that α/β-hydrolase fold proteins can have very different curvatures of the core β-sheet. Since the type II crystal form was more reproducible, this crystal form was used for a heavy atom search in order to solve the LipA structure using MIR. Quick soaking of LipA crystals in different heavy atom salts yielded derivatized LipA crystals. Each heavy atom salt chosen for soaking the LipA crystals was first standardized for soaking concentration and time. In addition, each salt of a particular heavy atom was treated independently since the binding to the protein molecule changes with the nature of the compound and not only the particular element. The soaking conditions and an assessment of the derivatization for the heavy atoms for which data was collected are listed in Table 2.2. The data collection statistics for each derivative data set are given in Table 2.3.

Heavy Atom	Salt	Soaking Concentration	Soaking Time	No. of crystals tested	χ^2 (Estimate of variation between native and derivative data)*
Platinum	K_2PtCl_6	2.5 mM	10 min	10	11.5
		5 mM	15 min	5	14.4
	K_2PtCl_4	2.5 mM	10 min	3	21.1
Samarium	$Sm(NO_3)_3$	5 mM	5 min	10	22.0
Mercury	Ethyl Mercuric Thio salicylic acid (Thimerosal)	10 mM	1 Hour	10	8.3

* χ^2 should have a range between 2 & 20 to signify an isomorphous substitution. χ^2 of ≤ 2 indicates no substitution and much larger than 20 indicates a non-isomorphous replacement.

Table 2.2. Soaking conditions used for obtaining heavy atom derivatives of LipA crystals and a preliminary check of substitution by estimating χ^2 that was calculated using a SCALEPACK program called 'scalenatder.com'.

	Platinum1	Platinum2	Platinum3	Samarium	Thimerosol
Soak condition (mM)	10	25	25	10	10
Soak time	10min	10min	10min	10min	30min
Data Collection Statistics					
Beamline	Rotating anode	Rotating anode	X11, Hamburg	Rotating anode	Rotating anode
Wavelength (Å)	1.5418	1.5418	0.8148	1.5418	1.5418
Unit-cell Parameters	a=103.3,b=54.7, c=65.8, β=92.6	a=103.5,b=55, c=65.8, β=92.4	a=102.5,b=55.2, c=65.4, β=92.7	a=103.4,b=54.5, c=66.3, β=92.8	a=103.4,b=54.3, c=66.2, β=93.2
Resolution (Å)	25-1.97 (2.0-1.97)*	25-2.6 (2.69-2.6)	30- 2.1 (2.14-2.1)	25- 2.0 (2.07-2.0)	25- 2.3 (2.38-2.3)
Unique reflections	26105 (2597)	11325 (1019)	22566 (1146)	24270 (2276)	15960 (1514)
Completeness (%)	99.9 (99.7)	98.3 (88.8)	99.6 (99.4)	91.9 (96.8)	96.9 (92.1)
R_{sym}	9.9 (32.0)	14.5 (39.7)	12.3(30.0)	8.8 (29.7)	16.1 (39.4)
Redundancy	5.5 (5.2)	6.2 (3.4)	6.2 (6.0)	3.1 (2.7)	3.7 (2.8)
I/(σ)I	15.1 (4.3)	10.4 (2.4)	14.3 (4.9)	10.4 (2.4)	6.9 (2.2)

* Values in parentheses are for the highest-resolution shell.

Table 2.3. Crystallographic statistics for the heavy atom derivative data sets used in the structure solution of LipA

2.4.3. LipA structure phasing and refinement

The first step towards obtaining phase information for LipA structure was to locate the heavy atom positions in each of the five derivative data sets (Table 2.4). The major inference that could be drawn from the SOLVE run was that the phasing power of the derivatives was weak and that the three Pt derivatives and one Hg derivative had a common site. Presence of even different heavy atoms at the same site does not add any phase information. This was reflected in the low SOLVE scores of 47.43 and a figure of merit (FOM) of 0.31. Scores above 50 and

FOM above 0.5 are generally considered significant. However, removing any data set reduced the scores even further. Several attempts to obtain newer derivatives with different sites failed. Nevertheless, the heavy atom positions were confirmed manually by checking the difference Patterson maps.

Heavy Atom	X	Y	Z	Peak Height (σ)	Occupancy (%)	Thermal B factor
Platinum						
Pt1	0.291	0.002	0.115	14.5	81	60.0
	0.267	0.450	0.489	9.1	31	35.9
	0.267s	0.039	0.091	10.5	30	22.1
Pt2	0.289	0.000	0.120	44.0	47	41.0
Pt3	0.288	0.000	0.120	40.1	39	24.3
Samarium						
Sm	0.676	0.233	0.019	20.3	27	25.2
Mercury						
Hg	0.282	0.009	0.115	6.5	8.4	8.8
	0.719	0.123	0.042	11.4	7.2	6.3

Table 2.4. The heavy atom locations in the five derivative data sets, as identified by SOLVE.

Major sites from the SOLVE output were treated as the real space coordinates (x, y, z; fractional distances relative to unit cell dimensions) and the peak positions from the Patterson maps as the reciprocal space coordinates (u, v, w). The coordinates of each peak corresponds to a vector difference between one pair of crystallographically related heavy atoms (*i.e.* atoms related by space group symmetry operators). The mathematical expression for the difference vector is given by taking the difference between each pair of symmetry operators in the space group, C-centered monoclinic in this case. The symmetry related equivalent positions were calculated to be x, y, z; x + ½, y + ½, z; -x, y, -z and ½ - x, ½ + y, -z. Difference vectors were calculated in the 16 possible ways, subtracting each vector with the rest three. Only vectors with y = 0 were compared with the peak positions (u, w) in sections v =0 (also called Harker section). In these cases, it was confirmed that the following condition is fulfilled: u, w = +/- 2x, +/- 2z. This exercise clearly showed that the peaks identified by the SOLVE run were

actually the heavy atom positions (Fig. 2.7). There were some minor peaks in the SOLVE output (such as the Hg peaks) that could not be detected in the Patterson maps, possibly due to the low occupancy levels and lesser peak heights.

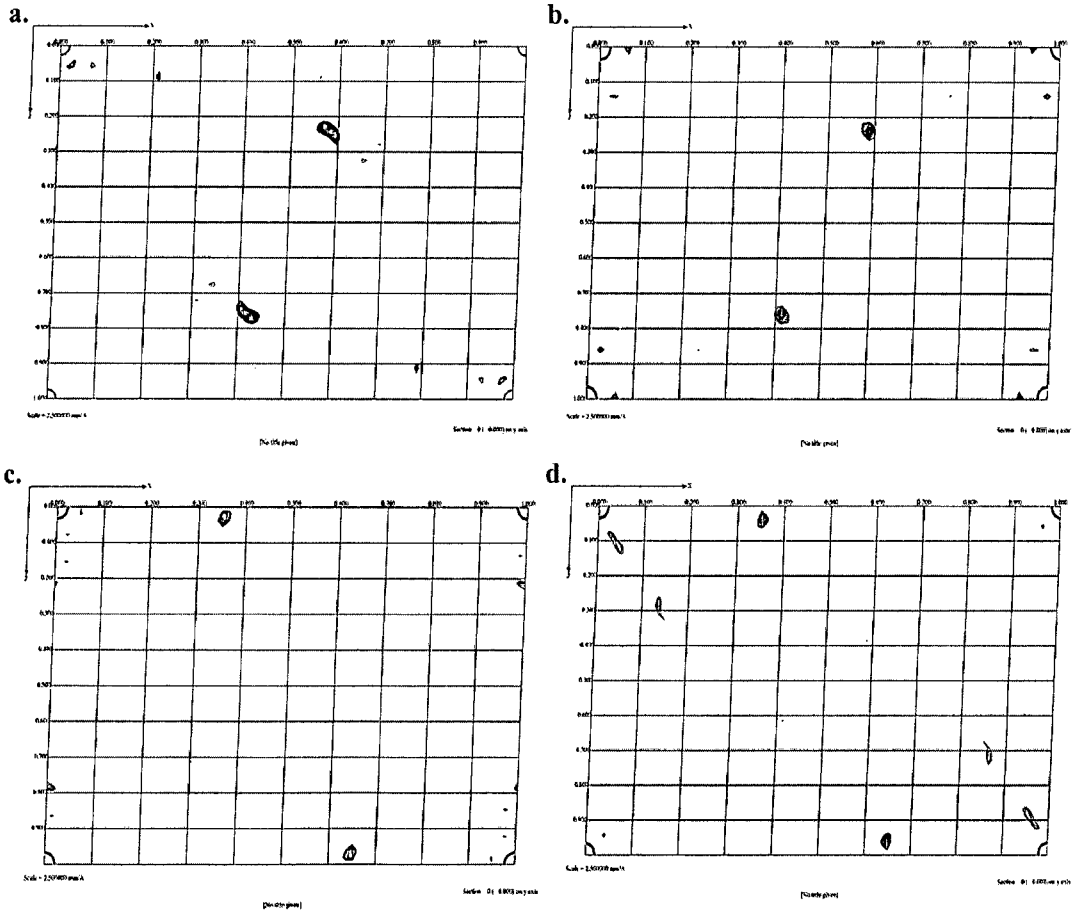


Figure 2.7. Patterson map of the platinum dataset. a. Isomorphous difference Patterson map for Pt1 derivative of Platinum at v = 0, showing a heavy atom peak (position u=0.58, v=0, w=0.23) corresponding to Peak1 in Table 4.4.3a (x=0.29, y=0, z=0.115). **b.** Anomalous difference Patterson map showing the coinciding peak at v = 0. **c.** Isomorphous difference Patterson map for Samarium derivative (u=0.35, v=0, w=0.15) at v = 0 **d.** Anomalous difference Patterson map showing the coinciding peak at v = 0.

The Fourier map (Fig. 2.8) calculated using the heavy atom positions from the SOLVE output was used as input for solvent flattening. The quality of the Fourier map after RESOLVE run was much higher with evident secondary structure elements (Fig.2.9). Automated model building by RESOLVE could build only 150 alanines and 30 side chain residues out of 397 expected amino acids. Iterative rounds of chain building using the experimental map for the rest of the 217 residues and RESOLVE runs for localised loop building were performed (Table 2.5). The final model containing 387 residues was refined to R/R_{free} values of 16.2/18.8% at 1.86Å resolution using the visualization software O (Table 2.6).

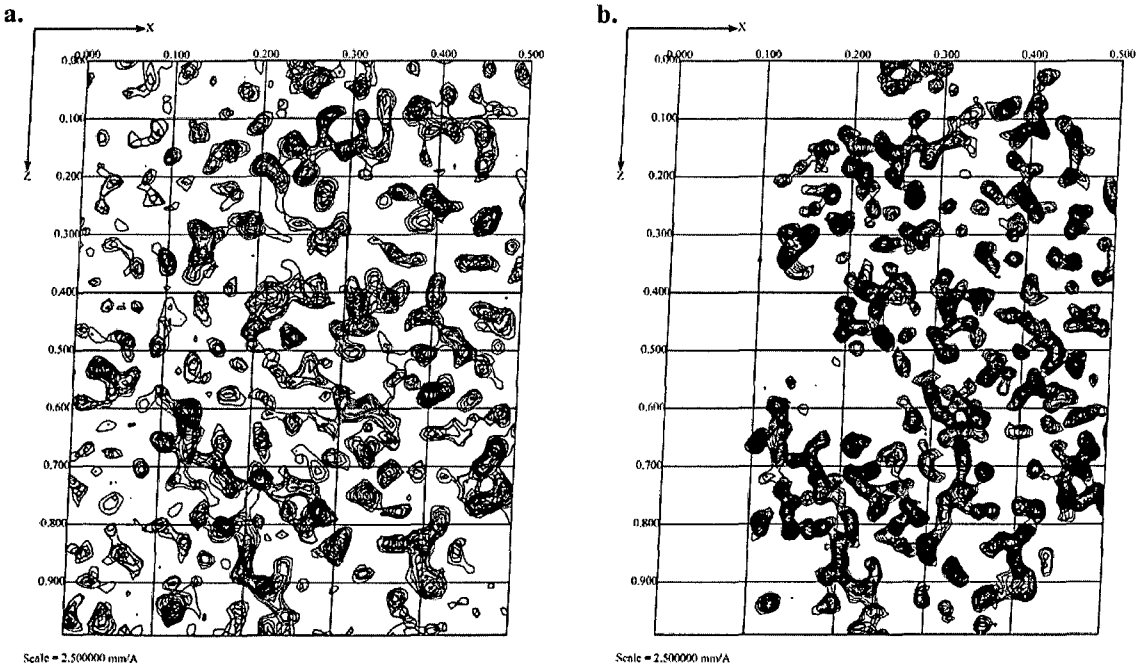


Figure 2.8. Comparison of fourier maps calculated after Solve and Resolve runs. a. Fourier m calculated using the heavy atom positions from the SOLVE output. **b.** Improvement in the phases evident in the RESOLVE Fourier map run after density modification by solvent flattening.

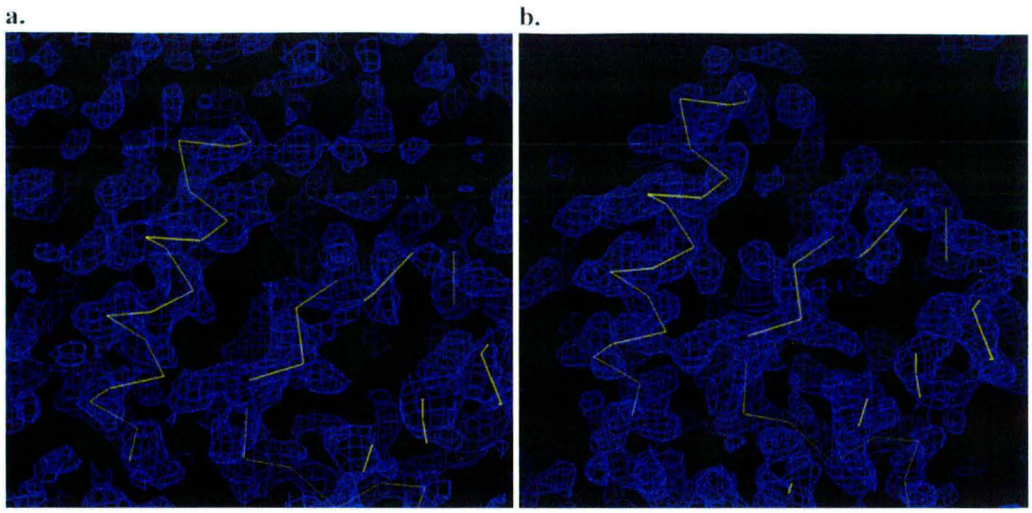


Figure 2.9 Comparison of Electron density maps calculated after Solve and Resolve runs. a. Electron density map calculated using the heavy atom positions from the SOLVE output. **b.** Improvement in phases after density modification by solvent flattening and RESOLVE.

Stage of Structure Solution	Resolution	No. Of Residues Built	R/R_{free}(%)[a]
Solve run Score = 35.91.; Figure of Merit = 0.32	2.5Å	70 aa; No side chains Map quality very poor	55.6/59.4
Solve run with density modification Score = 47.43; Figure of Merit = 0.31	2.5Å	156 aa; No side chains Map quality improved with several side chains and connectivities visible	48.8/53.1
Iterative Resolve runs with several rounds of manual model building	1.86Å	216 aa; 33 side chains	46.6/50.2
The whole model with alanine trace	1.86Å	394 aa; No side chains	40.0/43.6
Final refinement with waters added	1.86Å	394 aa; All side chains; All waters	16.2/18.8

[a] Throughout the refinement, 5% of the total reflections were kept aside for R_{free}.

Table 2.5. Progress of refinement during the various stages of structure solution of LipA.

Refinement Statistics	
R (%)	16.2
R_{free}(%)[a]	18.8
Mean B factor (Å^2)	
Protein	9.4
Water	24.9
Ligand	-
R.m.s deviation in	
Bond distances (Å)	0.004
Bond angles (°)	1.227
No. Of residues	387
No. Of atoms	3472
Protein	2941
Water	531
Heteroatoms	0
Ramachandran Plot Statistics	
Residues in most favoured regions	298 (90%)
Residues in additionally allowed regions	31 (9.4%)
Residues in generously allowed regions	0
Residues in disallowed regions	2 (0.6%)

[a] Throughout the refinement, 5% of the total reflections were kept aside for R_{free}.

Table 2.6. Final refinement statistics of the LipA structure.

2.4.4. General features of LipA structure

LipA is an α/β hydrolase-fold protein with a 9-stranded central mixed β sheet surrounded by α helices in a typical hydrolase topology (Nardini & Dijkstra, 1999) and an additional N-terminal short strand and helix (Fig.2.10). N-terminal sequencing indicated the cleavage of a 29 amino acid signal peptide and the corresponding electron density could not be detected. Electron density for the first residue after the signal peptide (residue 1) and the region from residues 28-36 was not detected in the LipA structure solved using either crystal forms.

The putative catalytic residues Ser 176, His 377 and Asp 336 (Fig.2.10; inset) are positioned similar to the canonical catalytic triads found in several other hydrolases. In general, the three residues of the classic catalytic triad each have a specific role in generating the nucleophilic potential at the seryl Oγ position and mediate a 'charge relay'. The aspartate residue forms a hydrogen bond with the adjacent imidazole group of the histidine. This interaction orients the imidazole group in the triad and increases the group's pK_a and basicity. The imidazole group (now a base) accepts the seryl Oγ proton, becomes positively charged and forms a transient salt bridge with the aspartate (Dodson & Wlodawer, 1998). The nucleophilic S176 of LipA lies on a 'strand-turn-helix elbow' forming a G-X-S-X-G motif that is conserved in hydrolases (Jaeger et al., 1999). It is interesting to note that the presence of the active site serine on such a bend forces it into the disallowed region of the Ramachandran plot, as was also observed in LipA. Tyr 125 was also found to be present in the disallowed region.

The distinctive feature of LipA structure is the presence of a 108 amino acid domain present as an insertion between the β6 and β7, corresponding topologically to the lid domain insertion position of triacylglycerol lipases (Fig.2.10; coloured green). This insertion domain consists of seven α helices with one 3_{10} helix (Fig.2.10b).

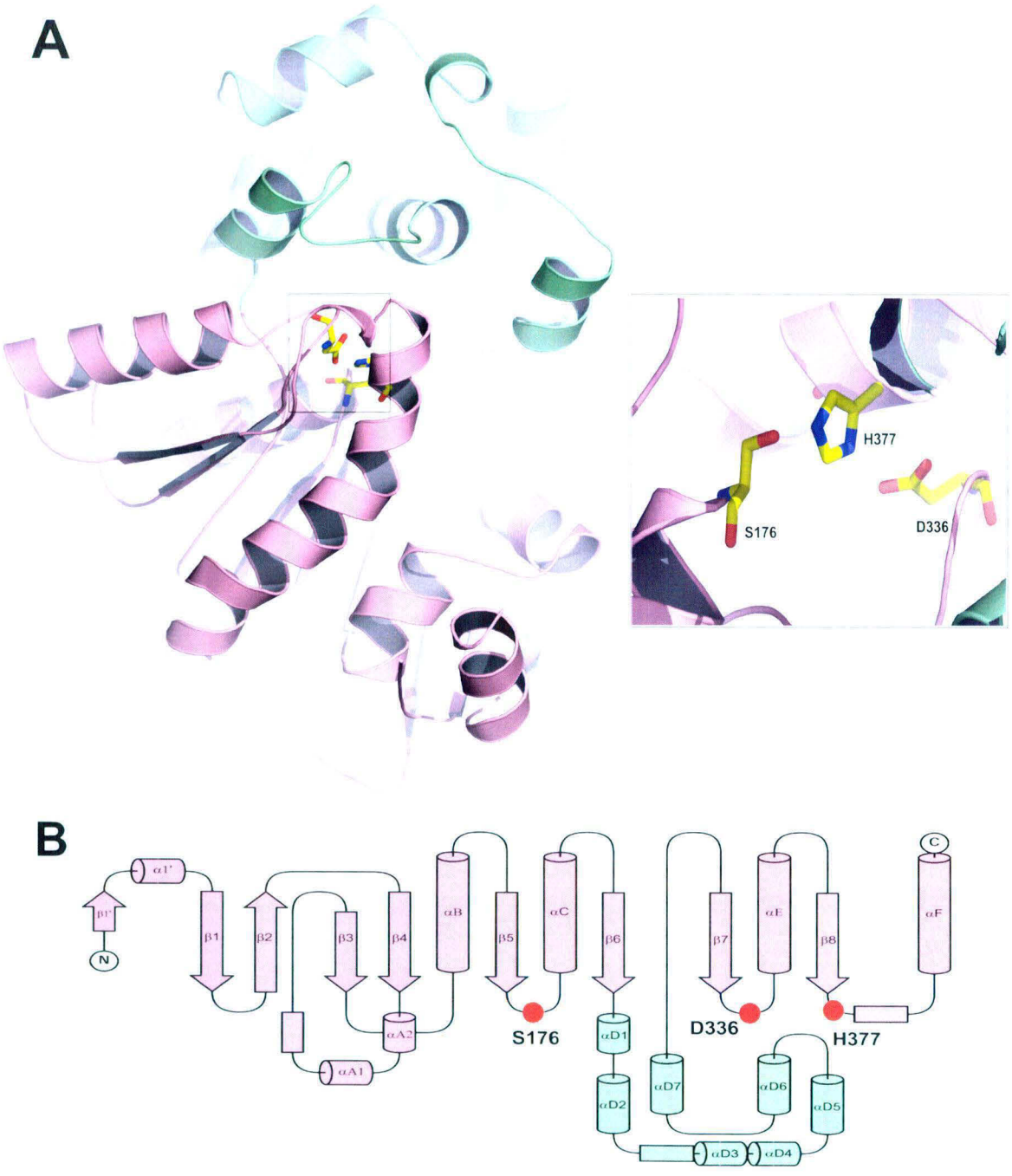


Figure 2.10. Three-Dimensional Structure of LipA (**A**) Cartoon representation of LipA structure showing the active site catalytic triad residues as yellow sticks (box). The hydrolase domain is depicted in pink and the lid-like domain in green. **Inset:** S176, D336 and H377 are given in a stick representation. (**B**) Topology diagram of LipA. The β sheets are shown as arrows, α helices as cylinders and 3_{10} helices as rectangles. The αD set of helices form the lid-like domain (green).

Figure 2.11. Structural superimposition of LipA with CalA. Cartoon representation of LipA structure (green) superimposed with its closest structural homolog, *Candida antarticus* lipase CalA (yellow).

Molecule	Z-Score	Aligned Residues	RMSD [Å]	% Seq. Identity	Known function (Organism)
2VEO	31.5	326	2.8	17	CalA Lipase
Lid (215-319)	2.5	92	3.2	9	(*Candida antarcticus*)
2QZP	22.2	248	2.8	17	Acyl-peptide hydrolase
Lid (473-504)	0.7	30	1.9	7	(*Aeropyrum pernix* K1)
1L7A	20.1	249	3.2	16	Cephalosporin C deacetylase
Lid (209-249)	0.5	35	3.6	3	(*Bacillus subtilis*)
1N1M	19.2	243	3.1	13	Dipeptidyl peptidase IV
Lid (657-687)	NA	NA	NA	NA	(Homo sapiens)
1JKM	18.8	251	3.3	16	Brefeldin A esterase
Lid (235-282)	NA	NA	NA	NA	(*Bacillus subtilis*)
1JFR	16.6	202	3.0	19	Lipase
NA	NA	NA	NA	NA	(*Streptomyces exfoliates*)
1VLQ	18.6	243	3.0	14	Acetyl xylan esterase
Lid (215-254)	0.3	33	3.7	3	(*Thermotoga maritima*)
1A88	15.6	212	3.2	14	Chloroperoxidase L
Lid (126-206)	0.7	48	4.3	4	(*Streptomyces lividans*)

Table 2.7. Structural homology searches of LipA using Dali server. The whole protein superposition parameters are listed. DaliLite pairwise alignment was used to calculate Z-scores and r.m.s. deviation between LipA ligand-binding domain and equivalent domains present at the same region in LipA structural homologs which showed up as top hits. Note that the LipA lid-like domain has no structural homologs as such.

Chapter Three

Insights into structure and function of the rice cell wall degrading esterase from *Xanthomonas oryzae* pv. oryzae

CHAPTER THREE

3.1 Abstract

The X-ray crystal structure of the secreted enzyme LipA from Xoo revealed an all-helical module as a distinct structural attachment to the canonical hydrolase catalytic domain. In order to understand the functional role of the accessory domain, an acyl glycoside ligand-bound LipA structure was determined. This structure demonstrated that the ligand binding is mediated through a rigid pocket comprising of distinct carbohydrate-specific and acyl chain-recognition sites where the ester linkage is situated 15Å from the anchored carbohydrate. LipA *in planta* function, exemplified by loss of both virulence and the ability to elicit host defense responses such as callose deposition and programmed cell death, was abrogated upon disruption of the carbohydrate anchor site or by blocking the pocket even at a considerable distance from the enzyme active site. This module is highly conserved across genus Xanthomonas, emphasizing the significance of this unique plant cell wall degrading function in this important group of plant pathogenic bacteria. A comparison with the related structural families illustrates how a typical lipase is modulated to act on plant cell walls and mediate the disease, thus providing a remarkable example of the emergence of novel functions around existing scaffolds for increased proficiency of pathogenesis during host-pathogen coevolution.

3.2 Introduction

The LipA structure has two distinct domains, a catalytic domain found commonly in hydrolases and a unique accessory domain with no characterized sequence or structural homologs. The catalytic domain belongs to the very well known family of α/β hydrolase fold proteins (Ollis et al., 1992). This fold is found in numerous hydrolytic enzymes of diverse substrate specificity and exhibits extensive sequence variability across several genomes (Holmquist, 2000). It is one of the most functionally flexible folds in SCOP database along with TIM-barrel, Rossmann fold, ferredoxin fold and P-loop NTP hydrolase fold (Hegyi & Gerstein, 1999) whilst it is classified into the highly promiscuous 3-layer (αβα) sandwich architecture in CATH database (Orengo et al., 1997). The α/β hydrolase family proteins possess a common structural fold, which consists of α/β sheet of five to nine parallel β strands, each connected by α helices. Loops situated at the C terminus of three β strands bear the active site serine, histidine and aspartate residues that form the catalytic triad, as observed in LipA structure as well.

Esterases and lipases form a large group of hydrolase fold proteins. Several insertions and deletions in the basic hydrolase scaffold are found in nature, each change evolving towards hydrolyzing esters found in the local habitats of the pertinent organisms, thereby altering the substrate-specificity (Nardini & Dijkstra, 1999). A small flap-like lid domain of lipases is a specific adaptation for long chain triacylglycerols and the movement of this domain in lipases is implicated in the phenomenon of 'interfacial activation' in response to highly hydrophobic substrates (Verger, 1997). Esterases act on less hydrophobic substrates and generally lack lid domains (Jaeger et al., 1999). When present, the lid remains 'permanently open' such as in fungal feruloyl esterases (Hermoso et al., 2004). The lid is absent in certain lipases, cutinases, acetylxylan esterases and cholesterol esterases (van Pouderoyen et al., 2001; Martinez et al., 1992; Ghosh et al., 1999). Association of the catalytic hydrolase domain with additional non-lid non-catalytic domains for specialized substrate binding has been seen in prolyl oligopeptidases (Szeltner & Polgar, 2008) and brefeldin A esterase (Wei et al., 1999) also.

The plant cell walls, where the Xoo LipA must act, comprise of several organic compounds such as cellulose, hemicellulose and lignin, heavily cross-linked by ester linkages. There is an abundance of heterogeneous polysaccharides such as arabinoxylan, the most extensive hemicellulose in cereals and grasses, consisting of a chain of β-1, 4 linked xylopyranosyl

residues. Many of the sugars residues are substituted with acetyl, arabinosyl and methyl glucuronyl residues, which further facilitate cross-links. The hydrolysis of this intricate plant cell wall structure requires an array of pathogen-secreted enzymes including different esterases (acetyl xylan esterases, feruloyl esterases, cinnamoyl esterases) along with main chain hydrolysing enzymes like cellulases, xylanases and sugar debranching enzymes, such as α-glucuronidase, α-L-arabinofuranosidase and β-xylosidase, working synergistically in all probability (Howard et al., 2003). Microbial acetyl xylan esterases hydrolyze the acetyl ester groups from the D-xylopyranosyl residues in xylan chains (Biely, 2003). Feruloyl esterases release ferulic acid esterified to α-L-arabinofuranosyl side chains in arabinoxylans (Salnier and Thibault, 1999). However, our current understanding of the chemistry of plant cell walls is far from complete and there are several uncharted linkages and moieties in this milieu (Somerville et al, 2004). Consequently, the functioning of microbial hydrolases involved in plant cell wall degradation also remains uncharacterized. Very little is known about the structure-function relationships of these enzymes leading to plant cell wall degradation and pathogenesis.

In an effort to understand the basis of LipA function, we cocrystallized LipA with β-octyl glucoside (BOG), a small molecule identified in ligand-screening experiments. Analysis of LipA interactions with BOG, led to the structure-based identification of a distinct ligand-binding module with a tunnel to recognize different components of BOG. We created several point mutations in this module and confirmed its role in ligand binding, virulence and elicitation of host defense responses, using a battery of *in vitro* and *in planta* assays, demonstrating the essentiality of this unique module in Xoo pathogenesis.

3.3 Materials & Methods

3.3.1. Cocrystallization of LipA with BOG

LipA was set up for a hanging drop vapour diffusion experiment containing 12% PEG 6000, 0.10 M MES pH 6.7 in the reservoir, the drops containing a 1:1 mixture of the protein and reservoir solution and each drop overlayed with a detergent or other hydrophobic ligands from Hampton additive screen (Hampton Research, USA). Hydrophobic and detergent ligands were chosen in lieu of the enzyme being a lipase/esterase. Of the various ligands screened,

LipA cocrystals were obtained in the presence of BOG alone and these crystals were obtained using a concentration of 17.5mM of ligand.

3.3.2. Structure solution of LipA-BOG cocrystals

The LipA-BOG cocrystal X-ray diffraction data was collected in a manner similar to the type II crystals of wild-type LipA. LipA-BOG structure was solved using Molecular Replacement (MR) method of identifying the protein phase. This method is dependent upon the existence of a previously solved protein structure and the softwares available for performing MR try to find the model that fits the experimental intensities best among the known structures. A Patterson map (an inter-atomic vector map; Section 2.3.6) created by squaring the structure factor amplitudes and setting all phases to zero (since the phase information is lost during X-ray diffraction) contains a peak for each atom related to every other atom, with a large peak at the origin, where self-vectors (vectors relating atoms to themselves) cluster. Such a map is far too noisy to derive any high-resolution structural information. In MR, the crux of the method is that the Patterson maps for the data derived from the unknown structure and from the structure of a previously solved homolog in the correct orientation and position within the unit cell are closely correlated. An MR search is typically divided into two steps: rotation and translation. The first step is to rotate and orient the Patterson map of the known structure on the structure factors from the target. As in other cases, here too the R-factors are used to assess the rotation function. The highest correlation (and therefore, minimum R factor) is obtained when the two structures (known and unknown) are in similar orientation(s). In the translation function, the now correctly oriented known model can be correctly positioned by translating it to the correct co-ordinates within the asymmetric unit. This is accomplished by moving the model, calculating a new Patterson map, and comparing it to the unknown-derived Patterson map. The derived phases (can be biased towards the model) are enough to derive initial electron density maps, which can be improved by density modification and refined. Molecular replacement was performed using MOLREP (CCP4 suite, 1994) on the dataset of the BOG-bound crystal.

3.3.3. Structure & sequence analysis

The LipA-BOG interaction was studied using the structure visualization software O. Important residues involved in BOG interactions were identified and marked for site-directed

mutagenesis. Closest sequence homologs of LipA were identified using NCBI BLAST. These sequences were used for generating homology-based structural models using Modeller9v4 software (Marti-Renom et al., 2000). These models were analysed for BOG-binding capabilities by examining residues identified as important for formation of the pocket and their evolution. ClustalW (EBI server) was used for multiple sequence alignments and MEGA v.4 was used for preparing phylogenetic trees and bootstrap analysis (Tamura et al., 2007). MOLE was used as a PYMOL plugin to visualize the tunnel (Petrek et al., 2007). Coordinates and structure factors of the BOG-bound LipA, G231F and N228W mutant are deposited at the PDB with codes 3H2K, 3H2H and 3H2I, respectively.

3.3.4. Isothermal Titration Calorimetry

The *in vitro* binding of BOG to LipA and the LipA point mutants was measured by isothermal titration calorimetry (ITC) using a VP-ITC calorimeter (MicroCal, USA). Buffer containing 50mM Tris pH 7.5 and 150mM NaCl was used for diluting the protein and dissolving BOG. All samples were degassed prior to titration. 1.8ml of 40μM protein sample was titrated against 150μl of 2mM BOG over 50 injections at 303K. The change in heat of the proteins upon titrating with BOG was measured, integrated and fitted into a one-site binding model for calculation of K_a, ΔH_{app} and $T\Delta S$ using the Origin 7.0 software (MicroCal, USA) for curve fitting and data analysis. The parameters K_d and ΔG were calculated using formulae $K_d = K_a^{-1}$ and $\Delta G = \Delta H - T\Delta S$, respectively. The heat of dilution of BOG was measured by a blank titration of ligand into the buffer and this was subtracted from the binding isotherms of the wild-type and mutant proteins. The β-octyl galactoside titration was also performed similarly and a sequential two-site model was used to fit the raw data in this case.

3.3.5. Site-directed mutagenesis and purification of mutant proteins

The *lipA* gene was cloned into pBSKS plasmid (Stratagene, USA) and mutated using the QuickChange site-directed mutagenesis kit (Stratagene, USA; primers listed in Table 3.1). The mutant *lipA* genes were excised as *Kpn*I-*Hind*III fragments, cloned into the multi-cloning site of broad-host range vector pHM1 and transformed into Xoo strain BXO2001 that has an insertion mutation in the *lipA* gene (Rajeshwari et al., 2005). Presence of the *lipA* point mutations was confirmed by sequencing of the LipA gene from each strain (listed as BXO2311-17 in Table 3.2). Expression of mutant LipA proteins was confirmed using rabbit

polyclonal anti-LipA antibodies. As described previously, Xoo strain BXO2008 containing the *lipA* gene on the broad-host range vector pHM1 was used as a source of wild-type LipA. The LipA mutant proteins were purified to homogeneity using similar protocols. Mutant proteins G231F and N228W were crystallized in the same crystallization condition as LipA wild-type crystals. The mutant structures were solved using Molecular Replacement method as described in section 3.3.2.

Name	Primer Sequence (5' - 3')
S176A FP	5' GCAAGGTCATGCTGTCGGGTTATGCGCAGGGCGGCCACAC 3'
S176A RP	5' GTGTGGCCGCCCTGCGCATAACCCGACAGCATGACCTTGC 3'
G231A FP	5'GCGGTCGGCGAAAACACGTTCGCTATCCTTCTGGGAAGCTATGCC 3'
G231A RP	5'GGCATAGCTTCCCAGAAGGATAGCGAACGTGTTTTCGCCGACCGC 3'
G231I FP	5'GCGGTCGGCGAAAACACGTTCATTATCCTTCTGGGAAGCTATGCC 3'
G231I RP	5'GGCATAGCTTCCCAGAAGGATAATGAACGTGTTTTCGCCGACCGC 3'
G231F FP	5'GCGGTCGGCGAAAACACGTTCTTTATCCTTCTGGGAAGCTATGCC 3'
G231F RP	5'GGCATAGCTTCCCAGAAGGATAAAGAACGTGTTTTCGCCGACCGC 3'
N228W FP	5'GGCAGCAATGCGGTCGGCGAATGGACGTTCGGTATCCTTCTGGGAAG 3'
N228W RP	5'CTTCCCAGAAGGATACCGAACGTCCATTCGCCGACCGCATTGCTGCC 3'
G221I FP	5' GCAAACATTCCTGGATAGCTGGAGCATCAGCAATGCGGTCGGCGA 3'
G221I RP	5' TCGCCGACCGCATTGCTGATGCTCCAGCTATCCAGGAATGTTTGC 3'
W219A FP	5'GCAAACATTCCTGGATAGCGCGAGCGGCAGCAATGCGGTCGGCG 3'
W219A RP	5'CGCCGACCGCATTGCTGCCGCTCGCGCTATCCAGGAATGTTTGC 3'

Table 3.1 List of primers used for creating point mutations in Xoo *lipA*

3.3.6 Western analysis of LipA using anti-LipA antibodies

The full-length wild-type LipA protein was used to generate anti-LipA polyclonal antibodies in rabbit. Preimmune serum was collected from the uninjected animal to serve as a negative control for the antibodies following subdermal injection of a 1:1 mixture of 0.5mg ml^{-1} of the protein and Freund's incomplete adjuvant (Sigma, USA) and a booster dose of the same composition 10 days later. The serum containing the polyclonal antibodies was collected 15 days after the booster dose and centrifuged at 15,000rpm for 30min to remove blood cells. Expression of mutant LipA proteins was confirmed using anti-LipA serum at 1:5000 dilution and 1:10000 dilution of anti-rabbit alkaline-phosphatase conjugated goat IgG from Sigma, USA.

Strain/Plasmid	Characteristic Features	Reference/Source
BXO4	*rif*-2; Rf^r derivative of an Indian isolate of Xoo	Laboratory collection
BXO2001	*lipA1*:: *bla rif*-2; Lip$^-$, Ap^r derivative of BXO43	Rajeshwari et al.,2005
BXO2008	BXO2001/pGJ5; Sp^r Lip$^+$ derivative of BXO2001	Aparna et al., 2007
pGJ5	pHM1 + 1501bp PCR amplified fragment containing *lipA* gene of BXO43	Aparna et al., 2007
pHM1	Sp^r	Innes et al., 1988
DH5α	F, *end* A1 *hsdR*17 (rk$^-$ mk$^+$) *supE*44 *thi*-1 *recA*1 *gyr*A *rel*A1 f80d*lac*ZDM15D (*lac*ZYA-*arg*F) U169	Laboratory collection
pS176A	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with S176A mutation	This work
pG231A	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with G231A mutation	This work
pG231I	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with G231I mutation	This work
pG231F	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with G231F mutation	This work
pN228W	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with N228W mutation	This work
pG221I	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with G221I mutation	This work
pW219A	pHM1 + 1501bp PCR amplified fragment containing the *lipA* gene with W219A mutation	This work
BXO2311	BXO2001/pS176A; Sp^r derivative of BXO2001	This work
BXO2312	BXO2001/pG231A; Sp^r derivative of BXO2001	This work
BXO2313	BXO2001/pG231I; Sp^r derivative of BXO2001	This work
BXO2314	BXO2001/pG231F; Sp^r derivative of BXO2001	This work
BXO2315	BXO2001/pN228W; Sp^r derivative of BXO2001	This work
BXO2316	BXO2001/pG221I; Sp^r derivative of BXO2001	This work
BXO2317	BXO2001/pW219A; Sp^r derivative of BXO2001	This work

*Rf^r, Ap^r, Sp^{r}and Km^r indicate resistance to rifampicin, ampicillin, spectinomycin and kanamycin resistance respectively

Table 3.2. List of strains & plasmids used in the structure-function analysis of Xoo esterase LipA

3.3.7. Substrate clearance assay

LipA and its point mutants were assessed for the hydrolysis of triacylglycerol derivatives that are usually used as lipase substrates. Due to their hydrophobicity, these compounds cause the lipase lid domain to 'open' and only after this interfacial activation does the lipase hydrolyze the ester bond. However, owing to small size, tributyrin can bind to other hydrolases that have no lid or immobile substrate-binding domains. Therefore, tributyrin can be treated as a non-specific α/β hydrolase substrate. Tributyrin (C4), tricaproin (C6), tricaprylin (C8), tricaprin (C10), trilaurin (C12) and tripalmitin (C16) (Sigma, USA) were used as substrates in plate assays for LipA activity (Smetlzer et al., 1992). Briefly, 0.5% suspensions of the triglyceride substrates were prepared in a buffer containing 100mM Tris Cl pH 8.0, 25mM $CaCl_2$, sonicated at 30W for 3 min to emulsify the substrates, mixed with equal volume of 2% agarose solution and solidified in Petri plates. 50µl of 0.5mg ml^{-1} of purified wild-type and mutant LipA proteins were added to the wells cut into each substrate plate and assayed for a zone of clearance. Triolein (C18:1) activity was assayed using the olive oil-rhodamine B plate assay (Kouker & Jaeger, 1987). In this assay, 2.5% ultra pure olive oil and 0.001% fluorescent Rhodamine B dye were mixed with 1% agarose solution, emulsified using sonication at 30W for 3 min and solidified in a Petri plate. Subsequently, 50µl of 0.5mg ml^{-1} of purified wild-type LipA was added to the wells cut into the plate and assayed for appearance of a zone of clearance. All assays were performed at room temperature.

3.3.8. Virulence assay

BXO43, BXO2001, BXO2008 and *lipA* point mutations expressed using pHM1 in BXO2001 background (Table 3.2) strains of Xoo were used for virulence analysis of rice. These strains were grown to saturation in peptone-sucrose medium supplemented with appropriate antibiotics. The cultures were pelleted down by centrifugation and resuspended in sterile water (3ml) at a concentration of ~10^9 cells/ml. Surgical scissors dipped in these bacterial suspensions were used to clip the leaf tips of greenhouse grown, 40 days old plants of the Taichung Native-1 (TN-1) rice variety which is susceptible to Xoo (Kauffman, 1973). Lesion lengths were measured 7 days after inoculation.

3.3.9. Callose deposition assay

LipA proteins and its point mutant variants were used to assess the host-innate immune responses triggered by the enzyme action in rice. Callose deposition and programmed cell death assay were used as markers for the rice innate immune responses. 0.1mg ml^{-1} of purified wild-type and mutant LipA proteins were infiltrated into leaves of 10-day old TN-1 rice seedlings using the blunt end of a 1ml syringe. Fourteen hr later, the infiltrated zone (~1cm x1cm) is cut from the leaf, destained with 70% ethanol at 65°C, stained with 0.5% aniline blue for 4h, and observed under Axioplan2 epifluorescence microscope, using a blue filter (excitation wavelength 365nm) and 10X objective (Hauck, 2003).

3.3.10. Programmed cell death assay

Seeds of TN-1 rice cultivar were germinated on 0.5% sterile agar in Petri dishes. After 2-3 days, root tips (~0.5cm) were excised from the seedlings and treated with 0.5 mg ml^{-1} of either the wild-type and mutant LipA proteins or buffer (10mM phosphate buffer pH 6.0). After incubation for 16h, roots were washed and stained with 1mg ml^{-1} propidium iodide for 20min and mounted in 50% glycerol on glass slides. The samples were observed under LSM-510 Meta Confocal microscope (Carl Zeiss, Germany) using 63X oil immersion objectives and He-Ne laser at 514nm excitation as desçribed (Jha et al., 2007).

3.4 Results

3.4.1. Identification of a distinct ligand-binding domain in LipA structure

A screen was conducted for identifying putative ligands of LipA using detergents and fatty acid additives that could mimic its natural substrate(s) *in planta* using cocrystallization. LipA cocrystallized with the glycoside detergent β-octyl glucoside (BOG) in the same crystallization condition as that of the wild-type. The cocrystal structure solution at 2.1Å resolution, with an R/R_{free} of 18.8/23.3% respectively, showed two molecules of bound BOG (Fig.3.1a). One molecule (referred to as BOG1) has the thermal B-factors in the range of 18-23 $Å^2$ while the other molecule (BOG2) is loosely bound and has higher B-factors (45-50 $Å^2$). $2F_{obs}$-F_{calc} map contoured around the ligands is shown in Fig.3.1b. The ligand acyl chains, placed very close to each other (6.9Å), disclose a 30Å 'tunnel' passing by the active site residues and ending very close to the outer surface of the protein (Fig.3.2a). BOG2 glucose moiety hangs out of the tunnel facing the solvent.

The proximity of BOG1 terminal methyl group with Ser 176 active site residue (3.8Å) strengthens the idea that this tunnel could be involved in substrate binding. The putative entry side opening is very broad (20Å) and a narrow exit point between Ile 287, Val 290 and Ser 220 is also evident. R.m.s. deviation upon Cα superposition of the wild-type and ligand-bound LipA structures is only 0.29 Å, indicating almost no structural changes upon ligand binding (Fig.3.1b). The side chains also superpose well between the two structures indicating a relatively rigid pocket employed for binding BOG. The lid-like domain may also not exhibit any domain motion with respect to the hydrolase-fold upon ligand binding, unlike large movements seen in conventional lid domains of lipases during interfacial activation (Nardini & Dijkstra, 1999).

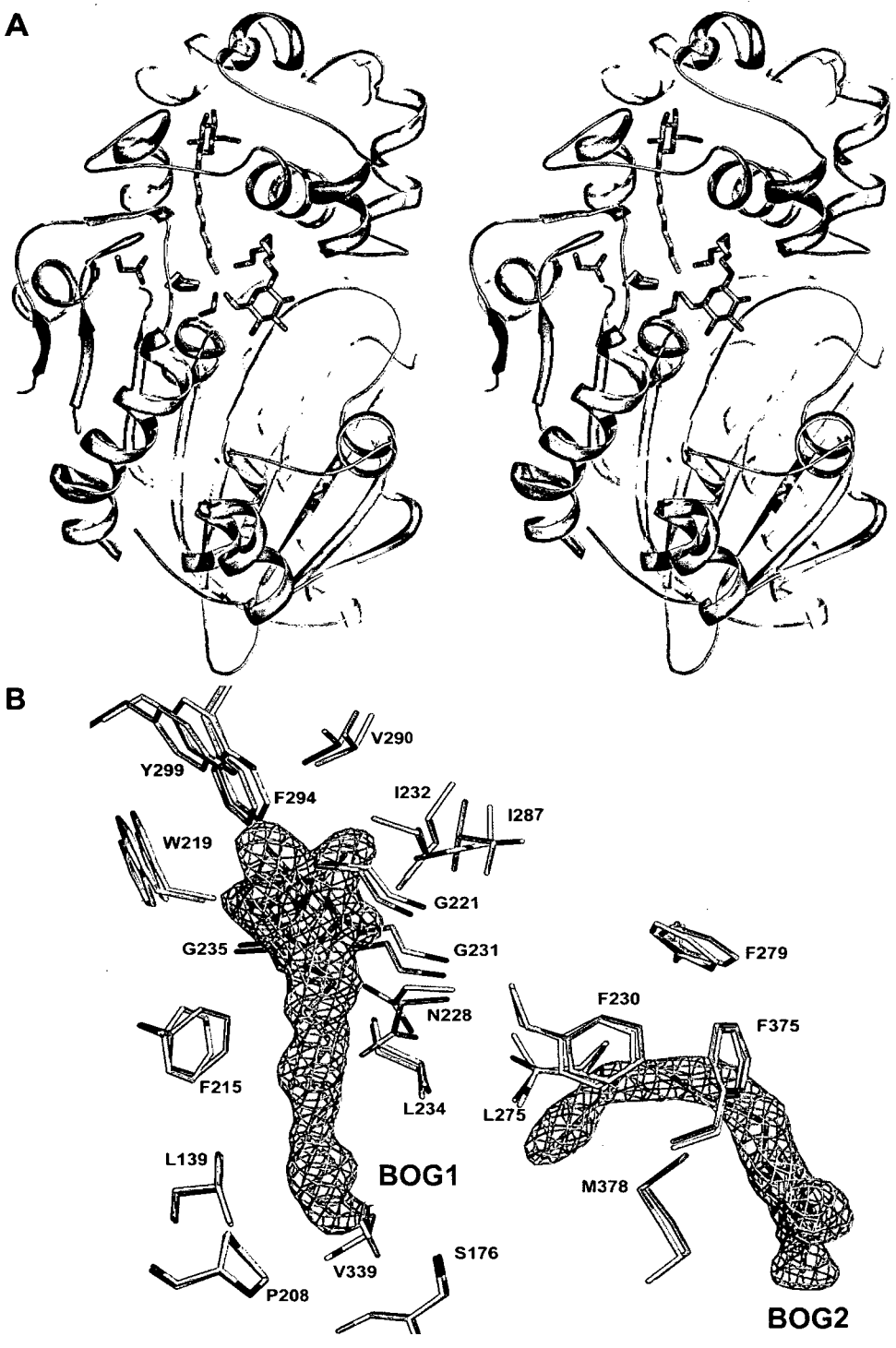


Figure 3.1. Three-Dimensional Structure of LipA. (A) Stereo-view of the ligand-bound structure of LipA showing two bound molecules of β-octyl glucoside and the catalytic triad. The hydrolase domain is depicted in pink and the ligand-binding domain in green. The ligand molecules (yellow), S176, D336 and H377 are given in a stick representation. **(B)** Residues lining the BOG-bound tunnel are shown. Electron density of the two BOG molecules is a $2F_{obs}$-F_{calc} map contoured at 1σ value. The amino acids are shown as yellow sticks. Active site S176 is highlighted. The wild-type LipA residues (grey) of the same region are superimposed.

An interesting feature common to both wild-type and BOG-bound structures is the absence of electron density for the region from residues 28-36 (Fig.3.2a black dots) in the vicinity of BOG2 site indicating that these residues are very flexible or unstructured, possibly due to absence of the natural substrate. This sequence is conspicuously absent in other closely related Xanthomonas strains (Fig.3.3a). Considering the position of this missing fragment, which marks the base of the entry region of the tunnel, this region might be involved in holding a long-chain substrate and positioning its entry into the tunnel.

The glucoside moiety of BOG1 interacts with the main chain atoms of LipA at the extreme end of the tunnel, ~15Å away from the active site serine, in a pocket made of three glycines and a few other polar residues (Fig.3.2b). In order to ascertain whether the pocket would confer carbohydrate specificity, we looked at the possibility of several other sugars besides β-D-glucose occupying the pocket and found that except β-D-xylose, all other sugars that were examined face clashes from the residues surrounding the pocket (Table 3.3). The BOG1 β-D-glucose O2 is held, albeit loosely, by main chain oxygens of Asn 228 and Ser 218, O4 by O: Trp 219 and O6 by O: Gly 231 (Fig.3.3b). Important intra-molecular interactions keep Gly 231 and Trp 219 in position even in the wild-type structure, indicating that the pocket is held in place and that substrate-induced fit is not occurring. Epimers of β-D-glucose, i.e., β-D-mannose and galactose have axially located O2 and O4 respectively and face a severe hydrophobic clash when superposed on the BOG1 sugar ring with the same pucker. Similarly, steric hindrances towards the axial O1 exclude the possibility of binding to α-D- sugars (Table 3.3).

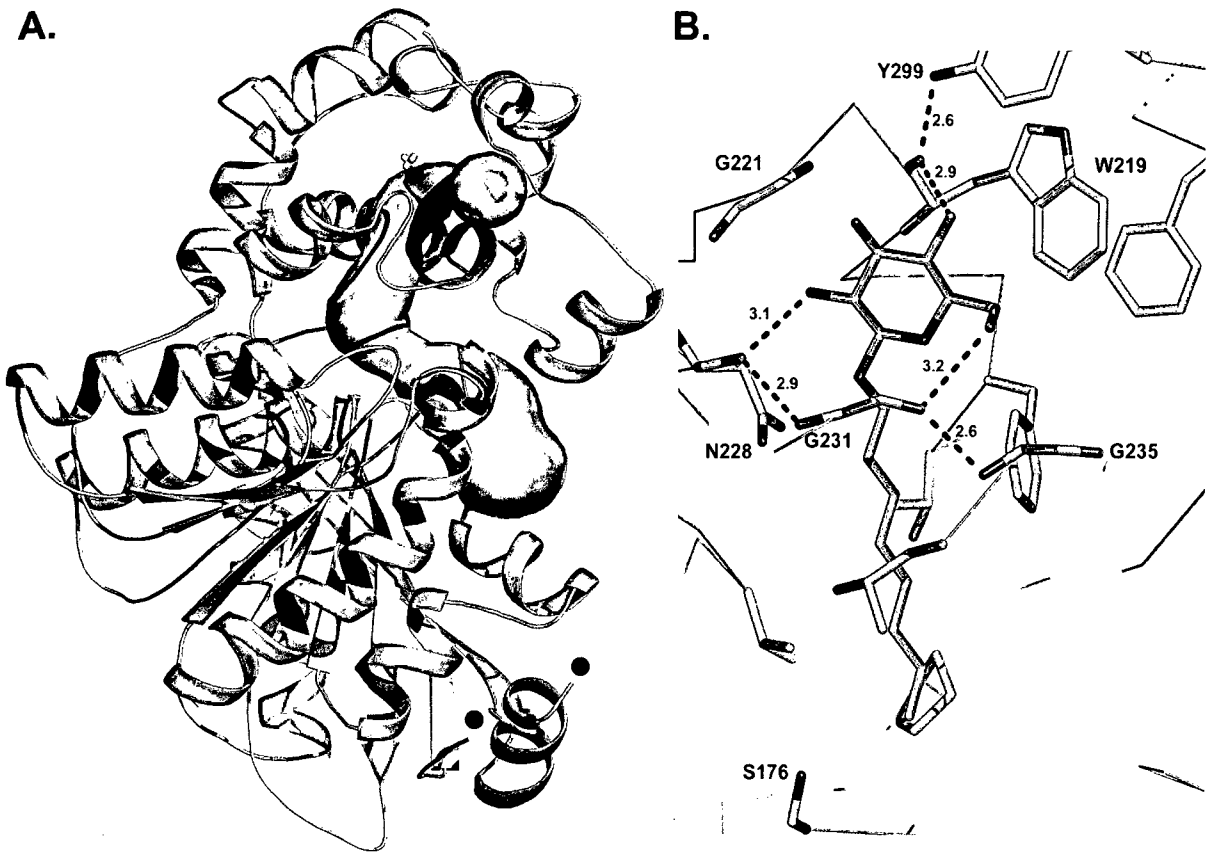


Figure 3.2. The LipA ligand-binding tunnel has a carbohydrate-anchoring pocket. (**A**) Cartoon representation of the ligand-binding tunnel as identified by the software MOLE. The two black filled circles indicate the stretch of residues 28-36 for which the electron density is missing. (**B**) Important BOG1-specific interactions in the carbohydrate-anchoring pocket of LipA are main-chain mediated and shown as dashed lines with distance in Å.

Type of sugar	Sugar hydroxyl :: Atom: residue of LipA (distances in Å)	Predicted binding
α-D- Sugars (O1 in axial position)	**O1 :: O:S218 (2.2)** **O1 :: C:S218 (2.8)** **O1 :: Cβ:S218 (3.1)**	NO (α-D- Sugars cannot be accommodated in the pocket)
β-D-Glucopyranose	O2 :: O:S218 (3.1) O2 :: O:S228 (3.1) O4 :: O:S219 (2.9) O6 :: O:G231 (3.2) O1 :: Cα:G231 (3.0)	YES
β-D-Xylopyranose (C6-O6 absent)	O2 :: O:S218 (3.1) O2 :: O:S228 (3.1) O4 :: O:S219 (2.9) O1 :: Cα:G231 (3.0)	YES
β-D-Galactopyranose (O4 in axial position)	O2 :: O:S218 (3.1) O2 :: O:S228 (3.1) **O4 :: Cδ1:I232 (2.8)** O6 :: O:G231 (3.2) O1 :: Cα:G231 (3.0)	NO
β-D-Mannopyranose (O2 in axial position)	**O2 :: Ca:G231 (2.7)** **O2 :: C:G231 (2.9)** O4 :: O:S219 (2.9) O6 :: O:G231 (3.2) O1 :: Cα:G231 (3.0)	NO
β-L-Arabinopyranose (O4 in axial position; O1 in axial position)	O2 :: O:S218 (3.1) O2 :: O:S228 (3.1) **O4 :: Cδ1:I232 (2.8)** **O1 :: O:S218 (2.2)** **O1 :: C:S218 (2.8)** **O1 :: Cβ:S218 (3.1)**	NO
Rhamnopyranose (O2 in axial position; -CH3 at C5 position) derivatives	**O2 :: Ca:G231 (2.7)** **O2 :: C:G231 (2.9)** O4 :: O:S219 (2.9) O1 :: Cα:G231 (3.0)	NO
Fucopyranose (O4 in axial position; -CH3 at C5 position) derivatives	O2 :: O:S218 (3.1) O2 :: O:S228 (3.1) **O4 :: Cδ1:I232 (2.8)** O1 :: Cα:G231 (3.0)	NO

Table 3.3. Interaction of various sugars found commonly in rice cell walls with LipA residues. Structural coordinates of different sugars (from PDB) were superposed with the glucosyl moiety of the BOG bound in LipA. Conformers with the same pucker as β-D-glucopyranose were used for analysis. Interactions shown in bold indicate predicted steric clashes among the query sugar hydroxyl(s) and LipA.

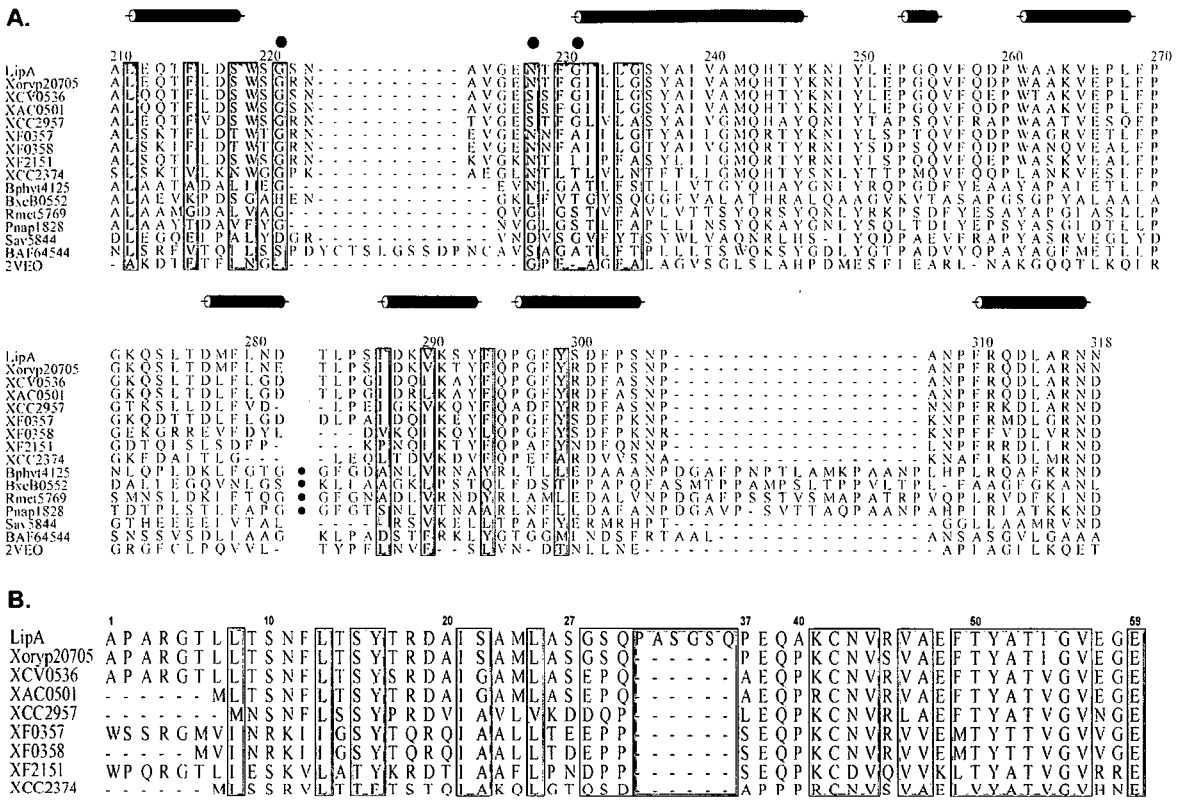


Figure 3.3. Sequence features of LipA (A) Multiple sequence alignment of LipA ligand-binding domain with homologous regions in other bacteria. *Xanthomonas oryzae* pv. oryzicola (Xoryp20705); *X. campestris* pv. vesicatoria (XCV0536); *X. axonopodis* pv. citrii (XAC0501); *X. campestris* pv. campestris str. ATCC 33913 (XCC2957 & XCC2374); *Xylella fastidiosa* 9a5c (XF0357, XF0358 & XF2151); *Burkholderia phytofirmans* (Bphyt4125); *B. xenovorans* (BxeB0552); *Ralstonia metallidurans* (Rmet5769); *Polaromonas napthalenivorans* (Pnap1828); *Streptomyces avermitilis* (SAV5844); *Ideonella* sp. (BAF64544) and *Candida antarticus* (2VEO). Residues marked in green are amino acids lining the tunnel and the red dots indicate the residues in the carbohydrate-anchoring pocket that were mutagenized. The black dots represent 39-45 aa inserts in the sequences of the corresponding proteins that are not shown in the figure. **(B)** An interesting region missing in the electron density of LipA structure. The sequence alignment of LipA N-terminal shows the region 28-36 (box), for which no electron density could be detected in the structure. The residues 31-36 (marked in red) are unique to Xoo and absent in other closely related Xanthomonas strains viz., *Xanthomonas oryzicola* (Xoryp20705); *X.campestris* pv. vesicatoria (XCV0536); *X.axonopodis* pv. citrii (XAC0501); *X.campestris* pv. campestris str. ATCC 33913 (XCC2957 & XCC2374) and *Xylella fastidiosa* 9a5c (XF0357, XF0358 & XF2151) LipA-like proteins. This implicates the evolution of a highly flexible region only in Xoo to perhaps, recognize its natural substrate. Note the highly conserved residues (marked in green) flanking the unique region.

Isothermal titration calorimetric (ITC) experiments confirmed these structural predictions. The calorimetric binding of LipA vs. BOG fits into a one-site model (Fig.3.4) while that of LipA vs. β-octyl galactoside could be fitted into a sequential two-site model indicating that the two modes of binding are different. Since the overall binding affinities can be used as indicators of the strength of binding, we compared LipA vs. BOG K_d (76.7 μM) to LipA vs. β-octyl galactoside K_d (739.5 μM; overall K_d for β-octyl galactoside was calculated using the formula $1/\sqrt{K1K2}$ where K_1 was 8.8 x 10^3 and K_2 was 230). The 10-fold reduction in the binding affinity of β-octyl galactoside as compared to BOG is in accord with our proposition that the tunnel has a sugar-anchoring pocket with a predetermined, though moderate, affinity for β-D-glucose.

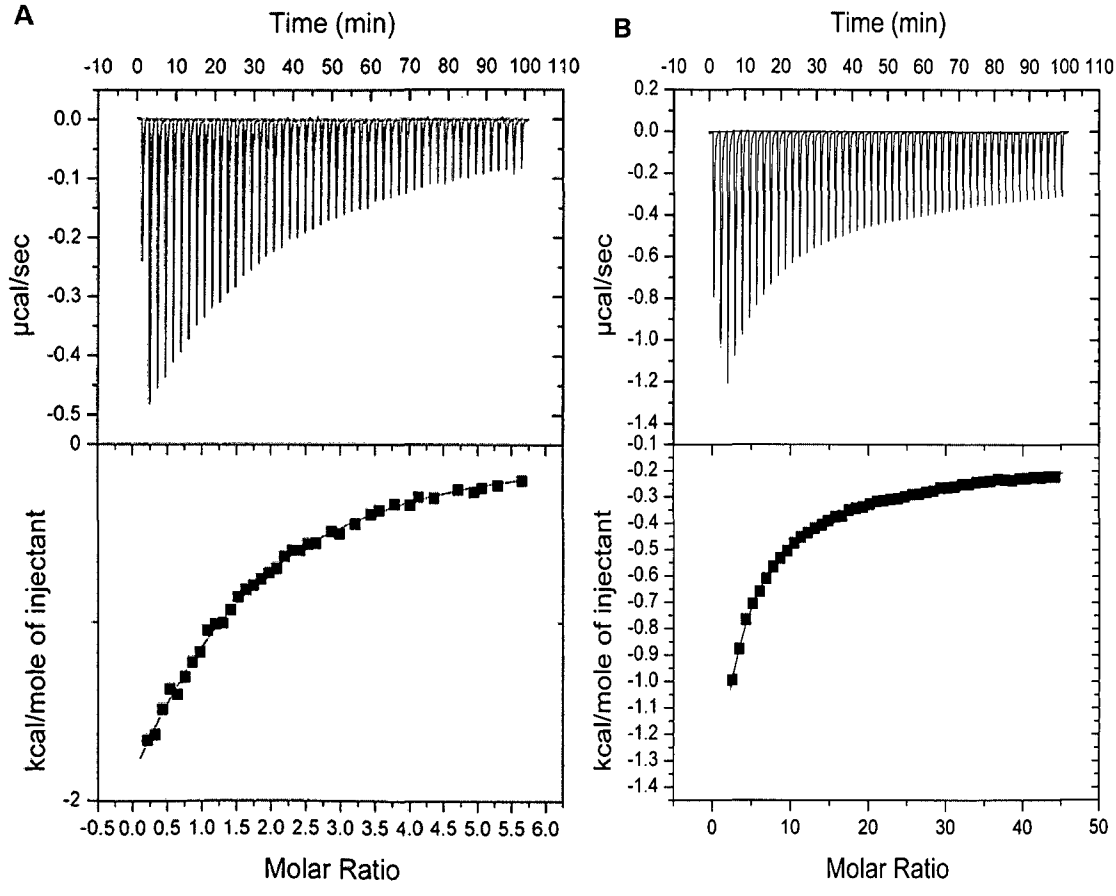


Figure 3.4. ITC analysis of LipA with acyl glucosides. (A) Binding isotherm of LipA titrated against BOG. Top panel: raw titration curve; Bottom panel: Heats fitted by non-linear regressional curve fitting using one site binding model. **(B)** Binding isotherm of LipA titrated against β-octyl galactoside. Top panel: raw titration curve; Bottom panel: Heats fitted by non-linear regressional curve fitting using sequential two-site model.

Apart from the sugar-anchoring pocket, the tunnel is lined with several hydrophobic residues that trace the tunnel from the entrance upto the sugar-anchoring pocket. Presence of this hydrophobic pocket suggests that the moderate specificity conferred by the few hydrogen bonds on the sugar moiety of the ligand is sustained by extensive hydrophobic interaction of the acyl chain with the rest of the tunnel residues. Glycoside ligand binding with recognition pockets for both carbohydrate and acyl chain components is extremely interesting. It also provides a rational for the lack of detection of free sugar binding in ITC assays. An interesting feature in the structure is the presence of at least two serine-histidine conserved clusters on the face opposite to the catalytic triad. These may be non-functional clusters but their position and partial similarity to general esterase active sites is certainly worth noting (Fig.3.5).

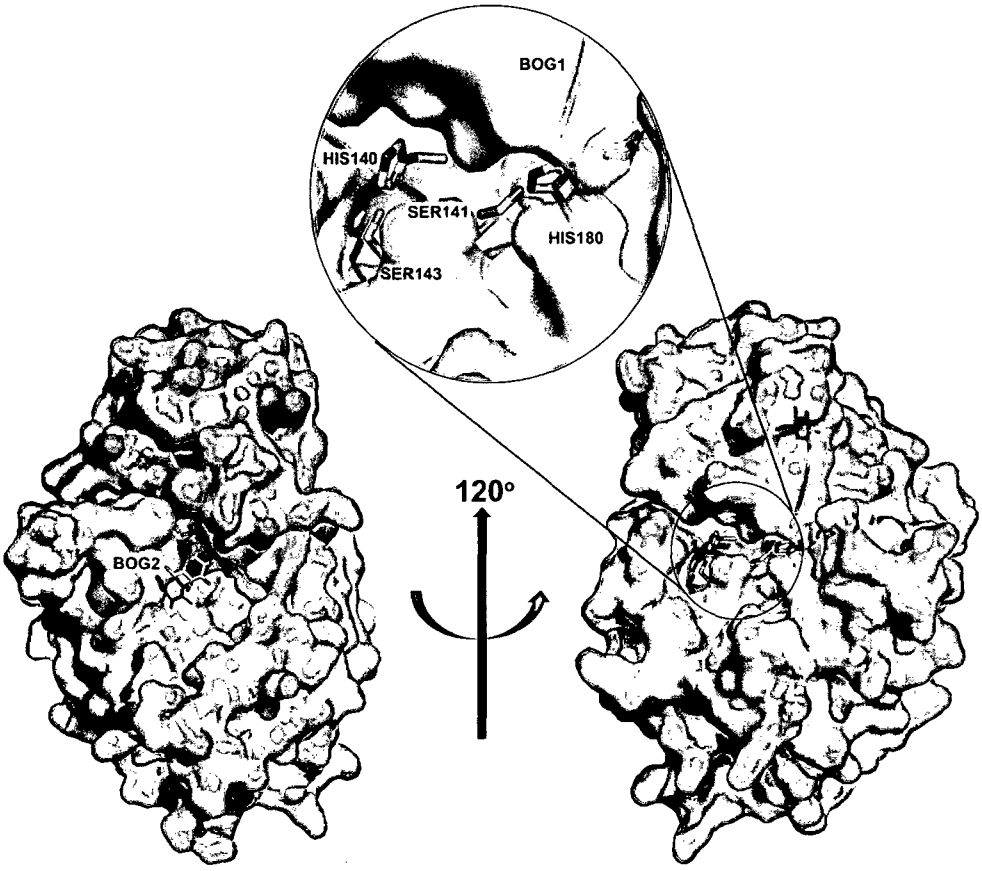


Figure 3.5. Surface view of LipA. This view highlights the entry to the ligand-binding tunnel in the front view (left panel). The ligand-binding domain is represented in green and the hydrolase domain in pink. BOG2 hangs out of LipA and highlights the entry point of the tunnel. Red patch inside the tunnel is the negatively charged patch around S176. A 120° rotation of the front view of LipA shows the two Ser-His clusters away from the active site (right panel). Note the placement of these clusters in solvent-exposed cavities similar to active sites in several other hydrolases. A close-up view of the Ser-His clusters is shown in the inset. A transparent surface is used in the inset and right panel to show BOG1 position.

3.4.2. LipA-like substrate recognition evolved in genus Xanthomonas

A search for LipA sequence homologs using BLAST in the non-redundant database of NCBI identifies a wide range of proteo- and actinobacteria with one or more LipA-like proteins in their genomes (with an E-value of 10^{-7} and lower). A structure-based sequence alignment and a phylogenetic tree was generated using these sequences (Fig.3.6), using CalA as an outgroup. LipA sequences of Xanthomonas strains and *Xylella fastidiosa*, a closely related plant pathogen, cluster together with a high bootstrap value. A small subset of these LipA homologs, consisting solely of the genus Xanthomonas and the organisms demarcated as the blue group in Fig.3.6, contains LipA-like ligand-binding domain. Interestingly, despite a high conservation of residues surrounding the tunnel in this subset, it appears that the carbohydrate recognition is present only in genus Xanthomonas.

The proposition of a conservation of the sugar-anchoring pocket across Xanthomonads is based on our structural analysis indicating Gly 231 as the crucial residue for the sugar ring positioning. Any substitution at this position would protrude into the pocket and abolish sugar anchoring. Gly 231 is invariant in all Xanthomonas LipA proteins with one exception (Fig.3.3a). *Xanthomonas campestris* pv. campestris XCC2374 has G231T substitution although it has an additional LipA-like protein XCC2957 that retains Gly 231.

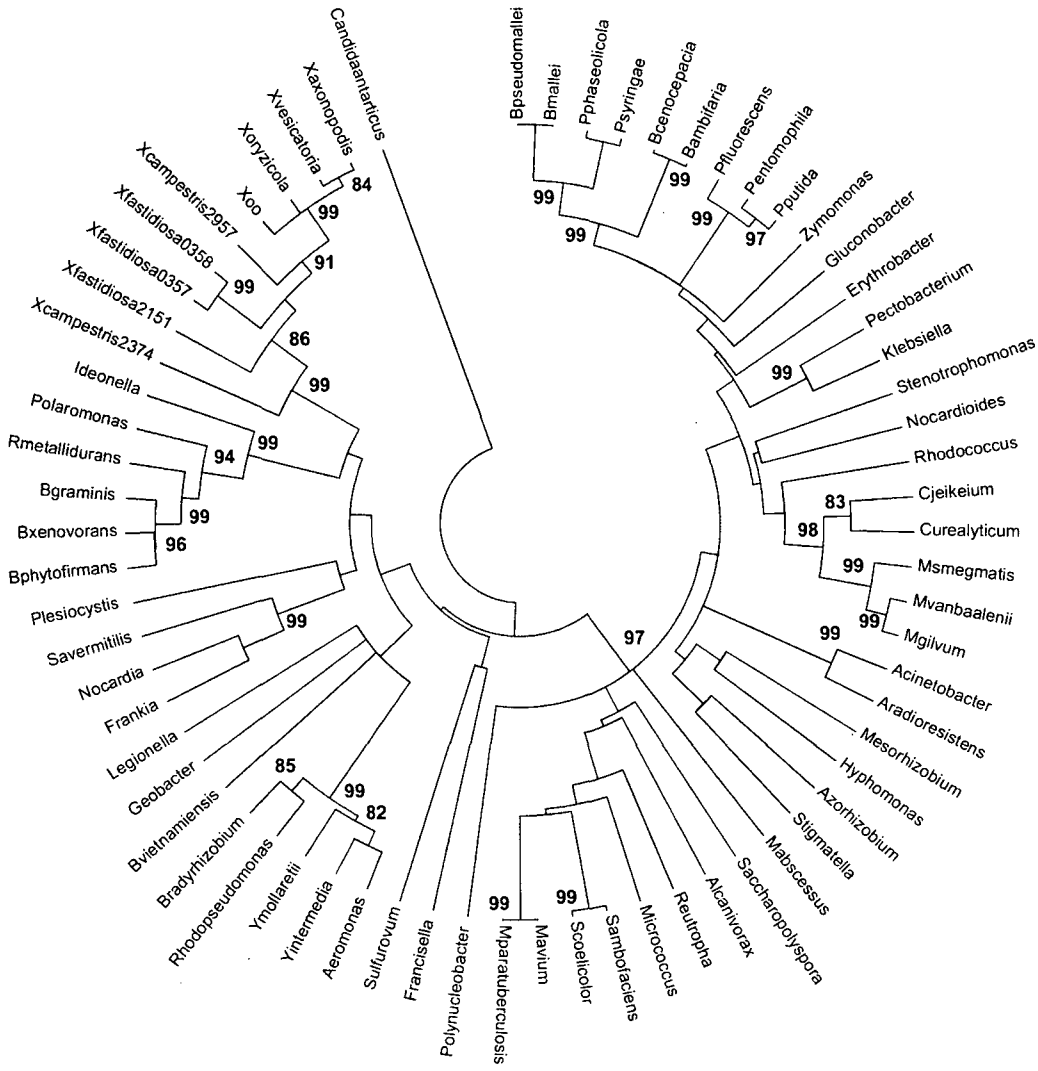


Figure 3.6. Phylogenetic analysis of LipA sequence homologs. Xanthomonas proteins (yellow group) cluster together in the phylogenetic dendrogram of LipA homologs in NCBI database. The yellow and blue groups contain the proteins that were picked up in homology searches using only the LipA substrate-binding domain sequence. Cluster analysis was performed using the UPGMA method after amino acid sequence alignment with ClustalW. The numbers (values >80 are shown) at branches indicate bootstrap values obtained from 1000 replicates. *Candida antarticus* CalA is used as an outgroup. The yellow group contains *Xanthomonas oryzae* pv. oryzicola (Xoryp20705); *X. campestris* pv. vesicatoria (XCV0536); *X. axonopodis* pv. citrii (XAC0501); *X. campestris* pv. campestris str. ATCC 33913 (XCC2957) along with Xoo LipA. The blue group has XCC2957; *Xylella fastidiosa* 9a5c (XF0357, XF0358 & XF2151); *Ideonella* sp. (BAF64544); *Ralstonia metallidurans* (Rmet5769); *Polaromonas napthalenivorans* (Pnap1828); *Burkholderia phytofirmans* (Bphyt4125) and *B. xenovorans* (BxeB0552). Interestingly, homology modelling of LipA-like proteins from the yellow group indicates the presence carbohydrate-binding module in the whole group. LipA-like proteins from the blue group also seem to have a similar tunnel but with either a blocked or obliterated carbohydrate-binding pocket.

A LipA-like protein in the soil bacterium *Ideonella*, with 27% identity to LipA, has been functionally characterized recently (Shinohara et al., 2007). Interestingly, it has been shown to degrade β-hydroxy palmitic (C16) acid methyl ester. A homology model of this protein based on the LipA structure was generated (Fig.3.7). The model indicates the presence of a longer hydrophobic tunnel lacking a carbohydrate-binding site in the LipA-like protein from *Ideonella* and is in agreement with its substrate specificity.

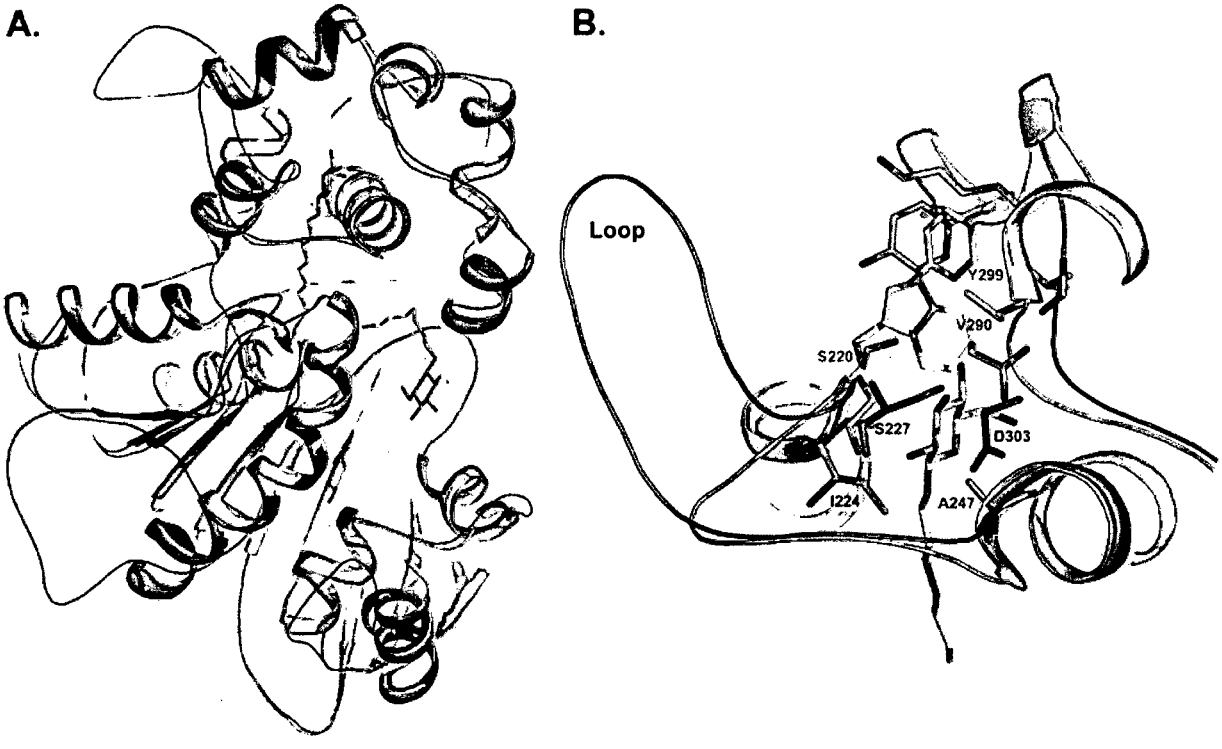


Figure 3.7. Homology model of Ideonella LipA-like protein. A. Superimposition of the homology model of Ideonella protein BAF64544 (green), generated using Modeller9v4, on LipA (yellow). The model shows 26% sequence identity, DALI Z-score of 63.2 and an r.m.s. deviation of 0.7Å over 378 Cα atoms with LipA. BOG molecules are shown as yellow sticks. Note the large loop-like regions in the model amidst very strongly superimposed secondary structure elements of the two structures, indicating regions present only in the Ideonella protein for which no structure prediction could be attained by the software. **B.** A close-up view of the region of the model that corresponds to the LipA ligand-binding domain. The 'Loop' region and the residues marked in green seem to be playing a crucial role in creating a hydrophobic and longer tunnel as compared to the carbohydrate-anchor region of LipA that is delimited by Ser 220, Ile 287 (not shown in the figure), Val 290 and Tyr 299.

The closest structural homolog of LipA is the *Candida antarcticus* lipase CalA, as discussed in Chapter 2. Structural superposition of LipA with CalA illustrates that the PEG molecule bound in CalA structure occupies a very different ligand-binding pocket (Fig.3.8). The analogous region in Xoo LipA is packed with hydrophobic amino acids and the PEG molecule will have substantial clashes with them. In addition, the CalA region that superimposes on the LipA carbohydrate-anchoring pocket is predominantly occupied by the main chain of a loop, indicating that there is no room for such a pocket in CalA. Therefore, it is clear that the pocket in LipA is unique in nature with a specific carbohydrate-anchoring site located far away from the solely acyl-binding pocket of CalA. This finding, along with the phylogenetic study, advocates for LipA-like proteins to be grouped as a distinct class of cell wall degrading esterases.

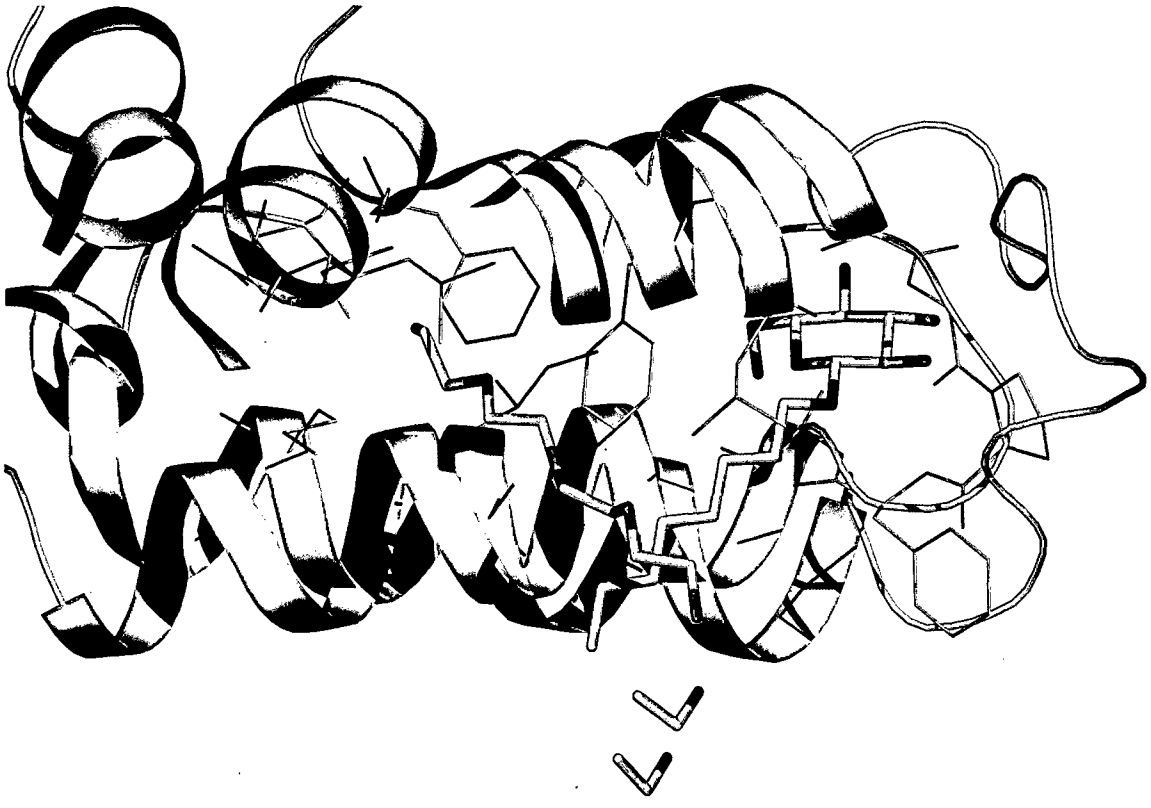


Figure 3.8. Superposition of the lid domains of LipA and CalA. BOG and PEG are shown as yellow and blue sticks respectively. Active site serines for LipA (yellow) and CalA (blue) are shown in yellow and blue respectively.

3.4.3. LipA exhibits esterase and not lipase activity

The presence of an unusually large 7-helical domain in place of the usual lid and the absence of any evident movement upon ligand binding in LipA prompted us to check whether LipA has lipase or esterase activity. It was found that LipA can degrade smaller chain length triacylglycerides like C4 (tributyrin) and C6 (tricaproin) but lacks activity on C8 (tricaprylin) or longer triacylglycerides (Fig.3.9a). It is negative for olive oil (triolein)-rhodamine B assay, for which all true lipases are positive. LipA shows maximum enzymatic activity on p-nitrophenyl (pNP-) butyrate (C4) as compared to pNP-acetate (C2), pNP-hexanoate (C6) and higher chain-length pNP esters (Fig.3.10a). The specific activity plot of LipA with pNP-butyrate (Fig.3.10b) also indicates esterase activity as it shows no evidence for interfacial activation (Martinelle et al., 1995).

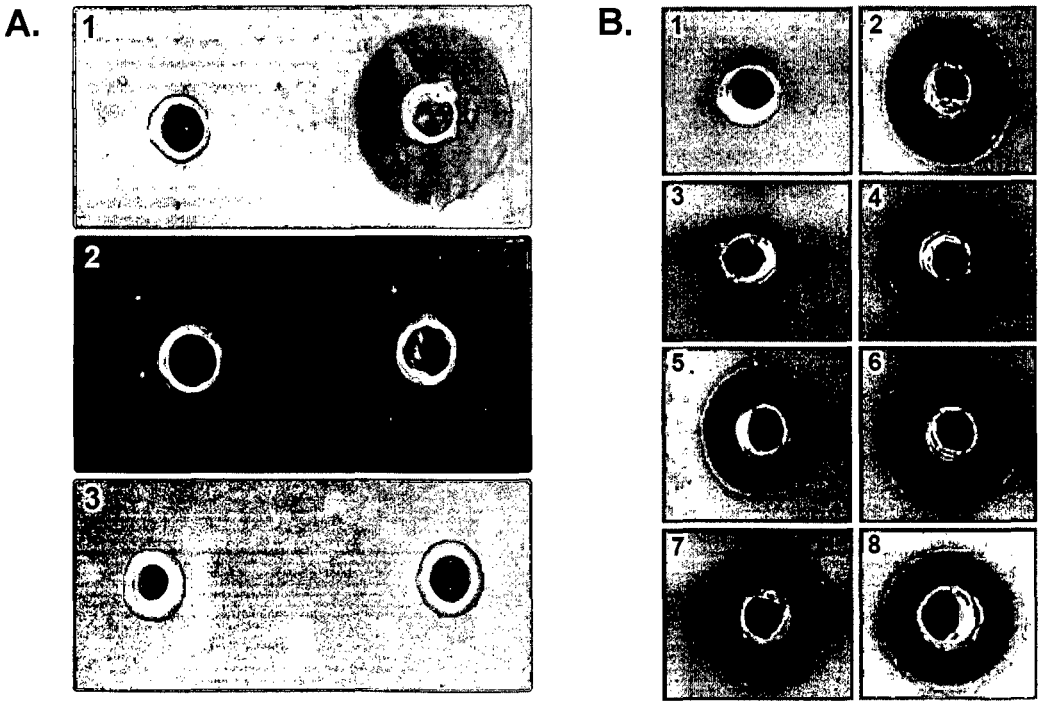


Figure 3.9. LipA exhibits esterase activity (**A**) Presence of zone of clearance indicates LipA activity on short chain triacylglycerides (1) C4 (tributyrin) and (2) C6 (tricaproin) while no activity is seen on (3) C8 (tricaprylin). Holes punched to the right side contain LipA and the left side contain buffer. (**B**) Loss of esterase activity in the S176A active site mutant protein. Tributyrin clearance activity of (1) Buffer; (2) LipA (wild-type) protein; mutant LipA proteins (3) S176A (4) N228W (5) G231A (6) G231I (7) G231F (8) G221I. Plates were photographed after 2hr of incubation at room temperature.

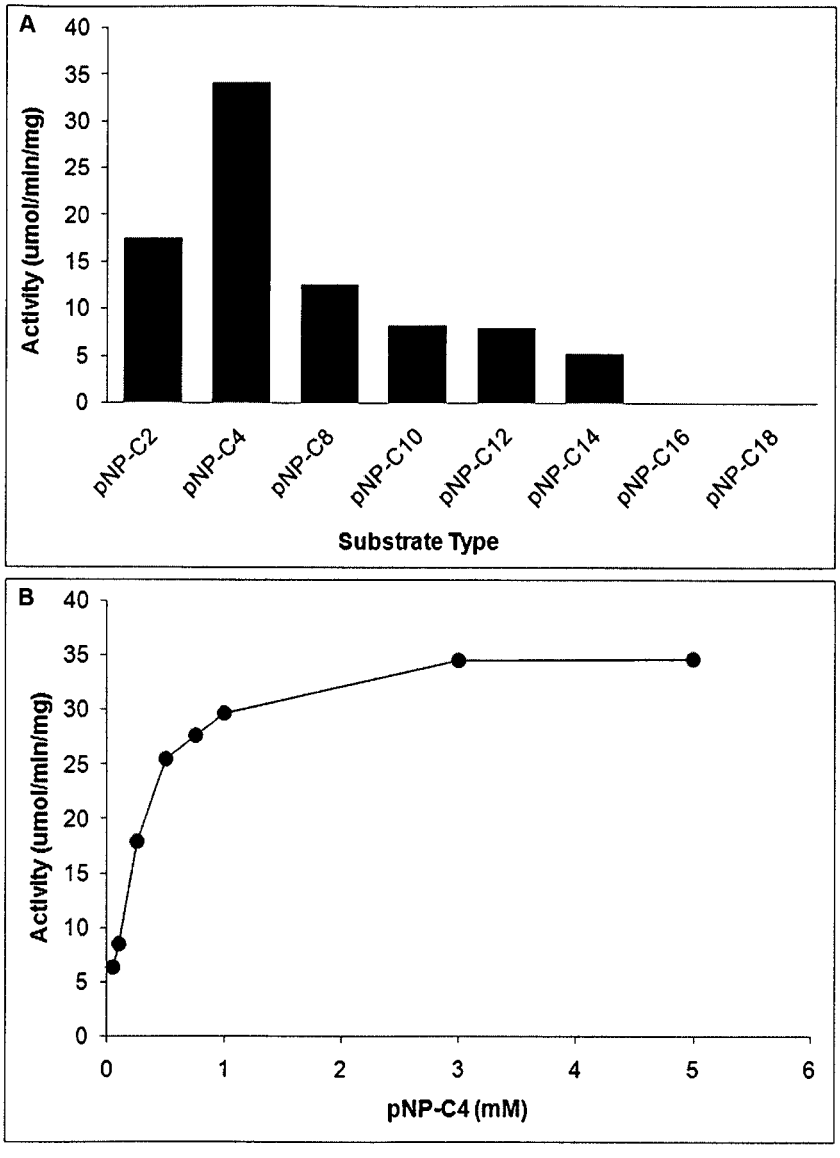


Figure 3.10. LipA substrate specificity and activity curve. (A) LipA shows maximum activity towards p-nitrophenol butyrate (pNP-C4) as compared to pNP acetate (-C2), pNP caprylate (-C8), pNP caprate (-C10), pNP laurate (-C12) and pNP myristate (-C14). No activity was detected for pNP palmitate (-C16) and pNP stearate (-C18). Optimal activity was found at 37°C in Tris-Cl pH 7.5 although the enzyme is active over ranges of pH 5.5 to 9.0 and from 10°C to 70°C. The spectroscopic assays were performed at O.D. 405nm after 10min incubation. **(B)** The activity curve of LipA with varying pNP-C4 concentrations shows an esterase-like behaviour as inferred from the absence of any surge in activity associated with lipase interfacial activation.

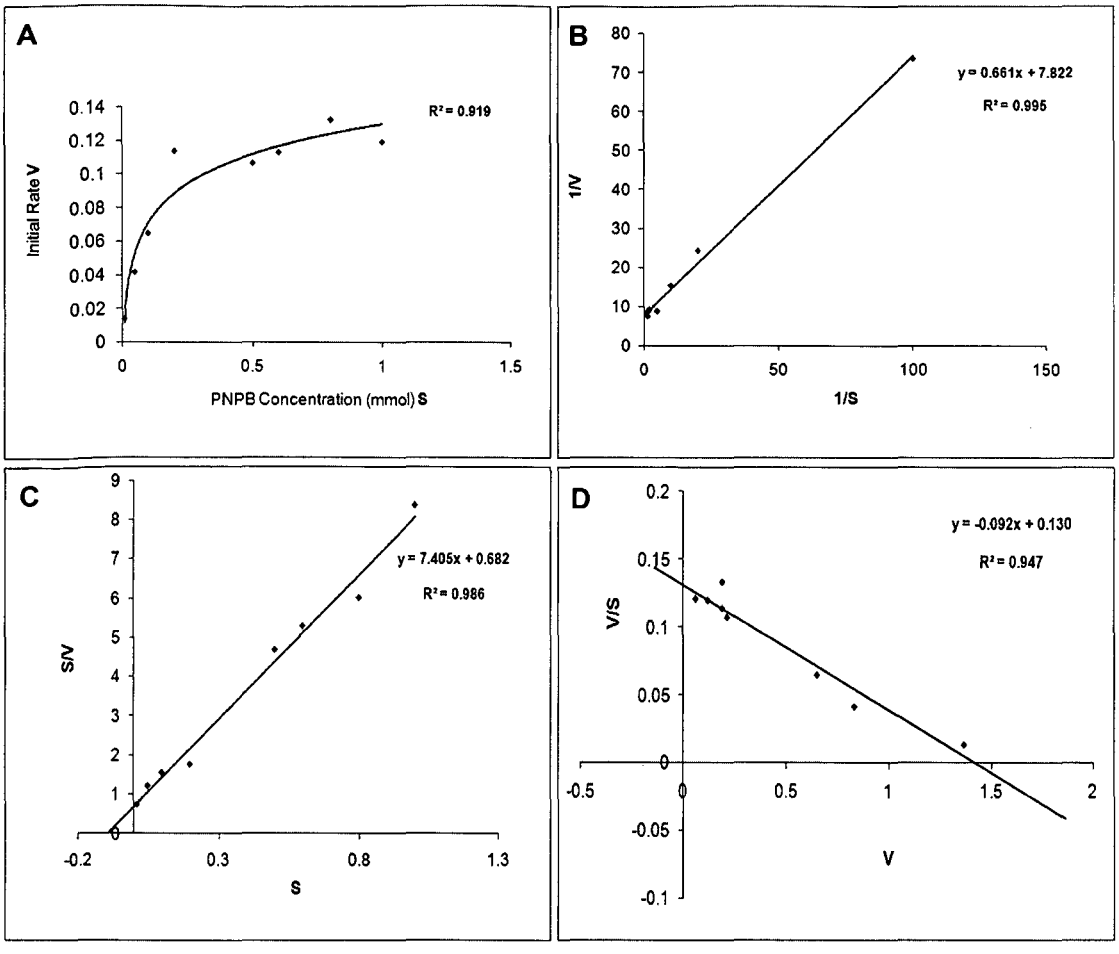


Figure 3.11. LipA enzyme kinetics. This assay was performed using pNP-C4 (p-nitro phenyl butyrate) as the substrate and Km was calculated to be 92 μM and Kcat is 45 s^{-1}. **(A)** Velocity vs. Substrate concentration plot; **(B)** Lineweaver-Burke plot; **(C)** Hanes-Wolfe plot; **(D)** Eadie-Hofstee plot. The line equation and fitting parameters for each plot (R^2) are mentioned.

3.4.4. Point mutations in LipA and assessment of *in vitro* activity

In order to validate our structural prediction that the tunnel is involved in substrate recognition, we generated several point mutations in LipA. Gly 231 was identified as a crucial residue mediating main chain interaction of LipA with BOG. The close proximity of this amino acid with the sugar ring of BOG suggests that the smallest replacement at this position would have a severe effect on LipA action and therefore, Gly 231 was mutated to Ala, Ile and Phe. The residue Asn 228 was mutated to Trp to block the tunnel just below the carbohydrate-anchoring pocket, which would protrude into the acyl-chain binding region. Gly 221, another invariant Gly located in the sugar-anchoring pocket (Fig.3.3a), was mutated to an Ile.

However, Gly 221 has no direct interaction with the sugar ring of BOG (Fig.3.2b). When tested for tributyrin degrading activity, all these mutants show zone of clearance (Fig.3.9b). It clearly indicates that a small substrate such as tributyrin would not require occupying the tunnel upto the sugar pocket that is located ~15Å away from the active site Ser 176. We also checked for enzyme activity in a S176A mutant and absence of a zone of clearance on tributyrin plates indicates that the LipA active site indeed employs the nucleophilic Ser 176.

3.4.5. Verification of BOG binding to LipA and its mutants

The *in vitro* binding of BOG to LipA or its mutants was assessed using ITC. LipA, upon titration with BOG, shows a moderate binding affinity with K_d value 76.7 μM (Fig.3.4). S176A has K_d values of BOG binding similar to that of wild-type LipA, indicating that the tunnel for BOG binding is intact despite a loss in enzymatic activity. All the three Gly 231 mutants display reduced binding of BOG with G231F showing an almost 25-fold reduction from wild-type enzyme (Table 3.4).Therefore, the presence of the tunnel is crucial for BOG binding, which is not affected in the context of the inactive S176A mutant. N228W and G221I bind to BOG with affinities similar (<2 fold reduction) to that of wild-type LipA suggesting that these changes in the sugar binding pocket do not have a significant effect on BOG binding.

Protein	K_a (M^{-1})	K_d (μM) *	ΔH_{app} (Kcal mol^{-1})	$T\Delta S$ (Kcal mol^{-1})	$\Delta G°$ (Kcal mol^{-1})
LipA	$1.3 \times 10^4 \pm 2.9 \times 10^2$	76.7	-5.3 ± 0.05	0.5	-4.8
S176A	$1.2 \times 10^4 \pm 4.5 \times 10^2$	83.3	-3.8 ± 0.07	1.8	-2.0
N228W	$8.8 \times 10^3 \pm 9.2 \times 10^2$	113.6	-2.8 ± 0.15	2.6	-0.2
G231A	$1.7 \times 10^3 \pm 2.2 \times 10^2$	588.2	-0.1 ± 1.0	-5.8	5.7
G231I	$1.3 \times 10^3 \pm 3.5 \times 10^2$	769.2	-0.1 ± 2.3	-6.6	6.5
G231F	$5.1 \times 10^2 \pm 6.2 \times 10^3$	1960.8	-0.01 (#)	-100.9	100.9
G221I	$8.4 \times 10^3 \pm 8.1 \times 10^2$	119.0	-4.8 ± 0.30	0.6	-4.2

* $K_d = 1/K_a$ # Very high error

Table 3.4. Binding parameters and thermodynamics of BOG titration with wild-type LipA and mutants calculated using ITC

3.4.6. LipA ligand-binding domain is essential for virulence on rice

Xoo LipA-deficient strain BXO2001 shows reduced virulence on rice as compared to BXO43, the wild-type strain (Rajeshwari et al., 2005). We mobilized plasmids expressing either wild-type LipA or various point mutants into BXO2001. Expression of LipA protein in these strains was confirmed using polyclonal rabbit anti-LipA antibodies (Fig.3.12). Upon inoculation on rice, the S176A mutant exhibits reduced virulence (at a level that is comparable to the LipA$^-$ BXO2001 strain), confirming that the enzymatic activity of LipA is essential for optimal levels of virulence on rice. Remarkably, the G231A, G231I, G231F and N228W mutants exhibit a virulence deficiency similar to that observed in the case of BXO2001. However, the G221I mutant exhibits wild-type levels of virulence (Fig.3.13). A reduction in virulence caused by either blocking the acyl chain-binding region beneath the sugar anchor site (N228W) or obstructing the sugar-anchoring pocket (Gly 231 mutants) specifically proves that the tunnel is indeed involved in natural substrate binding. Virulence proficiency of the G221I mutation indicates that this does not affect binding of the natural substrate in the host as indicated by *in vitro* binding studies also.

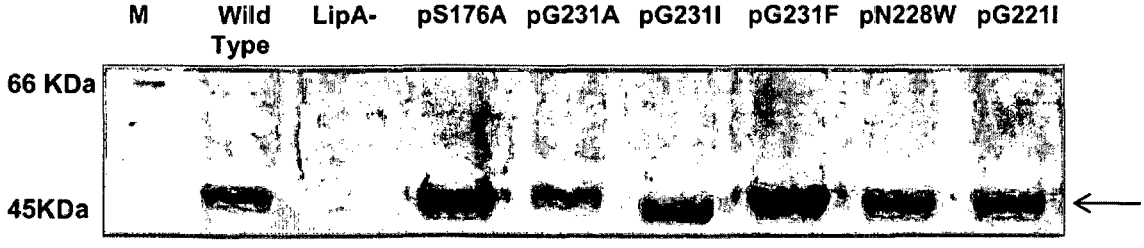


Figure 3.12. Western blot analysis of wild-type and mutant LipA proteins. Expression of LipA protein in the culture supernatants of various Xoo strains was assessed using polyclonal rabbit anti-LipA antibodies. Wild-type (Xoo strain BXO43), LipA$^-$ (Xoo strain BXO2001) and Xoo strains that express various mutant LipA proteins from the pHM1 plasmid in the BXO2001 background; M: marker.The arrow points to the LipA-specific band.

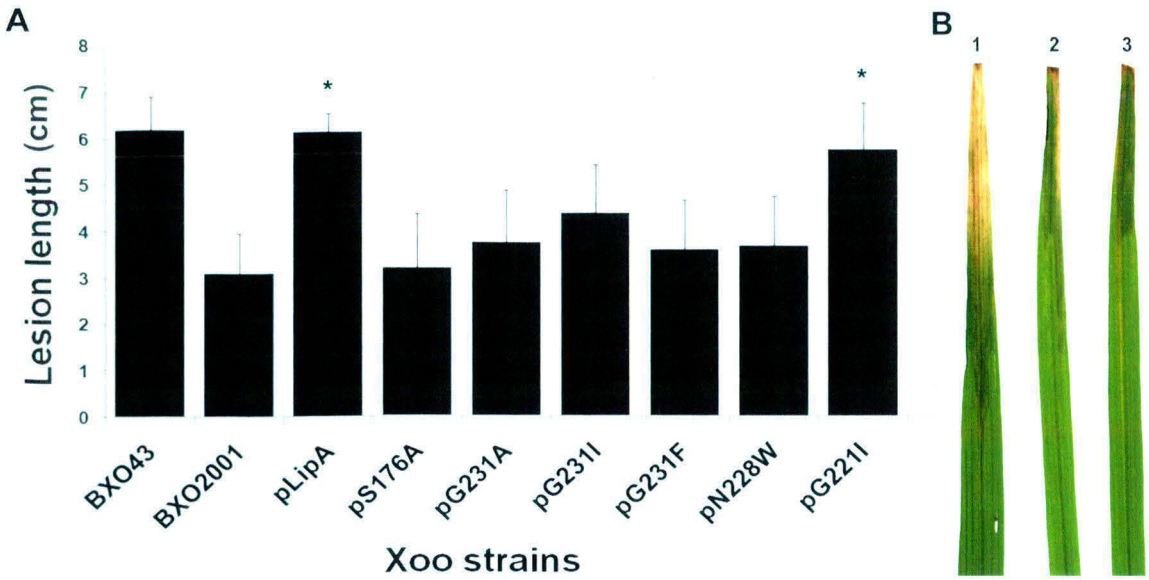
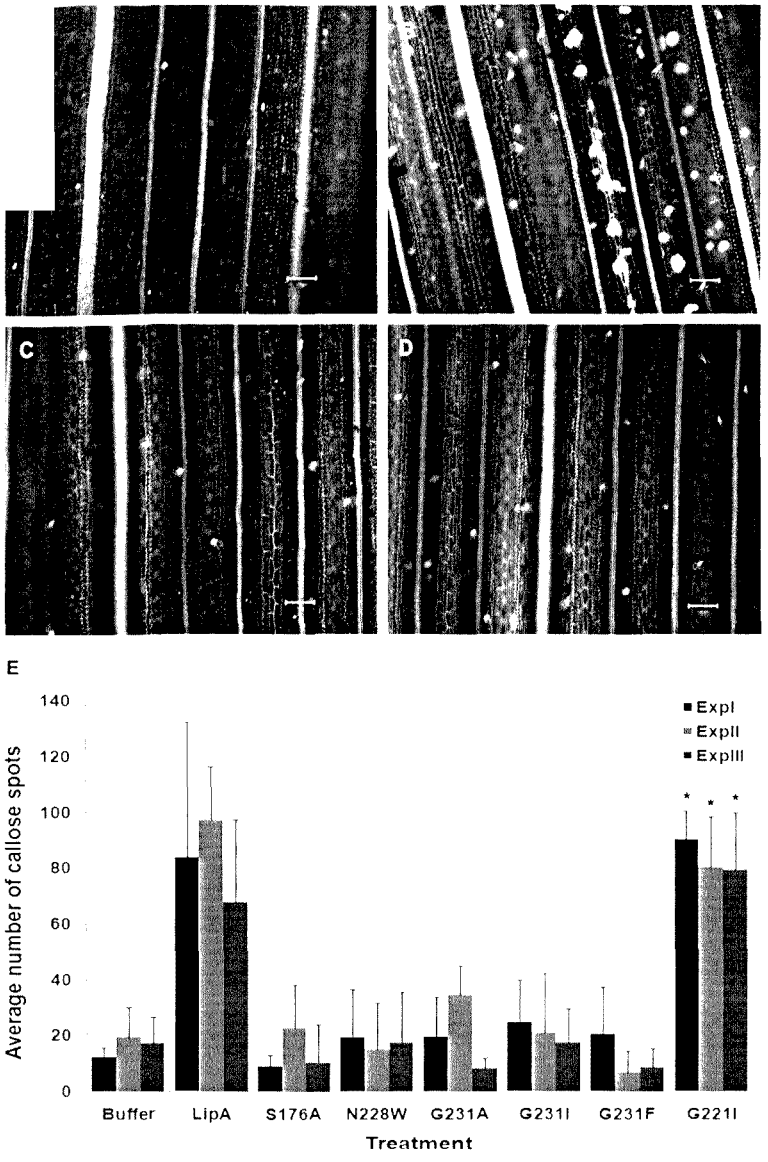


Figure 3.13. Virulence phenotypes of LipA mutants (A) Xoo strains were inoculated on rice leaves and lesion lengths were measured after 7 days. Atleast 10 leaves were used per strain. Similar results were obtained in another independent experiment. Values marked with an asterisk are not significantly different from the values obtained for BXO43 at $p<0.05$ in a student's two-tailed T-test for independent means. BXO43 (Wild-type Xoo), BXO2001 ($LipA^-$ mutant) and various strains that express either wild-type LipA or mutant LipA proteins from the pHM1 plasmid in the BXO2001 background. **(B)** Photographs of lesions caused by 1. BXO43; 2. BXO2001 and 3. pG231F. The latter two strains exhibit reduced disease symptoms (brown lesions) as compared to BXO43.

3.4.7. Loss of induction of rice innate immunity

The ability of LipA to degrade rice cell walls has been correlated with its capability to induce host innate immune responses (Jha et al., 2007). Cell wall fortification by rapid deposition of α-1, 3-glucan polymer called callose at the site of pathogen entry is a marker of plant defense response (Bestwick et al., 1995). The ability of various LipA mutants to induce callose deposition in rice leaves was investigated. Wild-type and mutant LipA proteins were purified and infiltrated into rice leaves. Callose deposits appear as fluorescent spots upon staining with aniline blue and observed under UV light. Infiltration with the G231A, G231I, G231F and N228W mutants shows background levels of callose deposition, similar to that observed in buffer treated leaves, while leaves infiltrated with G221I exhibit levels of callose deposition that are no different from those observed in leaves infiltrated with wild-type LipA (Fig.3.14).



gure 3.14. LipA mutant proteins are deficient at induction of defense response associated llose deposition in rice leaves. Light microscope images (10X resolution) of rice leaves infiltrated th (**A**) Buffer; (**B**) LipA (wild-type); (**C**) S176A; (**D**) G231F. Purified proteins were infiltrated into e leaves and stained for callose deposition. The callose deposits appear as sharp bright fluorescent ɔts as against the dull white spots representing the trichomes. (**E**) Number of callose deposits in leaf nes infiltrated with buffer, wild-type LipA or various mutant proteins. Mean and standard deviation ·re calculated for number of callose deposits from a leaf area of 0.60 mm^2. Data were collected from ır leaves in each experiment and four different viewing areas from the infiltrated region of each leaf. terisks indicate that the values obtained after treatment with G221I are not significantly different at 0.05, when compared with wild-type LipA.

Localized programmed cell death (PCD) is another important host defense response to pathogen attack (Pennell & Lamb, 1997). Rice roots were treated with the purified wild-type and mutant LipA proteins, stained with propidium iodide (PI) and examined by confocal laser scanning microscopy. PI is excluded from live cells and its internalisation is indicative of cell death. Rice roots treated with wild-type LipA take up PI and the PI staining material is dispersed within the cell. This is indicative of DNA fragmentation caused by PCD. S176A-treated roots showed absence of PCD. Similarly, the Gly 231 mutants and the N228W mutant resembled buffer-treated roots with no evidence of LipA induced PCD (Fig.3.15).

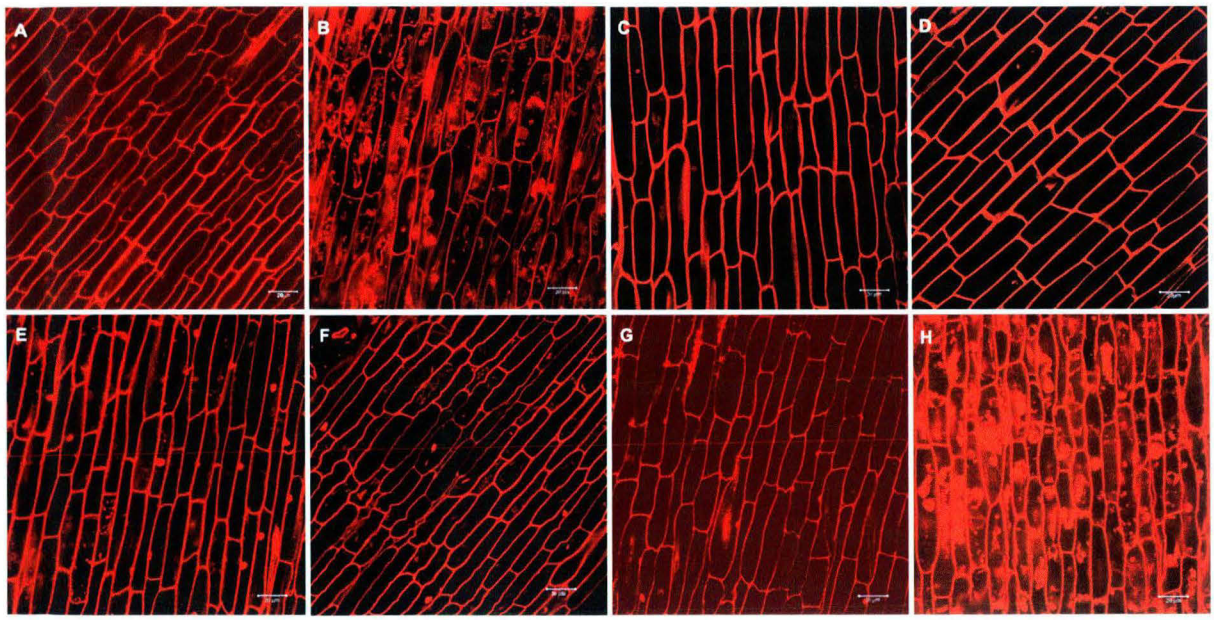


Figure 3.15. LipA mutant proteins are deficient at induction of defense response associated programmed cell death in rice roots. Rice roots were treated with purified proteins (either wild-type LipA or mutants), stained with propidium iodide (PI) and examined under a confocal microscope to assess the extent of DNA fragmentation. The control buffer-treated roots (**A**) exhibit a prominent cell wall associated autofluorescence but no internalization of PI into the cells. Treatment with either wild-type LipA (**B**) or G221I (**H**) resulted in cell death (intake of PI) accompanied by dispersed intracellular staining which is indicative of nuclear fragmentation. No cell death is seen in roots treated with S176A (**C**); N228W (**D**); G231A (**E**); G231I (**F**) and G231F mutant proteins (**G**).

The three Gly 231 mutants (to A, I and F) have similar *in planta* phenotypes, proving that substrate binding in the natural context would be perturbed by any bigger residue substitution that might protrude into the sugar binding pocket. The N228W mutation affects LipA function

in planta indicating that it disrupts binding to the natural substrate. However, this mutation does not affect BOG binding, possibly because BOG might be smaller than the natural substrate. The G221 mutant protein is able to induce PCD. The lack of any *in planta* phenotypes associated with the G221I mutation suggests that this change can be tolerated, possibly because of the presence of this residue on a less-packed region.

3.4.8. LipA point-mutant structures confirm the structural basis of ligand binding

Two mutants of LipA (G231F and N228W) were soaked with BOG at the same concentration as used for obtaining wild-type LipA-BOG co-crystals. $2F_{obs}$-F_{calc} map contoured at 1σ before the inclusion of the mutated residue in the refinement clearly shows the protrusion of the bulkier residue in the binding pocket. G231F and N228W structures could be superimposed with wild-type LipA with an r.m.s. deviation of 0.10Å and 0.38Å, respectively for 387 Cα atoms. Both structures show absence of BOG density in the lid-like domain, clearly due to the obstruction of the carbohydrate pocket in G231F (Fig.3.16a) and a blocking of the tunnel region by Trp in N228W (Fig.3.16b). Together, the mutant structures confirm that the phenotypes associated with these mutants are a direct consequence of blocking the pocket and not due to any major structural changes in the protein.

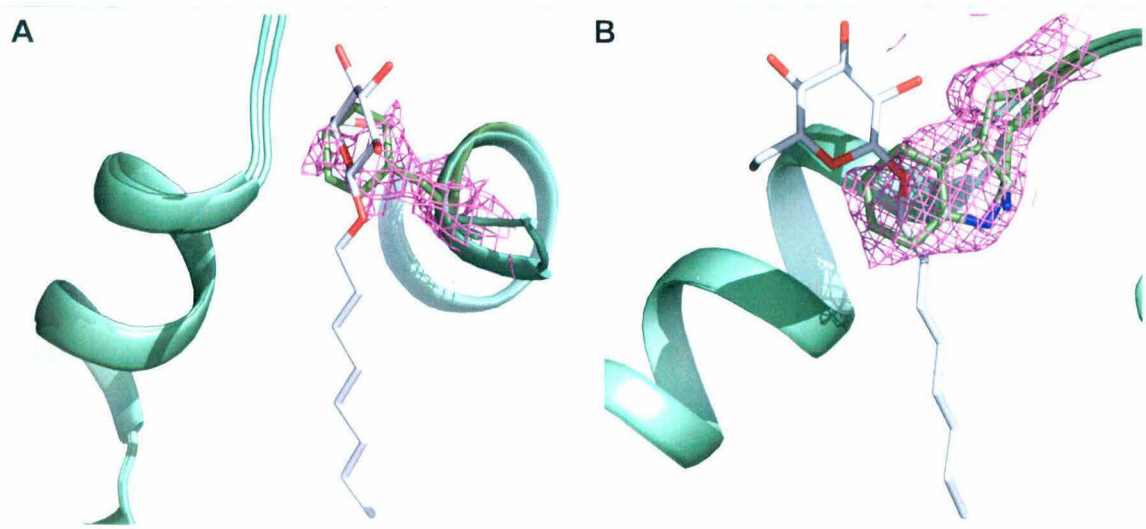


Figure 3.16. The LipA ligand-binding tunnel is blocked in the G231F and N228W mutant proteins of LipA. **(A)** An unbiased $2F_{obs}$-F_{calc} map contoured at 1s demonstrates disruption of the glucosyl ring-binding site in the G231F mutant. **(B)** N228W obstructs the tunnel as shown in the unbiased $2F_{obs}$-F_{calc} map contoured at 1σ electron density around the mutant residue. The mutant structures have been superimposed with the wild-type LipA structure in both figures.

3.5. Discussion & Implications of the study

The host environment imposes a strong selective pressure on bacterial pathogens for the evolution of a repertoire of divergent and specialised proteins that promote colonization and survival within their eukaryotic hosts. Gain of new functions in proteins is frequently accomplished by modulation of the pre-existing functionally versatile folds (Thornton et al., 1999; Goldstein, 2008). The ubiquity of the α/β hydrolases is also a consequence of the versatility of the fold. Xoo LipA crystal structure reveals a hitherto unidentified accessory glycoside-binding domain on a canonical α/β hydrolase domain, fine-tuning esterase activity for a plant-associated function. Structural comparison with CalA clearly demonstrates the lid domain has been functionally converted into a cell wall degrading esterase by engineering distinct carbohydrate and acyl chain-recognition pockets. This is a remarkable example of acquisition of a specialized function for increased proficiency in pathogenesis using an existing scaffold, to act on a complex polysaccharide-rich plant milieu of the host in this case.

The BOG-bound structure of LipA defines a clear path and space for substrate entry in the ligand-binding domain, an 18Å-wide opening that narrows to 8Å near the active site and remains so up to the end of the tunnel. The BOG complex, although obtained fortuitously, suggests that the natural ligand of LipA is very likely to have a stereochemistry resembling 'monosaccharide-alkyl chain-ester linkage-', since the Gly 231 residue found to be important for BOG-LipA interaction also plays a crucial role in the biological activity of LipA. The fact that the mutation of Asn 228 does not affect BOG binding while it disrupts *in planta* activity of LipA shows that there are differences between BOG and the natural ligand of LipA. Also supporting this architecture of the natural ligand is the fact that β-octyl galactoside shows much weaker binding to LipA in ITC experiments as compared with BOG, indicating that LipA has a higher specificity for β-D-glucoside as compared to β-D-galactoside. Therefore, mechanistically, LipA might act on an amphiphilic molecule with a glucose (or perhaps, xylose) moiety attached to a long (substituted) acyl chain (or aryl ring) of a length of 16-18 carbons (~30Å) with an ester bond situated ~15Å from the sugar ring. Presence of an aliphatic chain in the natural ligand is indicated by the stretch of hydrophobic residues lining the LipA ligand-binding tunnel just below the carbohydrate-anchor site as well as by the activity of LipA on p-nitrophenol acyl esters.

LipA could act on alkyl ester crosslinks between long polysaccharide chains, which need to be cleaved to make the polysaccharide chains accessible to other degradative enzymes. LipA may also chew the same chain at several points and thus may be able to accept substrates of varying lengths. Such a compound can be predicted to be a part of the lignin component of the rice cell walls, since short alkyl chains and alkyl groups esterified to aryl moieties are reported in the lignocellulosic crosslinks that connect the polysaccharide chains (reviewed in Buranov & Mazza, 2008). Several cyclo-alkyl rings linked to glucose, for example furostanol glucosides, have been identified from plant cell walls and are possible candidates for LipA action (Arthan et al., 2005). A comparison with the inhibitor-bound structures of other hydrolases in the PDB indicates that, upon binding to the substrate-binding tunnel in LipA, the carboxylate part of the ester would probably occupy the tunnel while the alcohol part would hang out of the enzyme. The alcohol part of the ester can be of varied characteristics since the LipA tunnel opens very wide towards the BOG2 binding site. Existing knowledge of the precise ester-linked polymers in plant cell walls, especially rice, is limited owing to their enormous complexity (Carpita, 1996; Somerville et al., 2004; Knox, 2008). However, as our understanding of the architecture of the plant cell wall improves, clarity on the selectivity aspect of LipA function mediated by the substrate-binding pocket will also increase.

This study demonstrates that the disruption of either the catalytic serine or the sugar-anchoring site 15Å away leads to the abrogation of LipA *in planta* functions. Therefore, LipA function requires grabbing a cell wall component, resembling BOG, in the ligand-binding tunnel and cleaving the ester linkage at the active site, illustrating the existence of a crucial functional interplay between these two distinct regions of the molecule. The cell wall damage and/or release of soluble elicitors is the cue for induction of host innate defense responses and a loss of LipA activity leads to an inability to provoke these responses. This substantiates the earlier observations showing that heat-inactivated LipA failed to induce rice innate immunity (Jha et al., 2007). Therefore, the result of LipA action, and not LipA molecule itself, is required for eliciting rice responses like PCD and callose deposition. Contrastingly, fungal xylanases can elicit plant defense responses through physically binding to a plant receptor (Ron & Avni, 2004).

LipA mediated Xoo pathogenesis is dependent on the presence of the glycoside recognition module. It is highly probable that the whole genus Xanthomonas employs this mode of substrate recognition for proficient pathogenesis. Recognition of both sugar and hydrophobic

components are essential for LipA function. Weak interactions between the BOG glucose and LipA are substantiated with a rigid open pocket and a long hydrophobic patch to hold the acyl chain. The carbohydrate ligand-binding domain of LipA is very distinct when compared to the carbohydrate-binding modules in the Carbohydrate Active Enzymes database (Cantarel, 2009). In general, non-catalytic carbohydrate-binding modules are reported to enhance the activity of their catalytic counterparts by facilitating substrate proximity and increasing the efficiency of substrate binding in plants (Boraston et al., 2004). This study shows that LipA substrate-binding domain is an essential modular non-catalytic domain appended to the catalytic hydrolase to recognize acyl glycoside substrates where the ester bond is located quite far away (15Å) from the sugar binding site. Taking into account the absence of structural homologs for the glycoside recognition domain and its conservation across the Xanthomonas group, LipA-like proteins constitute a new class of cell wall degrading enzymes with a unique mode of substrate recognition. Identification or generation of such distinct plant biomass degrading enzymes is an active area of research (Gilbert et al., 2008). Addition of LipA-like enzymes to the existing cell wall degrading enzymatic preparations provides a new dimension for the production of clean eco-friendly biofuels from agricultural wastes and other sources of plant material to improve the efficiency of the process.

LipA glycoside-binding domain is present in Xanthomonas pathogens that infect commercial crops like tomato, citrus, cabbage etc. This domain may have been acquired by a common ancestor of Xanthomonas and Xylella strains from evolutionarily distant but ecologically co-existing non-pathogenic bacteria like the soil/plant-associated Burkholderia and fine-tuned to act on carbohydrate-linked carboxylic esters. Diverse complex polysaccharides esterified to largely uncharacterized phenolics and aliphatics dominate the plant and soil interface environments, perhaps, enforcing evolution of such specializations in the hydrolase fold. This study offers the first glimpse into a plant habitat-specific adaptation of the common esterase activity and provides a solid platform to study secretory esterases from several other important pathogenic bacteria. It might also provide an opportunity to develop a battery of small molecule inhibitors for this family of enzymes, which could be developed for potential application in crop fields.

Chapter Four

In silico sequence analysis and homology modelling of Xanthomonas adhesin-like protein A (XadA) from *Xanthomonas oryzae* pv. oryzae

CHAPTER FOUR

4.1 Abstract

Xanthomonas adhesin-like protein A (XadA) is an outer membrane-located protein involved in the initial attachment of Xoo to rice leaves and hydathodal entry. XadA has high amino acid sequence homology to adhesins of animal pathogenic bacteria. It possesses sequence features of a Trimeric Autotransporter Adhesin (TAA), containing a C-terminal transporter region and an N-terminal passenger domain. The putative transporter region of XadA was found to have high homology to the corresponding region of the collagen-binding *Yersinia enterocolitica* TAA YadA. In addition, XadA passenger domain can be divided into four distinct regions, each homologous to the YadA passenger domain. Homology models of the four regions of XadA passenger domain were generated and each domain has a 'head' made of left-handed parallel beta-roll repeats and a transition loop region called the 'neck', which connects two tandem heads. XadA structure was found to be considerably different from that of YadA since the 'head-neck' of YadA is followed by a coiled-coil region called the 'stalk', which is missing from three out of the four 'head-neck' domains of XadA. Overall, XadA can be speculated as being a long trimeric molecule with four distinct domains for binding cognate partner(s) in rice, intermediate loop regions aiding in this binding. Comparison of XadA orthologs from the genomes of related Xanthomonas species shows similarity in the overall architecture of the adhesins. However, variability in the number of 'head-neck' domains of XadA among different species might be indicative of subtle differences in binding partners in plant hosts and be instrumental in bringing about host and tissue specificity.

4.2 Introduction

The first stage of infection is colonization of the host by the pathogen. The most favoured portals of pathogen entry are the openings through which host tissues are exposed to the external environment. Surface molecules and organelles involved in microbial adherence to the host are pivotal to pathogenesis since proper adhesion determines the chances of infecting the correct target tissues. The number and complexity of the paraphernalia on a pathogen surface is astounding and considered essential for the successful conquest of host niches and establishment of disease (Pizarro-Cerda & Cossart, 2006).

Gram-negative bacterial pathogens assemble adhesive organelles called pili on their surfaces. Pili, also known as fimbriae, are hair-like appendages that can assume two different morphologies; Type I and P pili are rod-like fibres with a diameter of ~7 nm, whereas the other type consists of thin flexible 2-5 nm wide fibrillae (de Graaf and Mooi, 1986). The host receptor-binding adhesin is exclusively present on the pilus tip and is presented on the bacterial surface by a multi-repeat pili subunit. While the adhesins on the pili tips determine host, tissue and cell-type specificity, the pili rods have been implicated in binding host tissue matrix proteins non-specifically to increase the efficiency of the binding (Hultgren et al., 1993).

Apart from this type of 'fimbrial' adhesins present on polymeric pili, there are scores of afimbrial or non-pilus adhesins on the pathogen surface that are monomeric or oligomeric proteins but never heteropolymeric in nature. Well-studied examples of afimbrial adhesins are mostly from animal pathogenic bacteria infecting human or other mammalian cells. These include the adhesins AfaD and AfaE from *Escherichia coli* involved in its attachment to the urinary tract, the filamentous hemagglutinin (FHA) from *Bordetella pertussis*, responsible for attachment to the lung epithelial and phagocytic cells and the opacity proteins (Opas) from Neisseria that are considered responsible for cell-type specificity (Finlay & Cossart, 1997). Simultaneous expression of several types of fimbrial and afimbrial adhesins indicates that pathogenic bacteria optimize the host adhesion by binding to a large variety of host surface molecules. The host cells, however, are not mere inert surfaces and sense the adhesin-receptor binding by a series of signalling events. Most surprisingly, in some cases, bacterial pathogens bind to host surface proteins that are expressed downstream of activation of the host cells following detection of bacterial entry.

4.2.1 Adhesin-like functions of Xoo

In silico analyses show the presence of multiple fimbrial and afimbrial adhesin-like functions in the Xoo genome. Xanthomonas adhesin-like protein A (XadA) was identified as an afimbrial adhesin with high sequence homology to *Yersinia enterocolitica* afimbrial adhesin YadA (Ray et al., 2002). In a later study from our lab, a paralog of XadA called XadB, a protein homologous to the Yersinia virulence factor YapH and a type IV pilus were also identified as possessing adhesin functions (Das et al., 2009).

XadA-deficient mutant Xoo strain was found to be virulence-deficient when allowed to infect rice leaves through the hydathodes (epiphytic inoculation, where the rice leaves are dipped in the bacterial culture). However, upon infection through wounding (clip inoculation, where scissors dipped in bacterial cultures are used to clip the rice leaves), XadA-deficient mutant strain causes disease to an extent similar to the wild-type Xoo (Ray et al., 2002). An ingenious confocal microscopy assay was developed to assess the efficiency of rice leaf attachment and entry of EGFP-tagged Xoo strains. In this assay, Xoo cells were found to enter into rice leaves within the first 1h of infection and that the colonization and leaf entry occurs preferentially via the dorsal where the hydathodes are particularly concentrated (Das et al., 2009; Mew et al. 1984). Upon comparison of the ability of various Xoo mutant strains to attach to and enter rice leaves, the number of bacterial cells within the first 1mm of the leaf tip were counted. It is evident from the virulence assays that XadA-deficient mutants exhibited a deficiency in leaf attachment and entry in the confocal assay while XadB-deficient strain exhibited a less severe deficiency. A *xadA- xadB-* double mutant exhibited an even greater deficiency in leaf attachment and entry when compared with either the *xadA-* or *the xadB-* single mutants. This additive effect indicates that both genes are independently contributing to leaf attachment and entry. However, the XadB-deficient mutant strain of Xoo shows no virulence deficiency in any type of inoculation on rice leaves. A *xadA- xadB-* double mutant strain is more virulence deficient than *xadB-* mutant but less virulent than *xadA-* mutant strain. Together, these data suggest that XadA and XadB proteins promote Xoo virulence when Xoo enters through the natural mode of Xoo entry through the hydathodes and not when the bacterium can bypass this route by entering through wounds. In the same study, Xoo YapH-deficient strain and the type IV pilus protein PilQ-deficient mutants were found to be more important in Xoo attachment and entry through wounds and even the *in planta* migration of the bacterium inside the rice xylem vessels (Das et al., 2009).

Apart from establishing the role of several adhesins in a plant-pathogen interaction, this study also revealed that these adhesins have redundant functional roles since none of the Xoo strains deficient in one or more adhesins is completely virulence deficient. This opens up a possibility of the presence of other uncharacterized adhesins in the Xoo genome and underlines the requirement for a complete characterization of the structural and functional features of these molecules. The focus of this chapter is on the sequence features of XadA that indicate a great deal about its functional role.

4.2.2 Preliminary characterization of XadA

XadA is a 1265 amino acid long, outer membrane-located protein expressed only in a minimal medium and not the standard laboratory rich medium used for growing Xoo (Peptone-sucrose) (Ray et al., 2002). XadA orthologs have been reported from the genomes of the two sequenced strains of Xoo (KACC10331 and MAFF311018) as well as from the genomes of all the sequenced Xanthomonas strains such as Xac, Xcc and *Xylella fastidiosa* (Simpson et al., 2000; da Silva et al., 2002; Lee et al., 2005; Ochiai et al., 2005).

XadA sequence was observed to be rich in alanine (22.3%), glycine (15.3%), serine (10.1%), valine (9.6%) and threonine (8.1%). Surprisingly, XadA, being a protein of relatively large size, lacks cysteines. The presence of either a derivative of the sequence TAVG or a nine-amino-acid derivative of TDAVNVAQL several times in XadA with little variation has also been noted. Homology of XadA to afimbrial adhesins of animal pathogenic bacteria, such as YadA of *Yersinia enterocolitica* and UspA1 of *Moraxella catarrhalis* was noted to be such that YadA/UspA1 homologous regions repeat several times in the XadA sequence. XadA N-terminal shows the presence of an unusually long putative signal peptide of 66 amino acids. The C-terminal has a stretch of hydrophobic amino acids that form the membrane anchor region in a class of afimbrial adhesins known as the autotransporters to which class belong YadA and UspA1 (Struyve et al., 1991; Henderson et al., 1998). In order to understand the relevance of these interesting observations, the structures of some of the autotransporter adhesins would be described first and then, compared with the known and more newly identified sequence features of XadA.

4.2.3 Afimbrial autotransporter adhesins

As the name suggests, 'autotransporters' are proteins that carry the whole information regarding their export and secretion through the bacterial cell envelope. Several proteins classified as afimbrial adhesins are known to be secreted through this elegant transport mechanism.

In Gram-positive bacteria, secreted proteins are commonly translocated across the single membrane by the Sec pathway that helps secrete unfolded/folding proteins or the two-arginine (Tat) pathway that secretes folded large complexes of enzyme-cofactors etc. However, Gram-negative bacteria have five distinct modes of protein secretion classified based on the mechanism of crossing the cell envelope that is made up of an inner membrane, a periplasmic space and an outer membrane. The type I, II, III and IV secretion systems have distinct mechanisms of secretion of proteins, polysaccharides, DNA etc. across the inner and/or outer membranes, through pores formed by multimeric complexes of several proteins (secretion systems reviewed in Tseng et al., 2009).

The type V secretion system is made up of 'autotransporter' proteins that contain the instructions for export and secretion in their amino acid sequences (reviewed in Dautin & Bernstein, 2007). Autotransporters are a large and diverse group, with almost 800 identified sequences to date, comprising of distinct N-terminal passenger domain and C-terminal membrane anchoring transporter domain. The transporter domain forms a β-barrel structure in the outer membrane (therefore called β-domain) and facilitates proper folding and secretion of the passenger domain, although the exact mechanism remains poorly understood (Pohlner et al., 1987; Bernstein, 2007). These proteins are generally dependent on the Sec pathway for inner membrane transport. The β-domain is the only common feature of all autotransporters and defines the family. The wide diversity in this family is due to the passenger domains that display a diverse range of sizes (<20 KDa upto >400 KDa) and functions such as adhesins, proteases, cytolysins, toxins and outer membrane proteins. These proteins are synthesized with an N-terminal cleavable signal peptide that directs their export into the periplasm via the Sec machinery. Passenger domains may be cleaved from the transporter domain after secretion or remain attached (Henderson et al., 2004).

Autotransporters have been classified into three categories, monomeric type Va proteins, two-component type Vb adhesins and trimeric type Vc adhesins (Desvaux et al., 2004a). The type Va proteins have a β-domain of ~30KDa size that can form the whole membrane-spanning translocator pore (Desvaux et al., 2004b). The type Vb passenger domain and the β-domain are translated as two separate proteins but retain the same mechanism of transport. The type Vc (referred to as Trimeric Autotransporter Adhesins; TAA) is an interesting group of proteins comprising of prominent bacterial virulence factors that are, to date, only adhesins in function. TAAs have a very small (~10KDa) β-domain and the translocator pore forms only when these proteins homotrimerize, trimerization then spreading to the passenger domains too (Linke et al., 2006).

4.2.4 Structural organization of trimeric autotransporter adhesins

TAAs from animal pathogens such as YadA from *Yersinia enterocolitica* (Fig.4.1), NadA and NalP of *Neisseria meningitides*, BadA from *Bartonella henselae*, Hia (Fig.4.2)and Hsf from *Haemophilus influenza* are very well characterized in terms of structural architectures and helped define TAAs from several other animal and plant pathogens. Initial scanning electron microscopic images revealed that these proteins form lollipop-shaped surface projections on the cell envelope (Hoiczyk et al., 2000). X-ray crystal structures of various regions of these proteins showed that the 'lollipop' comprises of head, neck, stalk that form the passenger domain and the membrane anchor β-domain.

The Passenger Domain:

X-ray crystal structures are available for only two TAA passenger domains, YadA and Hia. YadA forms a trimer of single-stranded left-handed β-helices in a novel β-roll fold (Fig.4.1a). The trimeric interface is held by strong hydrophobic interactions due to the periodically occurring conserved NSVAIGXXS sequence motifs (Fig.4.1b). These motifs or their variants are present in the heads of many TAAs, which suggests a structural (and probably functional) analogy in these regions. Following the head domain is a loop region called neck, which functions as an adaptor between the larger diameter of the head and the thinner stalk. The neck loops criss-cross below the trimers and create a 'safety-pin' like lock (Fig.4.1c) (Nummelin et al., 2004).

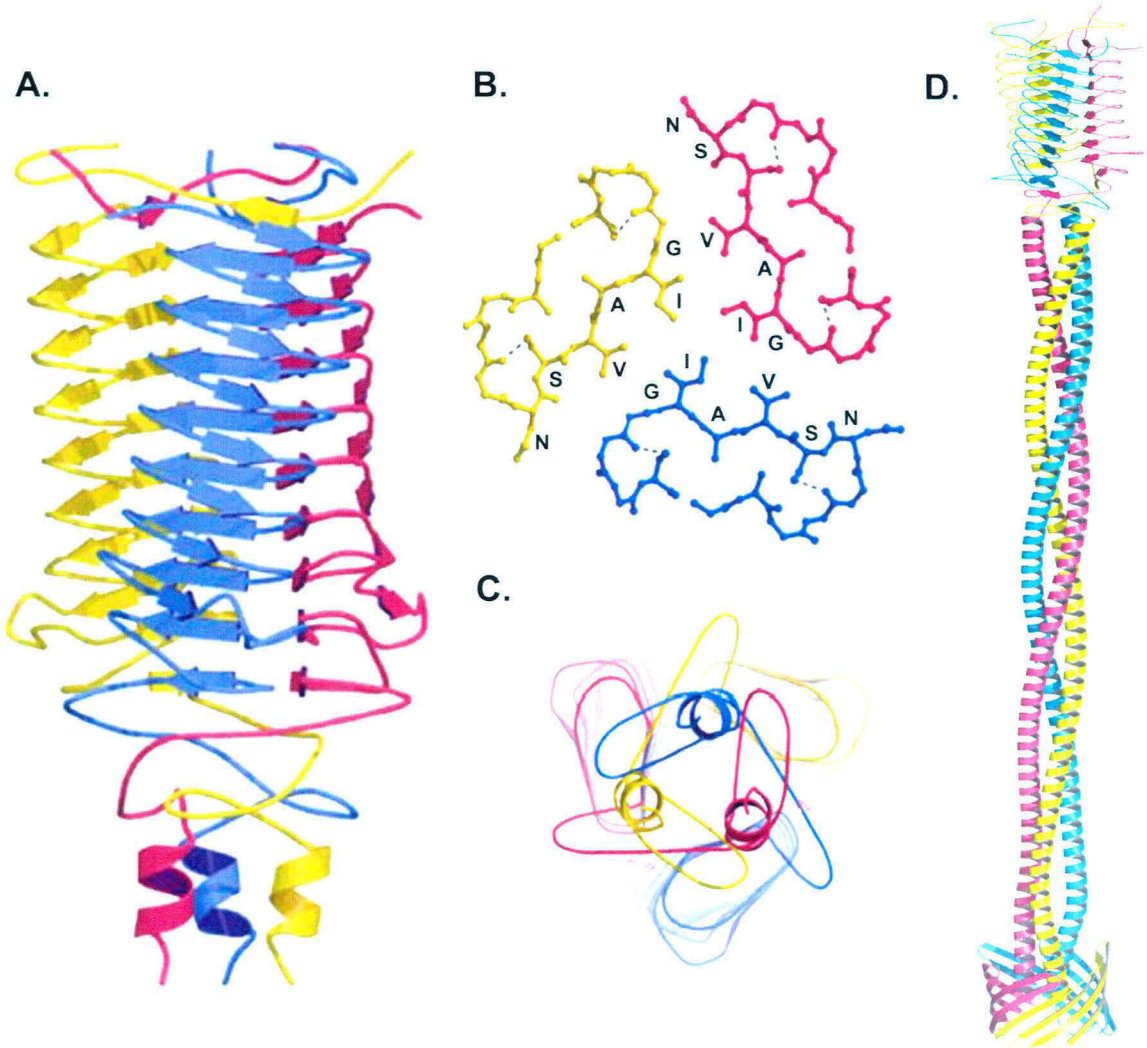


Figure 4.1. *Yersinia* YadA crystal structure. **(A)** The trimeric YadA head domain. The strands are drawn as arrows, helices as ribbons and the different monomers are in different colours. **(B)** One level of the β-roll in the trimer viewed along the z-axis showing the packing of the oligomeric core by large hydrophobic residues and the packing of the monomeric interior by small hydrophobic residues. The conserved 'NSVAIG' residues are marked. **(C)** Organisation of the neck region in the C-terminus of the head domain, viewed from the C-terminus along the z-axis. The safety-pin structures as well as the beginnings of the stalk domain helices are shown. **(D)** A model of the whole YadA trimer; note the long stalk and the transporter domain at the C-terminal. (Adapted from Nummelin et al., 2004; Linke et al., 2006)

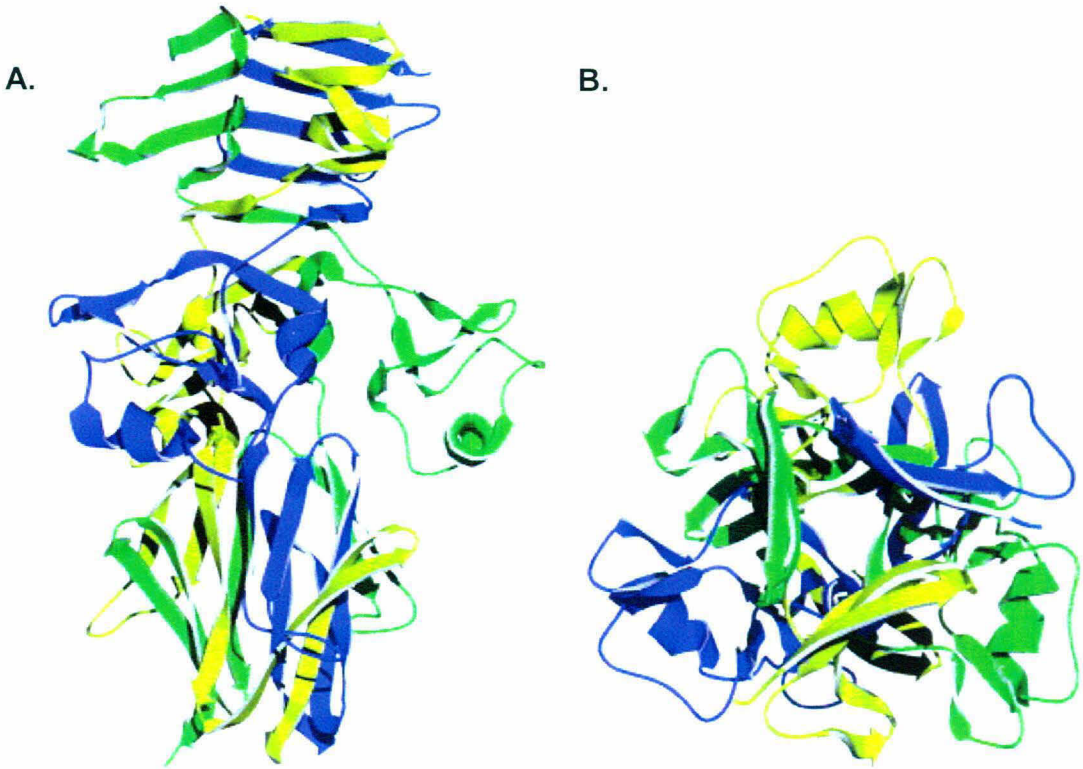


Figure 4.2. *Haemophilus influenza* Hia crystal structure. **(A)** The trimeric Hia binding domain (HiaBD1 with which it binds to host cell receptors) structure. β strands and helices are represented with arrows and thick helices, respectively. **(B)** Organisation of the HiaBD1, as viewed from the N-terminus along the z-axis. (Adapted from Linke et al., 2006)

The Hia head-neck architecture is very different, assuming a novel β-prism fold (Fig.4.2). The proximal part of the head domain is formed by a four-stranded β-meander perpendicular to the fibre axis and the distal part is formed by a 5-stranded β-sheet. The XadA head-neck sequence has a higher homology to YadA, as would be described in detail in the results section and therefore, Hia-like head-neck architecture will not be discussed further.

The stalk domains of TAAs are fibrous, highly repetitive structures that are rich in coiled coils and extremely variable in length (Fig.4.1d). They function as spacers to project the head domains away from the bacterial cell surface and towards its host binding partners. TAAs from animal pathogens bind to the proteins abundant in extracellular matrix of the corresponding hosts, for example, YadA binds to collagen, fibronectin and vitronectin. However, the ligand-binding sites on the head domains are yet to be identified. Several YadA-like proteins have been identified in genome sequences of various pathogens where the head-

neck and stalk domains are generally found in multiple copies within the same protein. The architecture of such proteins, the functional relevance and the binding partners of such proteins is yet to be characterized.

The β-Domain:

The type Va autotransporter NalP β-domain from *Neisseria meningitidis* was the first translocator structure to be solved and revealed a 12-stranded β-barrel with a hydrophilic pore of 10 x 12.5Å that is filled by an α-helix belonging to its passenger domain (Fig.4.3b). This domain and β-domains from several type V-secreted proteins are shown to have translocation activity *in vivo* and *in vitro* (Oomen et al., 2004; Roggenkamp et al., 2003; Surana et al., 2004). The 12-stranded β-barrel structure of EspA (also type Va) transporter domain from *E. coli* revealed that the cleavage of the passenger domain occurs within the pore and that there is a post-cleavage conformational change that stabilizes the β-barrel (Barnard et al., 2007). However, the mechanism and location of the cleavage of the passenger domains may not be common since the pore inner-face electrostatic charges vary from negative in EspA to positive in NalP. The only TAA translocator domain structure to be solved is the Hia β-domain from *Haemophilus influenza* (Fig.4.3a; Meng et al., 2006). In this case the domain is only 70 residues long as compared to the 250-300 amino acid-long domain of type Va proteins. As expected, the trimerization of the domain forms the whole pore, with each monomer contributing four out of the 12 strands constituting the β-barrel.

The sequence features of XadA described in earlier studies from the lab indicated that XadA belongs to the TAA class of adhesins. However, the presence of multiple repeat sequences in XadA and their relevance was not clear. Therefore, to understand the sequence features of XadA and obtain domain information for crystallization and structure solution, an in depth study of the XadA protein sequence was undertaken. Apart from the domain organization, an interesting hypothesis towards the diversity generation using the varying domains present in XadA sequences across the genus Xanthomonas has been put forth from this bioinformatics study.

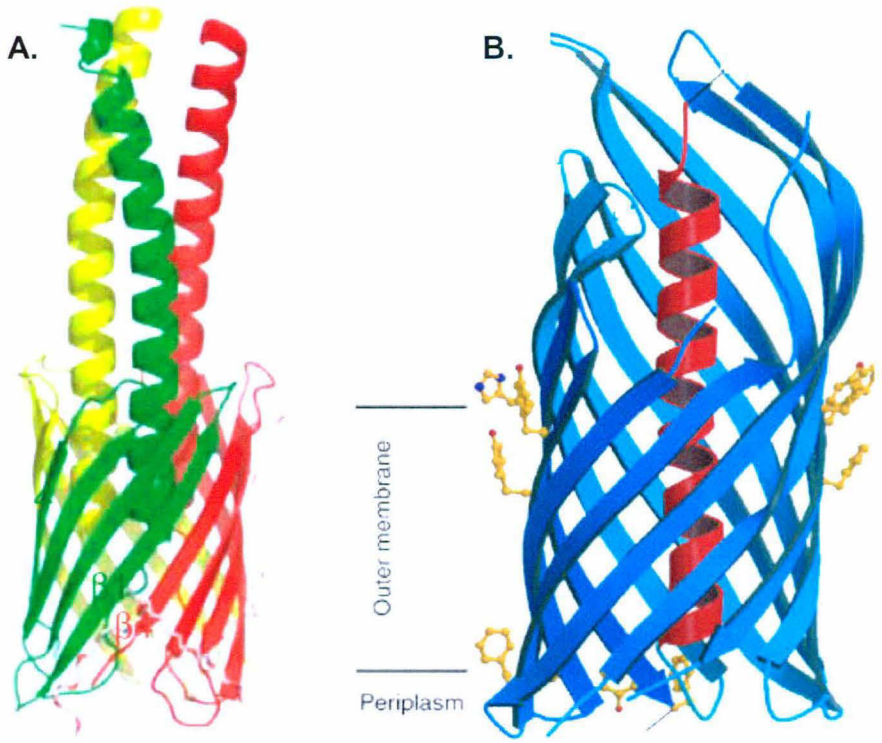


Figure 4.3. Structures of β-domains of autotransporter adhesins. (**A**) The trimeric β-domain of Type Vc *Haemophilus influenza* adhesin Hia. The three monomers are coloured red, green and yellow. (Adapted from Meng et al., 2006) **(B)** The monomeric β-domain of Type Va *Neisseria meningitides* adhesin NalP. (Adapted from Oomen et al., 2004)

4.3 Methods

4.3.1 XadA protein sequence analysis

The XadA sequence from BXO43 (lab strain of Xoo) is deposited in the National Center for Biotechnology Information (NCBI) database (Gene id. AF288222; Derived protein seq. id. AAG01335). This sequence was submitted to for whole protein BLAST searches (Altschul et al., 1997; http://blast.ncbi.nlm.nih.gov/Blast.cgi), PFAM server for identification of the XadA domains (Finn et al., 2008; http://pfam.sanger.ac.uk/) and CLUSTALW server for multiple sequence alignment of the domains of XadA with YadA and other autotransporter sequences (Larkin et al., 2007; http://www.ebi.ac.uk/Tools/clustalw2/). The sequence alignment of XadA domains with YadA was manually adjusted in the alignment editor software Jalview (Clamp et al., 2004). ALSCRIPT was used for formatting the alignments (Barton, 1993).

4.3.2 Manual alignments

A manual YadA structure-based alignment of the domains of XadA was performed to obtain more reliable alignments and secondary structure information about XadA. These features were not evident from the method described in section 4.3.1.

4.3.3 XadA homology modelling

The alignments obtained from the two methods described above were submitted to several homology-modelling servers that generate protein models based on the known structure and alignment of the query protein with the sequence of the protein with the known structure, YadA (PDB id: 1P9H), in this case. MODELLERv9 (Sali & Blundell, 1993), 3D-JigSaw (http://bmm.cancerresearchuk.org/~3djigsaw/), Swiss-Model (Schwede et al., 2003) and I-TASSER (Zhang, 2008) were used for this purpose. Structures of the homologous proteins were retrieved from Protein Data Bank (PDB; http://www.rcsb.org/pdb/). DALI server, freely available software in for structural alignments, was used for structural homology searches (Sali & Blundell, 1993). Structural models were generated as monomers for XadA in five different parts predicted to be domains. The transformation matrix for conversion of YadA monomer into trimer as provided in the PDB file on the RCSB server was used to generate XadA trimers for each monomer. The final model of the whole XadA trimerized molecule is a linear manual assembly of the different domain models.

Electrostatic surface potential of the individual trimers of each domain of XadA was analyzed using the software GRASP (Nicholls et al., 1991).

4.3.4 XadA sequence comparison across genus Xanthomonas

XadA sequences (protein and gene) from various members of genus Xanthomonas were retrieved from NCBI server (Xoo: MAFF 311018, KACC10331; Xcc str. ATCC 33913; Xcv str. 85-10; Xac str. 306). Sequences of the recently sequenced Xca and Xoo str. PXO99^{A} were made available under a collaborative Xanthomonas genomes annotation project (courtesy Dr. Adam Bogdanove). XadA ORFs were annotated in the Xca (cabbage pathogen *Xanthomonas campestris* pv. armoraceae) and Xoo str. PXO99^{A} genomes. XadA sequences from the seven genomes mentioned were submitted to the PFAM server to assess the number of YadA-like domains in each. Sequences from the same species of Xanthomonas having different number

of domains were analyzed further. Artemis Comparison Tool (ACT) was used to compare the whole coding regions and to find out whether there are any internal rearrangements within the corresponding genes (Carver et al., 2005).

4.4 Results and Discussion

4.4.1 XadA is a TAA with four distinct head-neck domains

A simple protein BLAST search with the whole XadA sequence identifies XadA-like proteins from other Xanthomonas members and closely related bacteria like *Xylella fastidiosa*, *Stenotrophomonas maltophila* and Burkholderia sp. or putative hemagglutinin-like proteins from other Gram-negative pathogens like Shigella sp. and *E. coli*. Homology to autotransporters from other bacteria like *H. influenza* and Neisseria sp. is mainly to the C-terminal transporter domain. The only solved structure with which XadA shows significant sequence homology is that of YadA from *Yersinia enterocolitica*. The result of a BLAST alignment with YadA is shown in Fig.4.4a. The peculiarity in this alignment is that the YadA homology is repeated five times in the XadA sequence, indicating the presence of repeats in XadA. Several 'TAVG' and 'TDAVNVAQL' sequence derivatives were identified in the XadA sequence as also reported by Ray et al., 2002. Interestingly, four out of five YadA homology regions were found to be centred around the 'TDAVNVAQL' stretches that constitutes a part of the 'safety pin'-like structure of the YadA 'neck' region just below the 'head' (Fig. 4.1c). Moreover, XadA homology to *Haemophilus influenza* adhesin Hia is also at similar stretches. The 'TAVG' repeats were found to be recurring after about every ten residues and resembled the hydrophobic patches that form the core of the head trimer in YadA. The fifth region of YadA homology is to the C-terminal transporter end of XadA sequence. The presence of four 'head-necks' and a transporter region in XadA was confirmed by PFAM analysis (Fig. 4.4b & c). Despite the identification of four 'head-neck' domains, a careful observation of the alignments indicated that several regions were misaligned and many 'TAVG' derivatives were not aligned with the corresponding YadA regions of homology. The reasons for this misalignment could be that the lengths of the XadA head-neck domains were much different from the single YadA head-neck domain and that many repeats in the XadA head region were longer than the YadA repeats.

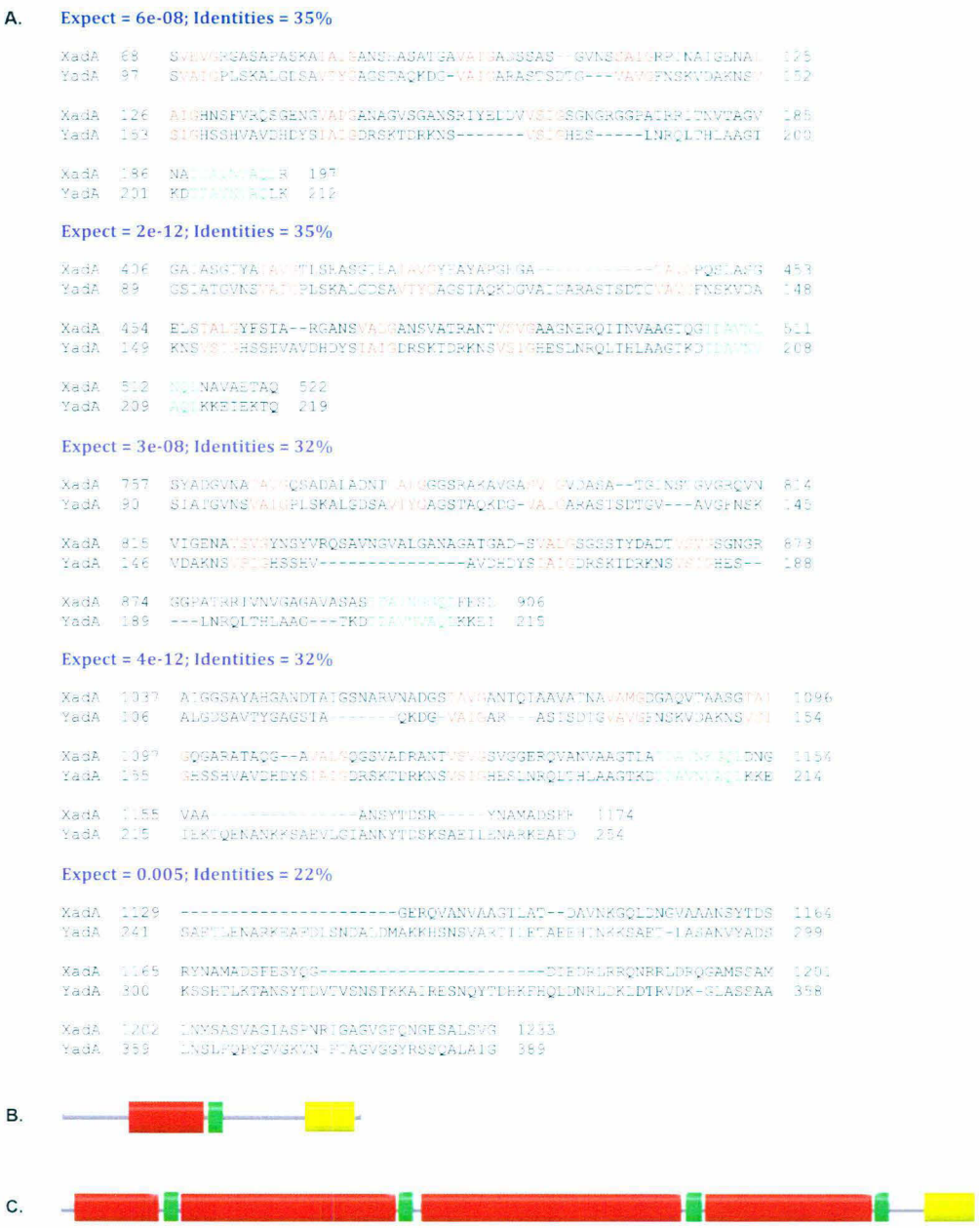


Figure 4.4. Preliminary sequence analysis of XadA. (A) The result of an NCBI-BLASTP alignment of XadA with YadA. Five discernible regions of homology and their respective E-values and percentage identities are mentioned. Note the presence of the highlighted highly homologous stretches; the ones marked in red correspond to the hydrophobic repeats found in the 'head' domains while the green correspond to the 'neck' domains. **(B & C)** PFAM analysis of the YadA (B) and XadA (C) sequences showing the presence of one or more 'head' domains (red), 'neck' domains (green) and autotransporter regions (yellow).

Therefore, a manual alignment of the whole XadA region was performed. The basis of this effort was that the YadA head is composed of a 14-residue repeat motif, of the form:

Turn (T1) ---Three-residue inner strand (IS) ---Turn (T2) ---Three-residue outer strand (OS)

The 'NSVAIG' face is the inner strand and hydrophobic in nature. The outer strands of YadA head are rich in charged residues. The variation in the repeat length is from 13 to 16 residues, is caused by a variation in the T1 length. The other turn, T2, is always the same length. The start of every T2 is marked by an invariant Gly since the conformation of T2 does not allow other residues at this position. In the equivalent position in T1, Gly is not conserved because the conformation is that of a normal β-strand.

YadA also has Ser residues at the fourth position from the invariant Gly to stabilise turns by hydrogen bonding to Gly carbonyl oxygen. In a tight β-roll, there is no space within the core of the head monomers for large residues, and therefore there is a strong preference for small ones, especially Ser in the turns and Ala in the middle of the strands. Conversely, the inter-trimeric core is formed by the Val and Ile (Nummelin et al., 2004).

The residues constituting the T1, IS, T2 and OS were identified for the four head domains of XadA manually. In Fig.4.5a, the residues marked in green represent the XadA inner β-strand IS and the residues in purple are the outer β-strand OS (Fig. 4.5). The invariant Gly is also seen throughout the four heads, repeating after every 13 residues and is marked in red. This alignment was found to be a better indicator of the extent of similarity between YadA and XadA head-neck domains since it revealed the essential structural features of XadA that neither PFAM analysis nor other alignments could identify. With this alignment, it can be safely concluded that XadA head domains have a left-handed parallel β-roll structure.

A.

>XadA Signal peptide
MNQIYRKVWNKSLGVWAVASELS
SGDSPGAVASASFIDRRHRLALT
>XadA Head-neck1
AAIA LALG GAG FATPLPA
NAQS VEVG RGA SAP
ASKA TAIG ANS HAS
ATGA VATG ADS SAS
GVNS SAIG RPT NAI
GENA LAIG HNS FVRQS
GENG VAPG ANA GVSGANSRTY
EDDV VSIG SGN GRGGPATRR
ITNVTAGVNATDAVNVAQL
>XadA Head-neck2
RHVADVAENTAQFFKASPGEESVGAYVE
GDSA LAAG EGA NAV
GTAT TALG TGA NAV
AENA TAVG TNA LAS
GQNS AAFG HNA QAN
GPAS VAVG GAA VNEDGEPLITNGGVPVTTGATSA
GVGG TAVG ASA KADGF
AASS FGVG AYA AGAQASAFGAVANAA
GDYA TAVG TQT RAS
GTSS TAVG GPV DLIPGLGLFVQTQAS
GEAS TALG AGA IAS
GTYA TAVG TLS EAS
GTEA TAVG YFA YAP
GEGA TAVG PQS LAS
GELS TALG YFS TAR
GANS VALG ANS VAT
RANT VSVG AAG NERQ
ITNVAAGTQGTDAVNLNQL
>XadA Head-neck3
NAVAETAQTTGKYFKASGSAKKDVGAYVE
GENA LAAG EGA KAV
GTGT TALG AGA HAV
VRKA TAVG VGA VAD
GIGA AALG NTA RAL
GDNS SAVG SNA VAS
DIGA TANG AGA QAL
SAYT IALG SKA VAS
DLQA IAEG FNS TAS
NVGS TALG GFS QSS
GRLS SALG YSA VAS
SVDS TAVG VAA QAT
GVSA VAIG EIS KAT
GEES VAVG GGA FSGWIPTQASGKGAAAFGAGAWAT
ADYT TAIG RDS YAD
GVNA TAVG QSA DAL
ADNT LALG GGS RAK
AVGA SVIG VDA SAT
GINS TGVG RQV NVI
GENA VSVG YNS YVRQS
AVNG VALG ANA GAT
GADS VALG SGS STY
DADT VSVG SGN GRGGPATRR
>XadA Stalk-like region
IVNVGAGAVASASTDAINGGQLF
ESLSNAASFLGGGAAIGAQGVFV
APTYLIQGASYNNVGAALTALDS
KVTELDARSGGTPANTAARTASL
RTATVPAMAATAVSAVSSNVAS
>XadA Head-neck4
TAID ATAG VQG TPTAAVVGSITP
AAIS TVVG TAA VAN
NVTG TAIG GSA YAH
GAND TAIG SNA RVN
ADGS TAVG ANT QIAAV
ATNA VAMG DGA QVT
AASG TAIG QGA RAT
AQGA VALG QGS VAD
RANT VSVG SVG GER
QVANVAAGTLATDAVNKGQLDNG
VAAANSYTDSRYNAMADSFESY
>XadA Membrane anchor
QGDIEDRLRRQNRRLDRQGAMSS
AMLNMSASVAGIASPNRIGAGVG
FQNGESALSVGYQRAISPRATVT
VGGALSSGDSSIGVGAGFGW

B.

>YadA Signal peptide
MTKDFKISVSAALISALFSSPYAF
AEEPEDGNDGIPRLSAVQISPNVD
PKLGVGLYPAKPILRQENPKLPPR
GPQGPEKKRARLAEAIQPQVLGGL
DARAK
>YadA Head-neck
GIHS IAIG ATA EAA
KPAA VAVG AGS IAT
GVNS VAIG PLS KAL
GDSA VTYG ASS TA
QKDG VAIG ARA SA
SDTG VAVG FNS KVD
AQNS VAIG HSS HVAAD
HGYS IAIG DLS KTD
RENS VSIG HES LNR
>YadA Stalk
QLTHLAAGTKDNDAVNVAQLKK
EMAETLENARKETLAQSNDVLD
AAKKHSNSVARTTLETAEEHAN
KKSAEALVSAKVYADSNSSHTL
KTANSYTDVTVSSSTKKAISES
NQYTDHK
>YadA Membrane anchor
FSQLDNRLDKLDKRVDKGLASS
AALNSLFQPYGVGKVNFTAGVG
GYRSSQALAIGSGYRVNESVAL
KAGVAYAGSSNVMYNASFNIEW

Figure 4.5. Manual YadA-structure-based-alignment of XadA. XadA **(A)** and YadA **(B)** sequences were spaced based on the 'Turn (T1)-Three-residue inner strand (IS)-Turn (T2)-Three-residue outer strand (OS)' arrangement seen in YadA structure. Residues marked in green are the IS hydrophobic amino acids, the invariable Gly are marked in red and the polar residues of OS are highlighted in blue.

Although four head-neck domains are present, only one equivalent of the YadA coiled-coil stalk region was found in XadA. The lack of stalk domains after the first two head-neck domains is interesting and may indicate a plant ligand-specific adaptation. Several long stretches of XadA head repeats were found not conforming to the YadA structure. These were predicted as loop regions.

XadA head-neck domains were individually aligned to the YadA head-neck domain and submitted to various comparative homology-modelling servers to obtain structural models for XadA. However, since the sequence identities of these regions with YadA are very low (Fig.4.4), the models obtained were of very low scores (by the standards of individual programs) with large aberrations in the secondary structures and deviations from the predicted YadA-like left-handed parallel β-roll. In order to generate better and more reliable models, manual alignments were done for short stretches of XadA sequence with YadA using the software Jalview and modelled using MODELLER v9. Since the sequence identities of individual domains with YadA increased to ~55% compared to only ~30% before the manual alignment (Fig.4.4), the resultant models are expected to be more reliable, resembled the YadA structure closely and superimposed well with YadA (Fig.4.6). The r.m.s. deviations of each of the XadA domain models were in the range of 1.0-1.5 upon superposition with YadA. As expected, the 'TAVG' hydrophobic patches formed the inner strand and the inner strands, in turn, formed the hydrophobic core upon trimer generation for each domain monomer. The predicted loop regions were found to constitute the T1 turns only and none was found to be present in the T2 turns. Some large loops have a propensity to form helices.

A working model of the whole XadA molecule was generated by organizing the trimeric form of each head-neck domain of XadA in a linear manner (Fig.4.7). The T1 regions of the head which were observed to loop out or form helices in the monomer models were found to be present away from the core β-roll, causing no obstructions in the trimerization and also, making XadA less compact than YadA. These loops and other secondary structures may have a role in mediating specific motif interactions with the plant ligand. It could be observed from the whole molecule model that the four head-neck domains can provide large ligand-binding areas and may help Xoo in binding to multiple ligands in the rice hydathodes at the same time, considering that XadA is involved in the very early entry of Xoo through rice hydathodes. The length of XadA molecule would allow it to cross the thick layer of exopolysaccharide surrounding the Xoo cells.

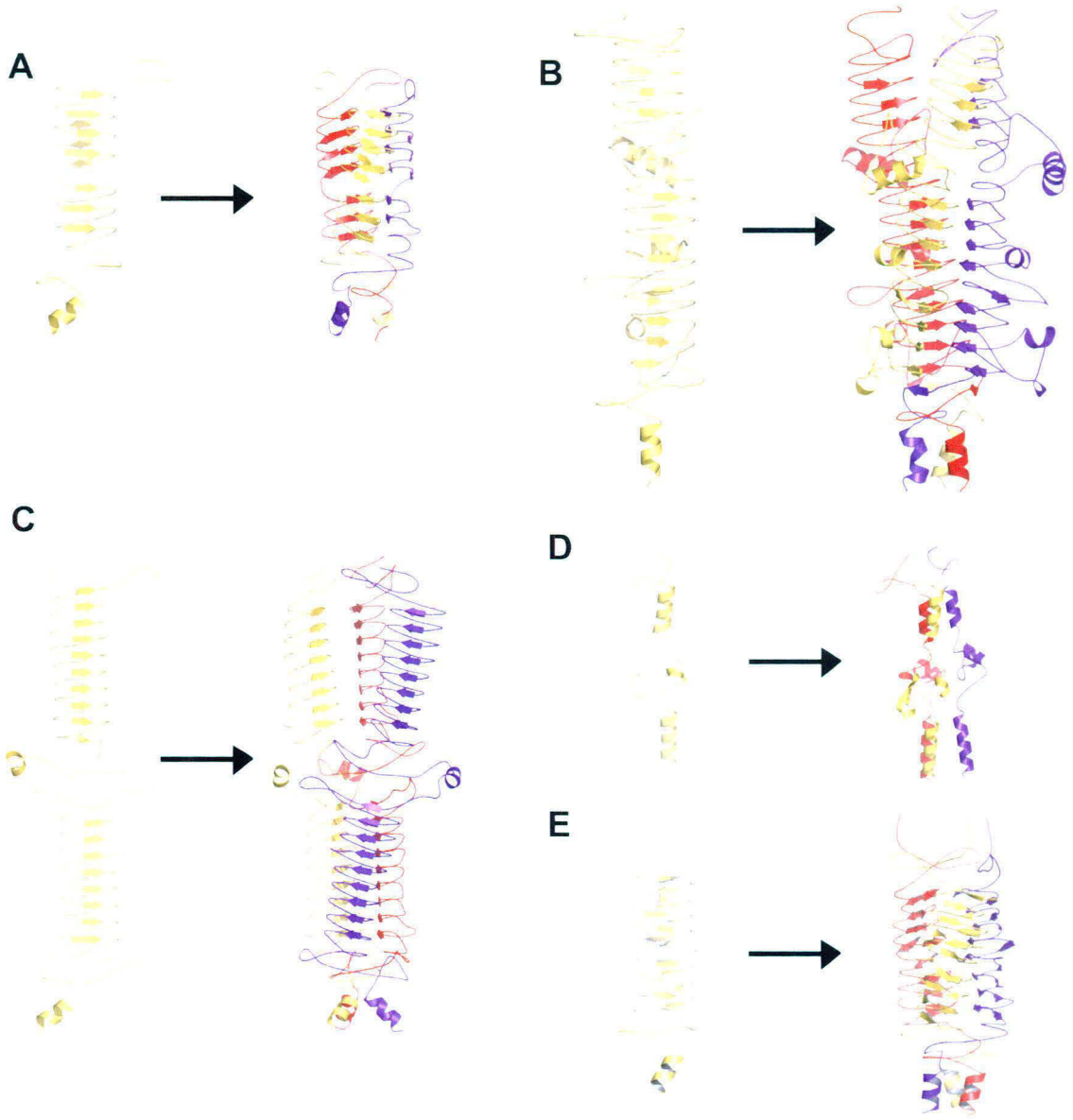


Figure 4.6. The homology models of the four head-neck domains of XadA. Models for **(A)** Head-neck 1, **(B)** Head-neck 2, **(C)** Head-neck 3, **(D)** a coiled-coil stretch between Neck 3 and Head 4 and **(E)** Head-neck 4 regions were generated as monomers by individually aligning the respective sequences to YadA. The individual monomers of each domain are shown in panels left of each arrow; Three monomers of each domain were generated based on the translation and rotation matrix of the three YadA monomers and used to create trimers (shown in the panels right of each arrow) with no Cα clashes. Note that the long loops between the β rolls in these domains might be artefacts of modelling in the regions of lower sequence identity. While each individual monomer is coloured in yellow, the three monomers of a trimer are coloured yellow, red and purple. β-strands are shown as arrows in the cartoon representation.

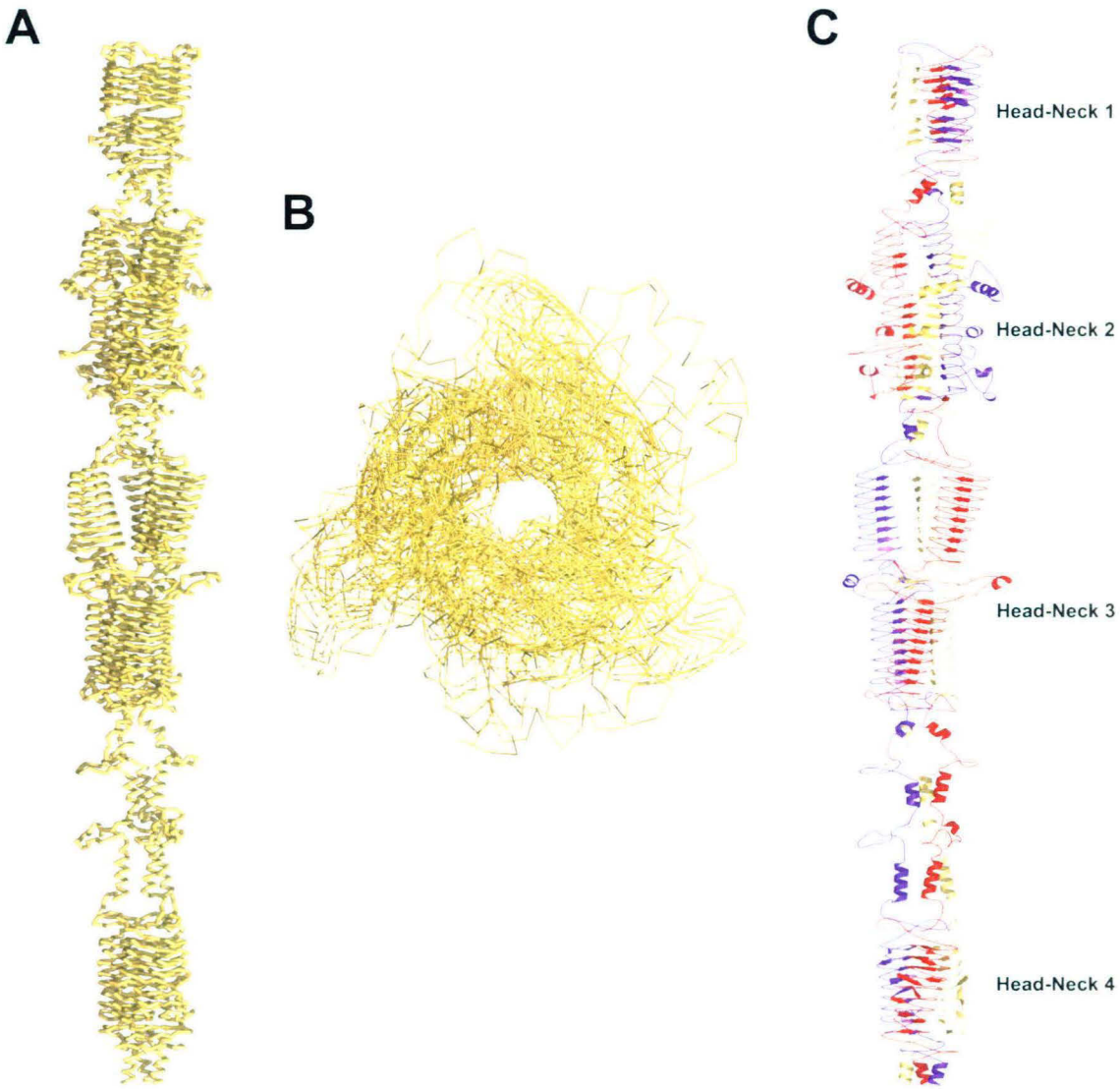


Figure 4.7. A homology model of the XadA passenger domain. This representation is the most probable structure of the XadA passenger as it might appear after being transported out of the Xoo cell outer membrane in a homotrimeric form. The trimers of individual head-neck domains shown in Fig.4.6 were joined to obtain this model. (A) Ribbon representation of the side-view; (B) Ribbon representation of the top-view (Note that XadA might also form a central pore similar to the YadA passenger domain; (C) Cartoon representation of the side-view. The three monomers of a trimer are coloured yellow, red and purple. β-strands are shown as arrows in the cartoon representation.

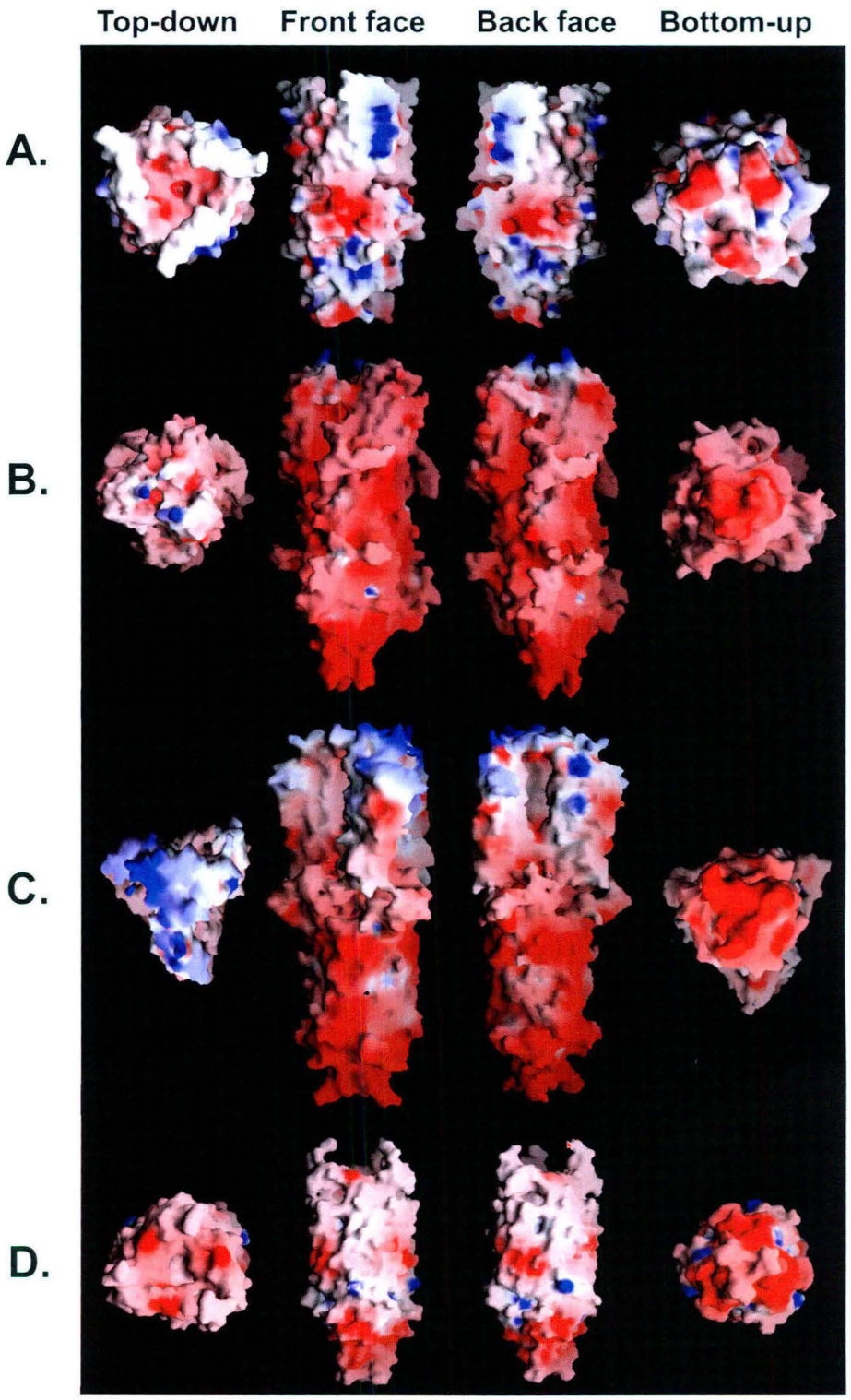


Figure 4.8. Electrostatic surface potential of the four trimeric domains of XadA. Surface potential of (A) Head-neck domain 1; (B) Head-neck domain 2; (C) Head-neck domain 3; (D) Head-neck domain 4; represented as Top-down view, front back faces and the bottom-up view. Red indicates negative charge and blue indicates positive charge.

Mapping the electrostatic surface potential of each trimeric domain of XadA revealed that while the second, third and fourth domains are highly negative in charge, the first domain appears to be positively charged (Fig.4.8). This might be an indicator of the first and the most exposed domain functioning as a ligand 'grabbing' domain or that it may bind to a ligand that is chemically different form the ligand(s) for the rest of the domains. In addition, there are several clefts and pockets in the four domains, which are more negatively charged than the rest of the surface and might presumably be involved in conferring specificity to the ligand binding of XadA. It would be interesting to study the effects of deletion of these highly charged stretches on XadA function in Xoo.

4.4.2 Why should Xoo need a XadA protein?

The collagen binding ability of YadA is well-established (Tamm et al., 1993; Roggenkamp et al., 1996). Collagen is a hydroxyproline-rich protein with a triple helical structure forming long rigid fibrils in the extracellular matrix of higher animals. The central assembly of YadA head seems to bind collagen and a protein docking analysis was used to speculate that the most probable site of binding is a long groove formed by two adjacent T2 turns in the YadA trimer (Nummelin et al., 2004). Trimerization and the 14-mer head repeats of YadA have been shown experimentally to be essential for collagen binding (El Tahir et al., 2000). Proteins containing hydroxyproline are important structural components of plant cell walls. These proteins are termed hydroxyproline-rich glycoproteins (HRGPs) and fall into four groups: extensins, proline-rich proteins (PRPs), 4-hydroxyproline-rich lectins and arabinogalactan proteins (AGPs). HRGPs are often covalently cross-linked into large meshworks, providing tensile strength for the plant cell walls (Showalter, 1993; Cassab, 1988). HRGPs, AGPs specifically, are highly glycosylated molecules involved in cell-cell interactions and might be putative binding partners for XadA-like proteins. Some of the XadA loops might be involved in recognizing HRGP surface glycosylated moieties.

4.4.3 Diversity in XadA proteins from various Xanthomonads

As mentioned previously, XadA orthologs have been reported from the genomes of all the sequenced Xanthomonas species such as Xoo, *Xanthomonas axonopodis* pv. citrii (Xac), *Xanthomonas campestris* pv. vesicatria (Xcv), *Xanthomonas campestris* pv. campestris (Xcc) and *Xylella fastidiosa* (Simpson et al., 2000; da Silva et al., 2002; Lee et al., 2005; Ochiai et

05). XadA sequences from these genomes and the newly sequenced *Xanthomonas campestris* pv. armoraceae (Xca), *Xanthomonas oryzae* pv. oryzicola (Xoc) and Xoo PXO99^{A} genomes were retrieved and analyzed. The three strains of Xoo viz., KACC10331, MAFF311018 and PXO99A were found to be identical and hence, were treated as one protein.

Multiple sequence alignment of the sequences revealed that the XadA proteins have lesser sequence homology amongst each other than expected from the highly homologous genomes of Xanthomonads. For example, although Xoo and Xoc genomes are considered about 90% identical, XadA proteins from the two are 70% identical. Interestingly, the variations in these sequences were found to be in the form of large stretches of sequences in the middle of the proteins and not as differences in residues, contributing to large differences in the lengths of XadA orthologs. PFAM analysis was also performed to assess XadA architecture among these bacteria.

The primary difference between the orthologs of XadA lies in the number of head-neck domains that they have and this leads to the different lengths of these proteins. For example, the Xoc XadA is 2117 aa long while Xoo XadA is 1265 aa long. The size difference is due to the deletion of two of the head-neck domains in Xoo XadA, possibly due to a recombination event involving sequences that are in/near the neck regions (Fig.4.9a). The putative site of recombination in Xoc beyond which the Xoc XadA is identical to Xoo XadA could also be traced (data not shown). Size differences between XadA proteins can also occur due to expansion of 14 aa repeats that comprise the head regions as observed between XadA of Xcc and Xca. Artemis comparison (Fig.4.9b) reveals that one region of the protein is very dissimilar between XadA orthologs of Xca and Xoc. Interestingly, this is the region (~2500bp or 850 aa long) that is present in Xoc XadA but absent from Xoo XadA.

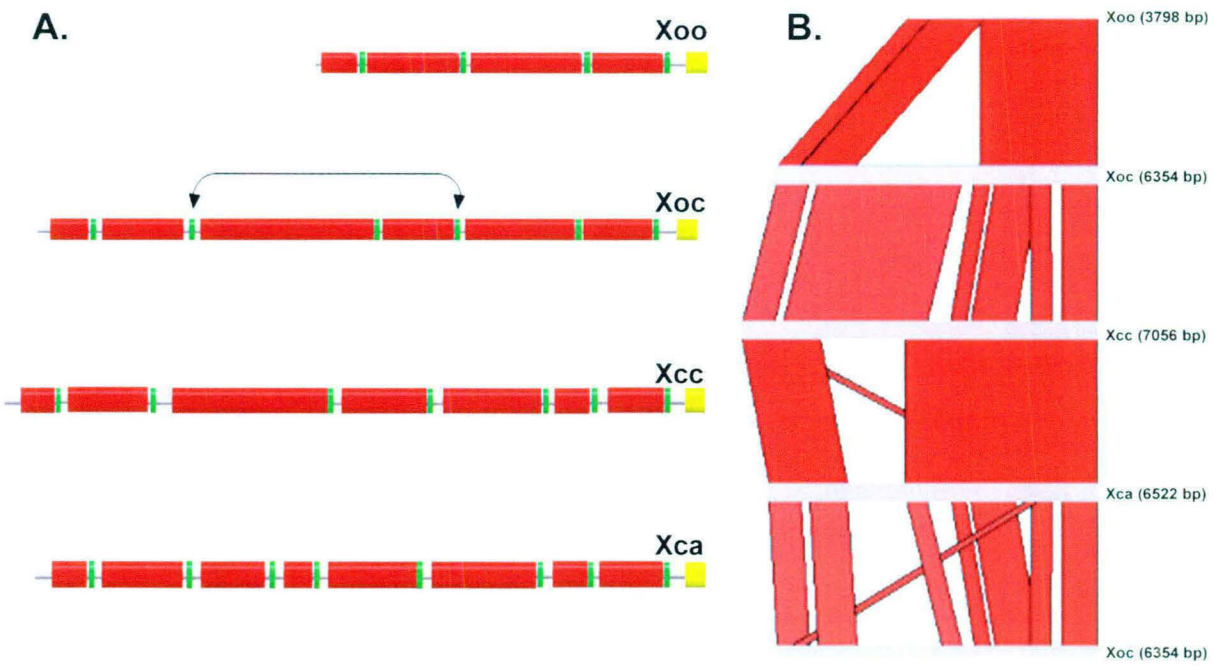


Figure 4.9. Comparison of XadA across Xanthomonad species. A. PFAM analysis of XadA of different Xanthomonads. The red blocks represent the 'Head' structural motifs of Yersinia adhesin YadA consisting of several 14 amino acid repeats. The green blocks indicate the 'Neck' motifs of YadA and the yellow blocks mark the C-terminal domain of YadA. The region marked by arrows indicates the part of Xoc XadA deleted in Xoo XadA. **B**. *xadA* nucleotide sequence comparisons using Artemis Comparison Tool (ACT). The red blocks indicate stretches of identical/similar sequences and white blocks show stretches of deletions/dissimilarities. *xadA* from each genome has been marked as a grey block.

Generation of molecular diversity is a common feature found in TAAs of several animal pathogens, achieved by phase variation, repeat expansion-contraction and even loss-reacquisition of whole TAA molecules (Hoiczyk et al., 2000; Martin et al., 2003; Lafontaine et al., 2001). Especially interesting is the observed variations in the total size of XadA-like Bartonella adhesin BadA due to variations in the length of the repetitive neck/stalk sequences; the sizes of the head and membrane anchor sequences remaining equal in all *Bartonella henselae* strains investigated (Reiss et al., 2007). In animal pathogens, escape from detection by host immune surveillance is a major virulence determinant and hence the variability in surface molecules, which are antigenic and easily detected. A plausible reason for diversity in XadA proteins from genus Xanthomonas might not be to evade plant immune system but to arbitrate host or tissue specificity. Deletion of two head-neck domains of XadA from the

vascular pathogen Xoo XadA that enters rice through hydathodes as compared to the mesophyllic pathogen Xoc that enters rice leaves through stomata also indicates a role for XadA in tissue specificity. Similarly, the vascular pathogen Xcc and mesophyllic pathogen Xca infect cabbage plants and exhibit large variation in the size of one of the head repeats as the major difference in XadA proteins. Apart from the expansion and deletion in both the cases, the rest of XadA and the flanking regions are 96-100% identical. However, such large variations in sizes were not observed upon comparison of XadA proteins from Xanthomonas pathogens, which infect different hosts i.e. Xoc and Xcc (Fig.4.9b). This finding suggests that these pathogens retain the more conserved ancestral versions and that the variations in the head-neck domains have led to evolution of varied XadA proteins to avoid competition with closely related species infecting the same host.

4.4.4 A bigger picture: Other afimbrial adhesins of Xanthomonads

Xoo, Xoc, Xac and Xcv have XadA and its paralog XadB, while Xcc and Xca encode only for XadA and lack XadB. XadB was also predicted to have XadA-like architecture, only shorter by one head-neck domain. It was found to be 97% identical between Xoo and Xoc while only 60% identical between Xoo and Xac/Xcv and again, 97% identical between Xac and Xcv. YapH, a member of type Va monomeric autotransporter family was also found to be similarly variable among Xanthomonads. A 38,766 bp genomic region containing genes *fhaB*, *fhaX, fhaB1* and *fhaC* that may encode predicted non-fimbrial adhesins of the two-component Type Vb system was found to be completely missing in the two Xoo strains BXO43, KACC10331 and MAFF311018 while present in Xoo strain PXO99^{A} and all the other Xanthomonas genomes (data not shown; Salzberg et al., 2008). The locus also includes several direct repeats of ISXo5 transposon elements, which in turn, are flanked by genes for a dual specificity phosphatase (DSP) and a DNA binding protein (DBP). In contrast, only one copy of the ISXo5 element is present between DSP and DBP in MAFF and KACC strains, indicating that the IsXo5 element was involved in the genomic rearrangement that led either to loss of the locus from BXO43, MAFF and KACC or gain of the locus in PXO99^{A}, the former being a likely case (data not shown; Salzberg et al., 2008). Together, this *in silico* analysis reveals that the orthologs of all the afimbrial autotransporter adhesins vary among the Xanthomonas genomes. These differences could have a major role in defining the host and/or tissue specificity among these pathogens.

4.5 Directions from the study

The analysis of Xoo XadA and other afimbrial autotransporter adhesins revealed several unique sequence features, distinct from other characterized adhesin structures and strongly advocated the need for detailed structural studies. Examination of XadA sequence convincingly identified four YadA-like head-neck domains with left-handed β-roll architecture and essential cues for trimer formation with a hydrophobic core. However, a possibility of XadA trimer taking up a more globular structure with the four domains interacting with each other to form a bulbous quaternary structure cannot be ruled out. It can be argued that a globular structure would form a very compact molecule that would be unable to cross the exopolysaccharide layer of Xoo cells and an elongated molecule with four distinct binding sites seems more appropriate as a structure for a surface adhesin. The homology models for XadA could be generated only in four parts and never as a complete molecule since overall identity and architectural similarity of XadA to the only characterized structural homolog YadA is very less. The manually generated working model of XadA (Fig.4.7) fits into the functional role of XadA well. Nevertheless, alternative quaternary structures for XadA could not be ruled out with this study and hence, X-ray crystal structure is necessary. As would be discussed in chapter 5, XadA was taken up for X-ray crystal structure analysis and the distinctive features identified in this chapter were used for cloning this membrane-anchored large protein.

Chapter Five

Expression studies on Xanthomonas adhesin-like protein (XadA) from *Xanthomonas oryzae* pv. oryzae

CHAPTER FIVE

5.1 Abstract

Growth medium-dependent differential expression of the outer membrane-located XadA protein of Xoo has been reported, the protein expression being induced by a minimal medium but not by the peptone-sucrose (PS) medium used to culture Xoo. In the present study, the differential XadA expression was also found in a specialized minimal medium called XOM2 that is considered to mimic growth conditions inside the plant. A careful deletion of each media component of XOM2 revealed that XadA expression is glutamate dependent. The addition of glutamate to PS medium induces XadA expression within one hour. Surprisingly, *xadA* transcripts were found to be present under all nutritional conditions, indicating the involvement of a post-transcriptional control for XadA. Consistent with this possibility, functional deletions of the master transcriptional regulators of Xoo protein expression *in planta* were found to have no effect on glutamate-dependency of XadA expression. Primer extension studies show that the XadA 5' UTR is 186 bp long and present under all growth conditions. Together, the results suggest that rapid expression of XadA is achieved by keeping the transcripts ready under all conditions but expressing XadA only after sensing the glutamate that is present in exudates from the hydathodal pores of the rice leaves.

5.2 Introduction

Dynamic environmental conditions like temperature, wind and rain regulate the conditions on the exposed aerial surfaces of the plant hosts. The key factor that attracts several microorganisms and promotes epiphytic growth on such hostile plant surfaces is the presence of organic nutrients that exude out from the plant sap (Andrews & Harris, 2000). Bacteria can sense these small molecules as signals to communicate with the plant hosts to regulate cellular responses of the whole pathosystem (Dunn & Handelsman, 2002). Foliar plant pathogens are more proficient in forming larger colonies on the leaf surfaces, even under environmental stress and tend to colonize openings such as stomata, hydathodes or wounds better than the non-pathogens (Wilson et al., 1999). Signals in the form of sugars, amino acids or some aromatic compounds have been found in the plant exudates and are considered as important chemoattractants for the bacteria to detect their hosts (reviewed in Andrews & Harris, 2000).

The internal milieu of the plant host is more conducive to microbial growth, being abundant in sugars, minerals and water (Marschner, 1995; Raven, 1984). *In planta* growth leads to the induction of bacterial genes that have virulence-related functions (Stachel et al., 1985; Osbourn et al., 1987; Beaulieu & van Gijsegem, 1990; Marco et al., 2003). In *Pseudomonas fluorescens*, some of these genes were shown to respond to small molecules like sugars, amino acids and organic acids that are presumably present in the wheat root exudates (van Overbeek & van Elsas, 1995).

Plant-inducible genes facilitate *in planta* growth of the pathogen through the expression of efficient mineral and sugar uptake systems (Expert et al., 1996; Blanvillain et al., 2007). The other essential functions associated with these plant-responsive genes are to encode toxins or hormones that interfere with the normal plant metabolism or development, in addition to proteins that can disrupt host-signalling pathways (van Gijsegem, 1997). The *hrp* (for hypersensitivity reaction and pathogenicity) gene cluster is a plant-inducible set of genes responsible for the causation of disease on host plants and eliciting a hypersensitive response (HR) on the non-host plants (reviewed in Lindgren, 1997). This cluster codes for genes involved in suppressing the innate immune responses that arise due to the detection of the pathogenic invasion by the host plant and promoting expression of host susceptibility factors (Mudgett & Staskawicz, 1998; Yang et al., 2006; reviewed in Jha & Sonti, 2009). In plant pathogens, the *hrp* genes encode a Type III secretion system (T3S) that forms a protein export

apparatus called the injectisome for delivery of T3S-secreted proteins directly into the plant cells (reviewed in Cornelius, 2006). Interestingly, expression of *hrp* genes is observed not only *in planta* but also during growth in synthetic minimal media, where the level of expression depends on the nature of the carbon source provided (Arlat et al., 1992; Rahme et al., 1992; Wei et al., 1992). These media conditions are thought to mimic physiological conditions encountered by bacteria during *in planta* growth.

A major mechanism of bacterial 'response regulation' is the two-component sensory transduction system (Stock et al., 2000). HrpG, is the two-component transcriptional activator responsible for the expression of *hrp* genes *in planta* as well as in the synthetic plant mimic media (Brito et al., 1999). HrpX is an AraC-type sensor kinase that functions in response to activation by HrpG and in turn, activates the *hrp* effector operons (Wengelnik & Bonas, 1996). It is believed that HrpG can assimilate signals from various sensor genes that can detect plant milieu. One such sensor is PrhA that can sense physical contact with plant cells (Brito et al., 1999). Although no other sensors have been reported that may relay plant signals to HrpG, it is understood that there are contact-independent metabolic signal sensors functioning upstream of HrpG signalling pathway.

Xanthomonas adhesin-like factor A (XadA) is a surface-anchored protein, which belongs to the trimeric autotransporter adhesins (TAA) family, as inferred from the results of Chapter 4. An important observation regarding XadA expression in Xoo was that this protein could be detected only in a minimal medium and not in the peptone-sucrose (PS) medium routinely used to culture Xoo (Ray et al., 2001). This finding suggests that XadA is a differentially expressed protein responding to some nutrients in its growth medium. This result indicates a milieu-dependent variability at the level of XadA protein expression. In this chapter, a detailed analysis of the factors governing XadA expression is described. This study indicates that the expression of XadA is glutamate-dependent, post-transcriptionally regulated and might be essential for Xoo entry into the rice hydathodes.

5.3 Materials and Methods

5.3.1. Growth media and strains used in the study

Strain/Plasmid	Characteristic Features	Reference/Source
BXO43	rif-2;Rf^r derivative of BXO1	Laboratory collection
BXO836	*xadA*1::Tn5 *gus*A40 rif-2 ; *xadA*$^-$ strain derived from BXO43	Ray et al., 2001
BXO846	*zxx*-110::Tn10 rif-2 ; *xadA*$^+$ recombinants obtained from BXO836	Ray et al., 2001
DH5α	F , *end* A1 *hsdR*17 (rk$^-$ mk$^+$) *supE*44 *thi*-1 *recA*1 *gyr*A *rel*A1 f80d*lac*ZDM15D (*lac*ZYA-*arg*F) U169	Laboratory collection
pK18Mob	Km^r	
BXO2301	*hrpG*::pK18mob rif2 ; *hrpG*$^-$ strain derived from BXO43	This work
BXO2302	*hrpX*::pK18mob rif2 ; *hrpX*$^-$ strain derived from BXO43	This work
BXO2303	*hfq*::pK18mob rif2 ; *hfq*$^-$ strain derived from BXO43	This work
BXO1844	*colS*::pK18mob rif2 ; *colS*$^-$ strain derived from BXO43	S. Sujatha, M. R. Vishnupiya, H. Patel, A. Pandey and R. Sonti
BXO1843	*colR*::pK18mob rif2 ; *colR*$^-$ strain derived from BXO43	

Table 5.1. List of strains & plasmids used plasmids used in the structural and functional analysis of XadA

Escherichia coli strains were grown in Luria-Bertani medium (Miller 1992) at 37°C and Xoo cultures were grown at 28°C in Peptone-Sucrose medium containing 10gm peptone and 10gm sucrose per litre (Ray et al. 2002). For induction of XadA, plant mimic medium XOM2 containing 1.8gm D (+) xylose (0.18%), 0.1gm L (-) Methionine (670 μM), 1.852gm Na-glutamate (10mM), 12ml of 20mM Fe-EDTA, 0.0068gm $MnSO_4$ (40 μM), 2.0gm KH_2PO_4 (14.7mM) and 1.0185gm $MgCl_2$ (5mM) were dissolved per litre. The pH was adjusted to 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0 and 7.5 using 1N NaOH. 20mM Fe-EDTA solution was prepared by mixing 2.5gm $FeSO_4.7H_2O$ and 3.36 gm Na-EDTA in 450ml water, boiling and cooling with constant stirring. Basal XOM2 medium (XB) was prepared by adding 1.8gm D (+) xylose (0.18%), 1.852gm Na-glutamate (10mM), and 2.0gm KH_2PO_4 (14.7mM) per litre and adjusting the pH to 5.5. The concentration of other amino acids used to replace Na-glutamate in the XB medium was also 10mM. Antibiotics rifampicin (50 μg ml^{-1}), kanamycin (15μg ml^{-1} for Xoo and 50 μg ml^{-1} for E. coli), cephalexin (10 μg ml^{-1}l) and cycloheximide (100 μg ml^{-1}) were used in this study (Sigma-Aldrich, USA). Cyclohexamide was used to minimize fungal contamination. The Xoo strains used in the study are listed in Table 5.1.

5.3.2. Cloning of XadA in various domain combinations into *E. coli*

The *in silico* analysis of XadA (Chapter 4) indicated that this protein has four distinct head-neck domains. With the aim of overexpression and crystallization, XadA was cloned in different combinations of the head-necks, apart from the full-length passenger domain of XadA. XadA was cloned with an N-terminal 6x-His-tag using pET28b expression vector under an isopropyl thiogalactoside (IPTG)-inducible T7 promoter (Novagen, USA). The various constructs of XadA used in this work are elaborated in Fig. 5.1.

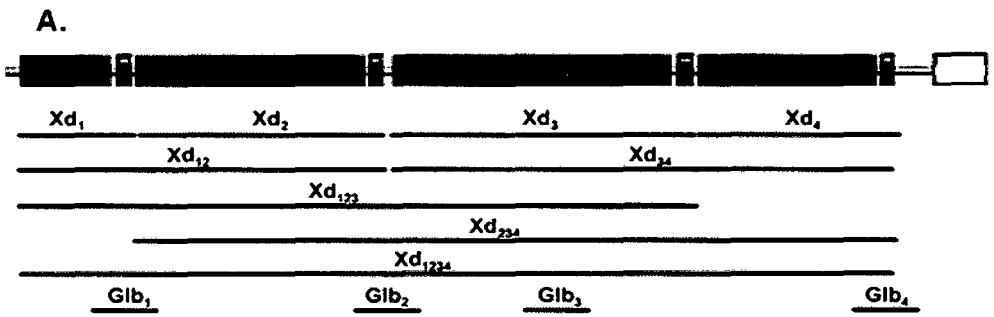
The N-terminal 66-amino acid predicted signal peptide and the C-terminal membrane anchor domain were not included in any of the constructs. PCR primers for each construct were designed according to the *xadA* sequence of Xoo strain BXO43 (Genbank accession no. 288222) with NdeI and XbaI restriction endonuclease sites. The PCR products were obtained using regular PCR protocol using high fidelity Phusion polymerase (Finnzymes, USA), cloned into pET28b vector and transformed into BL21DE3 *E. coli* strain. DNA sequencing was performed to confirm the sequence of each construct.

5.3.3. Expression and purification of the XadA constructs in *E. coli*

A meticulous scheme was followed to overexpress the various XadA constructs, initially in small 10ml culture volumes and then, scaled up. Each clone was assessed for maximum expression under varying temperatures, time of induction and concentration of IPTG used. The different temperatures and time combinations used were 37°C for 4 hr, 30°C for 6hr, 25°C for 10hr, 18°C for 24hr and 15°C for 30hr while the IPTG concentrations were varied from the standard 1mM to 0.1mM. The induced cells were harvested, sonicated at 24W for 3min and centrifuged. The supernatants were analysed on SDS-PAGE gels.

The constructs Xd_{12} and Xd_4 were purified using native affinity purification protocol. Lysis buffer containing 100mM Tris-Cl pH 7.5, 300mM NaCl with 1mM PMSF was used to solubilise the induced cells harvested from one to eight litres of cultures and sonicated at 24W until complete lysis of the cells was visible (30-40min). The lysate was centrifuged at 18,000rpm for 45min, allowed to bind to the Ni-NTA^{2+} affinity column (Qiagen, USA) and washed thoroughly with wash buffer containing 100mM Tris-Cl pH 7.5, 300mM NaCl. Protein of interest was eluted using 50mM, 100mM and 250mM imidazole (step elution) in 100mM Tris-Cl pH 7.5, 300mM NaCl. The protein peaks obtained during the

chromatography run were assessed using SDS-PAGE and fractions containing pure XadA proteins were pooled. The pooled fractions were concentrated using Amicon-Ultra 15 concentration units (Millipore, USA) and loaded on a Superdex-75 gel filtration column. In this step, the proteins were buffer-exchanged with 20mM Tris-Cl pH 7.5 and 20mM NaCl. The fractions containing pure protein were pooled and concentrated to obtain 1-3mg ml^{-1} protein.



B.

XadA construct	Amino acid span
Xd_1	67-224
Xd_2	197-518
Xd_3	515-905
Xd_4	903-1154
Xd_{12}	67-518
Xd_{34}	515-1154
Xd_{123}	67-905
Xd_{234}	197-1154
Xd_{1234}	67-1154
Glb_1	177-271
Glb_2	447-598
Glb_3	734-845
Glb_4	1049-1160

Figure 5.1. Schematic representation of the various constructs of XadA. A. The PFAM annotation of XadA representing the four 'head-neck' motifs is shown in cartoon form. The lines represent the various fragments of XadA that were cloned as 6x-His-tagged proteins. (Note: the various fragments are not drawn to scale). **B.** The region of XadA present in each XadA construct used in this study.

For XadA constructs that expressed largely in insoluble form, 6M Guanidine HCl was used to solubilize the inclusion bodies completely and the proteins were allowed to bind to the Ni-NTA^{2+} column. The wash step was also a column refolding step with slow gradient between the buffer containing 8M Urea, 100mM Tris-Cl pH 7.5 and 300mM NaCl and another buffer containing 2M Urea, 100mM Tris-Cl pH 7.5, 300mM NaCl. The proteins were eluted very slowly with 250mM imidazole in the wash buffer containing 1M Urea, 100mM Tris-Cl pH 7.5, 300mM NaCl. Fractions containing pure proteins were pooled and subjected to slow dialysis at 4°C with buffer containing 1M urea and finally, 100mM Tris-Cl pH 7.5, 300mM NaCl and 1mM PMSF. Absence of precipitation at this step was considered as an indication of complete refolding. The proteins were treated like the native proteins after this step and subjected to gel filtration chromatography and SDS-PAGE analysis. The final concentrations of the proteins were 1-3mg ml^{-1}.

5.3.4 Assessing protease contamination in purified XadA proteins

Casein zymogram analysis was performed to rule out the possibility of protease contamination in the purified XadA preparations. For this analysis, the protein samples were mixed with sample buffer lacking reducing agents like DTT (4X sample buffer: 250mM Tris pH 6.8, 8% SDS, 40% glycerol, 0.4% bromophenol blue). α-Casein was added to a normal 12% SDS-PAGE mix at a concentration of 1mg ml^{-1}. After electrophoresis, gels were incubated for 30 minutes with 2.5% Triton X-100 and subsequently for 16hr at 37°C in 100 mM Tris-Cl, pH 7.4, containing 15 mM $CaCl_2$. Gels were stained with Coomassie Blue R-250 and destained with water.

5.3.5 Protocol used for attempting crystallization of XadA

XadA proteins Xd_{12} and Xd_4 were found to be stable for longer periods than other XadA constructs. Purified Xd_{12} and Xd_4 were screened against 198 crystallization conditions from Hampton Research Crystal Screen I and II, 96 conditions from Hampton Research Index Screen (Hampton research, USA) and 240 conditions from JBS screen (Jena Biosciences, USA). Each screen was monitored for atleast three months for the presence of crystals.

5.3.6 Assessment of polydispersity in purified XadA proteins

The distribution of molecular mass in a protein sample is a good indicator of its quality, wherein the presence of molecules of various radii is a sign of polydispersity or multiple species in multiple conformations and molecular weights in the sample. Only single fractions collected from the protein peaks obtained in the final step of purification, i.e., gel filtration were used for this analysis in order to reduce the chances of introducing multiple conformers in the starting samples. Dyna Pro99 Dynamic Light Scattering (DLS) instrument (Wyatt Technology Corporation, USA) was used to record the intensity distribution plot, which is a measure of the molecular radius vs. Amplitude. Concentration of the proteins used was 0.2 mg ml^{-1}. The samples were thoroughly centrifuged at 16,000 rpm and filtered using 0.2μm Anodisk filters (Millipore, USA). The protein samples were incubated for 5 min at room temperature before data acquisition over an acquired time of 10 min.

5.3.7 *In silico* analysis of disorder in XadA sequence

In order to assess whether XadA sequence was responsible for the disorder observed in these constructs, an *in silico* GlobPlot disorder analysis was performed (http://globplot.embl.de/). This software indicated that large stretches of XadA sequence are disordered and only very small globular regions were present. These globular regions are marked in Fig. 5.1 as 'Glb'. The four globular stretches were cloned and overexpressed as 6x-his-tagged proteins and used for setting up crystallization screens as described above.

5.3.8 Secondary structure analysis of purified globular regions of XadA

The secondary structure elements of the Glb constructs were analysed using Far UV circular dichroism (CD) spectroscopy. These spectra were recorded on a Jasco J-715 spectrophotometer (Jasco Inc., USA) at room temperature. Concentration of the proteins used was 0.2 mg ml^{-1}. 0.1cm path-length quartz cuvette was used to record the far UV spectra over 195-260 nm.

5.3.9. Polyclonal anti-XadA antibody generation

0.5mg of Xd_{12} and Xd_4 were mixed thoroughly with incomplete Freund's adjuvant and used for subdermal injection of two different healthy rabbits. Preimmune sera were collected before the injections. Booster doses of equal amounts of proteins were given to the animals 2 weeks later. After 3 weeks, a small amount of blood was withdrawn from both animals and test Western blotting was performed to assess the antibody efficiency as compared with the preimmune sera. 10ml of serum was finally collected from each animal to yield separate polyclonal antibodies against N- and C-terminal of XadA. The antibodies against Xd_{12} were effective at 1:3000 and Xd_4 at 1:2000 dilution.

5.3.10. Xoo outer membrane protein preparations and Western blotting

Various Xoo strains were grown to saturation (24hr) in PS medium. The cultures were divided into 1ml aliquots and centrifuged at 8,000 rpm for 2min at 4°C. The supernatants were discarded; the pellets were washed with 1ml sterile water three times. The pellets thus obtained were resuspended in various growth media. For induction in the different minimal media, the pellets equivalent of 1ml Xoo cultures were resuspended in 1ml of the requisite medium and grown at 28°C for 36hr. Since Xoo cultures grow very fast in PS as compared to the minimal media, different strategies were followed to grow the cultures in PS. For short-term induction in PS medium (for assessing XadA expression within 1-12 hours), the pellets equivalent of 1ml Xoo cultures were resuspended in 1ml of the requisite medium and grown at 28°C. For long-term induction in PS medium (for assessing XadA expression 12-36 hr after growth in PS), the pellets equivalent of 1ml Xoo cultures were resuspended in 1ml of PS, of which a 10% inoculum was transferred to 1ml PS and grown at 28°C.

After induction, equal number of cells (normalised using O. D. units) were pelleted and resuspended in an outer-membrane protein extraction buffer containing125mM Tris-Cl pH 7.5 and 2% SDS. The samples were heated at 95°C for 3min and centrifuged at 13, 000 rpm for 5 min (method modified from Hussain et al., 2001). The supernatants were treated as outer-membrane protein preparation and used for Western blot analysis. The samples were loaded on 8 % SDS-PAGE gels and electro-transferred using ECL semi-dry transfer unit (Amersham Biosciences, USA) on a Hybond C-extra nitrocellulose membrane (GE, USA). The blots were blocked with 3% bovine serum albumin (BSA) made in TBST buffer (TBST: 50mM Tris pH

7.4, 150mM saline, 0.5% TWEEN-20), followed by treatment with primary anti-Xd_{12} antibody diluted in 1% blocking solution. After three rigorous washes with TBST for 15min each, the blots were treated with secondary goat anti-rabbit IgG alkaline-phosphatase conjugated antibody at a 1:10000 dilution (Sigma, USA), followed by three rigorous washes again. The blots were developed with a developing solution containing 0.05mg ml^{-1} nitroblue tetrazolium (NBT) and 0.05mg ml^{-1} bromo- chloro- indoyl phosphate (BCIP) dissolved in 100mM Tris pH 9.5, 100mM NaCl and 5mM $MgCl_2$. After the appearance of the bands, the blots were washed with water, air-dried and photographed.

5.3.11. Enrichment and partial purification of XadA from Xoo cells

Cells were harvested from a 100ml PS-grown saturated 24hr culture of Xoo BXO43 and transferred to 250ml XOM2 medium. After 36hr of induction, the cells were harvested again and resuspended in 10ml of outer-membrane protein extraction buffer (125mM Tris-Cl pH 7.5, 2% SDS). The samples were heated at 95°C for 3min and centrifuged at 15,000rpm for 30min. The supernatants were treated as outer-membrane protein preparation and subjected to ammonium sulphate precipitation with 5%, 10%, 20%, 30% and 70% cuts. A western analysis was performed to trace XadA in the supernatants obtained after every step of ammonium sulphate precipitation and the fractions containing XadA were pooled. These fractions were dialysed for 36hr with several changes of 100mM Tris-Cl pH 7.5, 100mM NaCl to remove SDS, although it is known that complete removal of SDS from protein samples is very difficult.

5.3.12. RNA isolation and Real time PCR analysis of *xadA* transcripts

Xoo saturated (24hr) cultures grown in PS medium were pelleted and induced in different media conditions for 36hr. After induction, equal number of cells (normalised using O. D. Units) were subjected to total RNA isolation by using Trizol (Invitrogen, USA), following the manufacturer's instructions. RNase-free DNase (Invitrogen, USA) treatment was given to the samples to ensure no genomic DNA contamination. The quality of the isolated RNA was assessed by 0.8% formamide-agarose gel electrophoresis and quantified with a Nanodrop spectrophotometer (NanoDrop technologies, USA).

2μg of freshly isolated RNA were used for first strand cDNA synthesis using Superscript II reverse transcriptase (Invitrogen, USA) and random hexamer primer. One μl of the 1:400 fold

diluted cDNAs were further used for real-time PCR (RT-PCR) analysis. RT-PCR enables both detection and quantification (as absolute number of copies) of a specific sequence in a cDNA sample using intercalating double stranded dyes such as SYBR green fluorescent dye. An increase in product is proportional to the increase in fluorescent intensity and the number of cycles required is measured, resulting in the quantification of the amount of cDNA. Triplicates of each cDNA sample were taken to set up the RT-PCR. 2.5picomole of each of the primers and 2µl of DNA sample were taken and volume of the reaction mix was made up to 17.5µl with water followed by addition of equal amount of SYBR green master mix (Invitrogen, USA) to make up the final volume of 35µl which would be sufficient for 3.5 PCR reactions. As a positive control, genomic DNA amount (Experiment positive control) was diluted to get a threshold cycle (C_T) value equivalent to the cDNA samples. RNA samples with no cDNA synthesis step were used as internal negative controls while 16s rRNA was used as an internal positive control. Average C_T value was calculated for triplicates of each sample. A difference greater than 0.3 in C_T value was rejected as an outlier. For each primer pair:

Step1: $^{AVERAGE}\ C_T\ (cDNA) - {}^{AVERAGE}\ C_T\ (RNA) = \Delta C_T$

Step2: $\Delta C_T^{(Experiment)} - \Delta C_T^{(16s\ rRNA)} = \Delta\Delta C_T$

Step3: Relative amount of transcript in the experimental samples = $2^{-\Delta\Delta CT}$

The values obtained from the final step for each Xoo strain under different nutrient conditions were plotted.

5.3.13. XadA primer extension and transcriptional start site mapping

Primer extension was performed using a reverse primer annealing to the 5' end of the *xadA* transcript (annealing in the region from -1 to -22 of the translational start site). 5-10µg of each RNA sample isolated as described previously was used as the template. The 5'end- ^{32}P labelled primer was allowed to anneal to the transcript in the presence of 10mM Tris pH 7.8, 1mM EDTA, 250mM KCl at 50°C, followed by incubation at 42°C for reverse transcription in the presence of 400µM dNTPs and 5 units of AMV reverse transcriptase. At the end of 30min, the extension products were processed and resolved on a 10% polyacrylamide-8M urea denaturing gel in 1X TBE. A parallel manual sequencing reaction was performed using

Perkin-Elmer sequencing kit with the previously described labelled primer and genomic DNA as the template. The gels were dried and subjected to phosphor imaging analysis. All bands were quantified using the Image Gauge program (Fuji, Japan).

5.3.14. Generation of mutations in *hrpG*, *hrpX*, and *hfq* genes of Xoo

Mutations in *hrpG*, *hrpX*, and *hfq* genes of Xoo strain BXO43 were generated by integration of recombinant pK18mob plasmid through homologous recombination (Fig.5.2). An internal fragment of each gene to be mutated was polymerase chain reaction (PCR) amplified from BXO43 genomic DNA and cloned in pK18Mob plasmid. The plasmid DNA isolated from the positive clones was transformed by electroporation into BXO43 electrocompetent cells and the recombinants were isolated by selection for kanamycin resistance. Gene disruption was confirmed by PCR and sequencing of the PCR products. This process resulted in nonpolar mutations when the transcriptional orientation of the gene of interest and the promoter of *lacZ* gene in pK18mob were in the same direction (Wei et al. 2007; Windgassen et al. 2000).

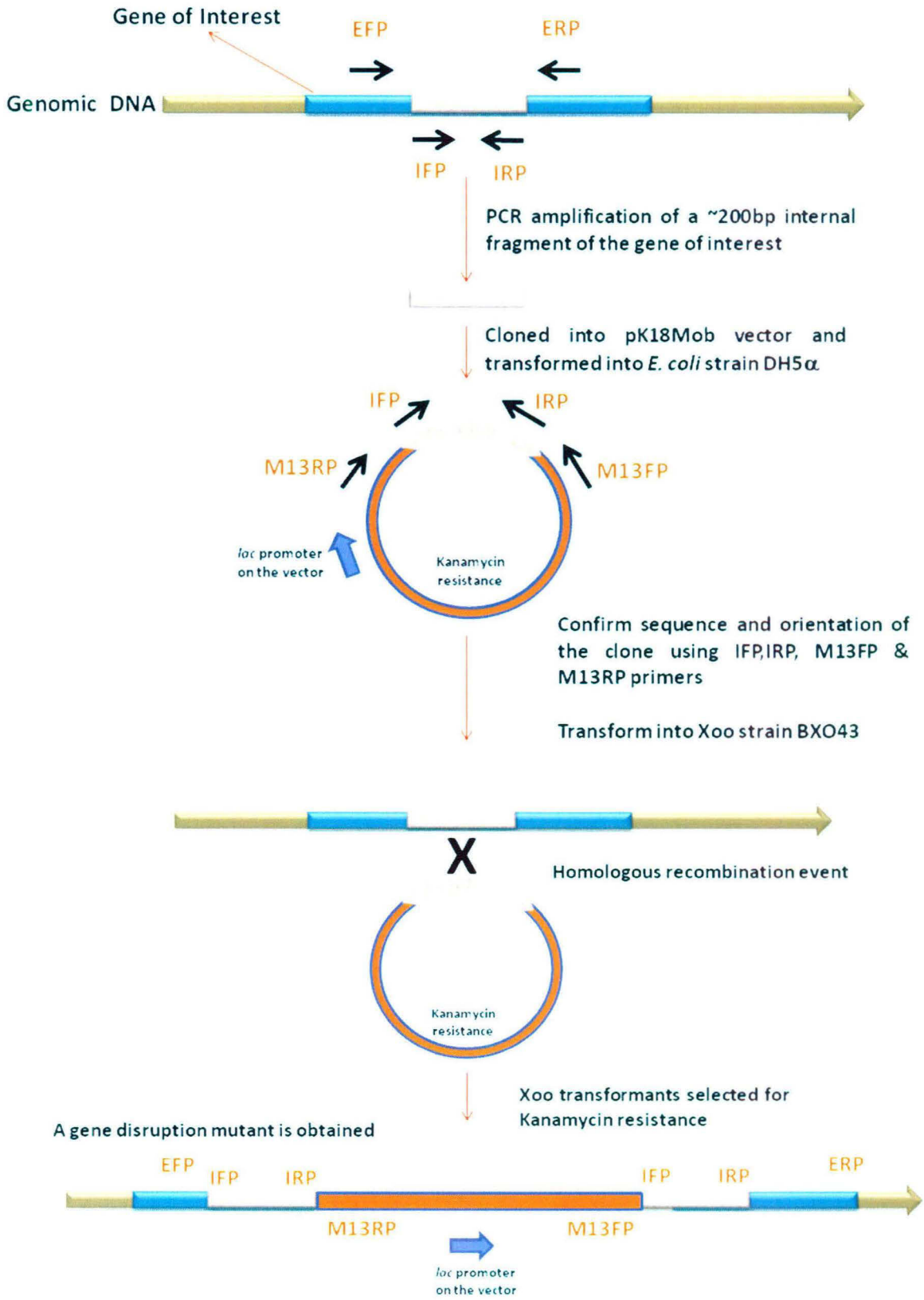


Figure 5.2. Schematic representation of the strategy used for generation of mutations in *hrpG*, *hrpX* and *hfq* genes of Xoo using pK18Mob vector.

Description	Name	Sequence (5'- 3')
HrpX external primers	HrpXFP	5'ATGATCCTTTCGACCTACTTT 3'
	HrpXRP	5'CCGTTGCAAGGTTTCCATCG 3'
HrpX internal primers	HrpXIFP	5'CGAAACGCCCCCCAGCCTGTG 3'
	HrpXIRP	5'GCGCGCCCGGTCACCCCAATCT 3'
HrpG external primers	HrpGFP	5'TTCCTCGTTGCAATGGTAAATAG 3'
	HrpGRP	5'ATGTGCCGGGCTGAGACTTAG 3'
HrpG internal primers	HrpGIFP	5'GACGCTCCCCCGACAACCAACTG 3'
	HrpGIRP	5'GCCCGCGCGATACCAGTCCAGGATGTC 3'
Hfq external primers	HfqFP	5'GCCGCTGCTGCAACTGTTCGACGC 3'
	HfqRP	5'ACGCGCGGTCGGTTATCGGCAAGC 3'
HrpX internal primers	HfqIFP	5'CCACCACCCGGCCCCACGCGCACA 3'
	HfqIRP	5'CGGCGTGAGCGGGTTCCTGTGTCG 3'
Primer for Primer Extension	XPG6R	5'TGCTCTAGACGGCGCACGATTCCTGAGTTGTTG3'
RT- PCR from region 15-96bp of *xadA*	XRT5FP	5'TCGCAAGGTCTGGAACAA 3'
	XRT5RP	5'ACGCAACGGCTCCGGGGCTG 3'
RT- PCR from region 2884-2985bp of *xadA*	XRT3FP	5'CCCGCCAACACCGCAGCACGCA 3'
	XRT3RP	5'CACCGGCCGTTGCATCGATCGC 3'
16sRNA control for RT- PCR	16srRNAFP	5'GGCCTAACACATGCAAGTCG 3'
	16srRNARP	5'AAAGAGTAGATTCCGATGTA 3'
Primers used for cloning the most stable and globular regions of XadA into pET28b	Glb1F	5'CATGCCATGGTCACTGCCGGCGTCAAT 3'
	Glb1R	5'AAGGAAAAAAGCGGCCGCTCATGCGCTGTTCTG CCCGGA 3'
	Glb2F	5' CATGCCATGGCCAGTGGTGAACTGAGC 3'
	Glb2R	5'AAGGAAAAAAGCGGCCGCTCATGCCCGCGCGG TGTTACC 3'
	Glb3F	5' CATGCCATGGGCGCAGCGGCCTTC 3'
	Glb3R	5'AAGGAAAAAAGCGGCCGCTCATGCACCGGCATT GGCACCCAG 3'
	Glb4F	5'CATGCCATGGACACCGCAATCGGCAGC 3'
	Glb4R	5'AAGGAAAAAAGCGGCCGCTCAGCTATTGGCGG CGGCAACGCC 3'

Table 5.2. List of primers used in the structural and functional analysis of XadA

5.4. Results and Discussion

5.4.1. XadA protein overexpressed in *E. coli* is intrinsically unstable

Various constructs of XadA (Fig. 5.1), containing different combinations of the four head-neck domains described in Chapter 4, were cloned as His-tagged proteins and overexpressed in *E. coli*. While constructs Xd_{12} and Xd_4 yielded completely soluble proteins, the rest of the 'Xd' series of constructs showed insoluble proteins that could not be retrieved into soluble form under native conditions (without any denaturing agents). The level of expression of XadA in all the constructs was found to be low. In general, detergents such as Triton X-100, TWEEN-20 and Na-deoxycholate or polyols like glycerol and sorbitol can help in the solubilisation of inclusion bodies and bring more amount of protein into soluble form. However, the total XadA expression in every construct was very less and the effect of each additive on the solubility of XadA was only partial. Therefore, other than Xd_{12} and Xd_4, affinity purification was performed under denaturing conditions for the rest of the constructs.

A common feature that was observed in all XadA constructs was the rapid degradation, despite the addition of protease inhibitors like PMSF and cocktails containing leupeptin, pepstatin and aprotinin even at low temperatures. The stability was found to be less than 24hr in most cases, with the appearance of protein ladder within a couple of hours after purification, resulting in complete degradation (Fig.5.3a). No protease contamination was found in any protein preparation, as indicated by casein zymograms (Fig.5.3b). Therefore, the instability of XadA proteins seems to be intrinsic.

Xd_{12} and Xd_4 were found to be more stable than the rest of XadA proteins and were used for setting up crystallization screens, wherein no crystals were obtained. Dynamic light scattering analysis indicated the presence of multiple species of different oligomeric nature in the peaks showing pure and single protein bands immediately after purification (Fig.5.3c). For crystallization purpose, the protein of interest should have conformational purity and almost no polydispersity or in other words, the protein used for crystallization should be of a single molecular species with no oligomers and degradation products that can hinder crystallization. An *in silico* disorder analysis tool, GlobPlot, indicated the presence of very short globular stretches in XadA while suggesting frequently repeating regions of disorder. The globular stretches (Glb proteins), when cloned as His-tagged proteins, yielded highly stable and well-

expressing proteins (data not shown). However, these stable globular XadA proteins were very short in lengths and found to contain largely coiled coil secondary structures as predicted by the *in silico* analysis in Chapter 4 as well as from the CD spectrophotometric analysis (data not shown), rendering these proteins unfit for X-ray crystallography experiments.

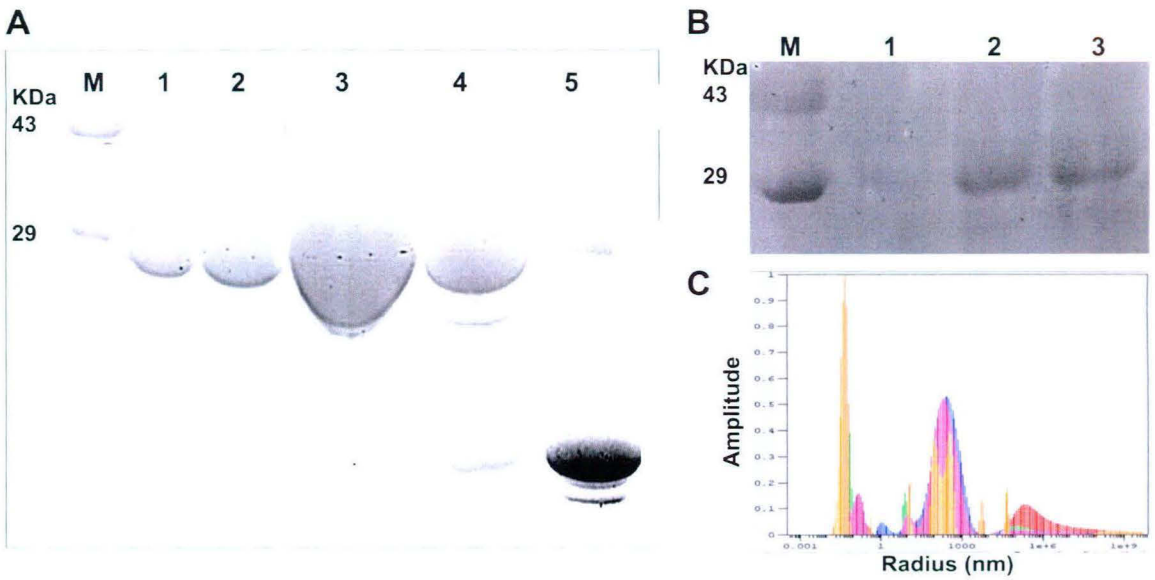


Figure 5.3. Xoo XadA protein shows intrinsic instability when expressed in *E. coli*. **(A)** SDS-PAGE gel photograph showing the degradation of the 6x His-tagged 27 KDa Xd_4 construct of XadA after purification under native conditions. Lane 1: Xd_4 after NiNTA affinity purification; 2: Xd_4 after gel filtration using Superdex75; 3: The purified protein after concentration using a 5KDa-cutoff Amicon Ultra filter; 4: The purified concentrated protein after 24hrs incubation in 50mM Tris-Cl pH 7.5, 150mM NaCl, 1mM EDTA and 1mM PMSF; 5: The purified concentrated protein after 48hrs incubation in the buffer condition mentioned above. Note the appearance of lower bands (found to be degradation products by Western blot analysis using anti-XadA antibodies) and the decrease in the concentration of the 27KDa XadA band in lanes 4 & 5. **(B)** Casein zymogram of purified Xd_4 showing an absence of protease contamination (expected to be seen as zones of clearance in the position equivalent to the 27KDa Xd_4 band) in the protein preparation. Lane 1, 2 and 3 contain 50ng, 500ng and 750ng of a very freshly purified Xd_4 protein. **(C)** The intensity distribution plot obtained by dynamic light scattering analysis of a freshly purified Xd_4 protein. Note the presence of multiple peaks in the plot that represent the proteins of various radii even in the pure fractions of Xd_4 immediately after gel filtration (lane 2 Fig.5.2A), clearly indicating polydispersity.

5.4.2. XadA expression in Xoo is nutrient condition-dependent

It is reported that XadA expression is absent in peptone-sucrose (PS) rich medium that is regularly used for culturing Xoo and that it is induced in modified minimal medium (Ray et al., 2002). It was found that XadA expression is more consistent in a plant-mimic medium called XOM2 (Fig. 5.4a). This medium contains xylose as the sugar source and glutamate, methionine, Fe^{3+}, Mg^{2+} and Mn^{2+} as growth supplements. Tsuge et al. (2002) showed that this synthetic medium is capable of inducing the expression of *hrp* genes that are otherwise expressed only during growth inside the plant host. The secretion of the Hrp proteins in synthetic media by several plant pathogens is reportedly more at acidic pH values and reasoned to be so due to a closer approximation of the pH conditions inside plant apoplast (Furutani et al., 2003; Rossier et al., 1999; Huynh et al., 1989; Gopalan et al., 1996). Taking this into account, XadA expression was tested in XOM2 medium at different pH and was found to be induced at pH 5.0, 5.5, 6.0 and 6.5, with maxima at pH 5.5 (data not shown).

5.4.3. Glutamate regulates XadA expression

In order to find out whether the absence of XadA expression in rich medium is due to some inhibitory activity in PS or because of an inducer in XOM2 medium, the wild-type and *xadA*- strains were grown in a 1:1 mixture of PS and XOM2. Surprisingly, robust XadA expression was observed in the $XadA^+$ strains in this mixture, clearly suggesting that XadA expression is due to an induction effect in XOM2 while its absence in PS is mainly because of the lack of induction and not due to some inhibitory effect (Fig.5.4b).

Therefore, a search for the inducer in XOM2 medium was performed by reconstituting various XOM2 components and deleting each component one at a time. For this purpose, a basal XOM2 derivative (XB) containing 0.18% xylose and 14.7mM KH_2PO_4 (pH 5.5) was made. It was confirmed that XadA does not express in this medium. Xylose was replaced with equal amounts of glucose, sucrose and arabinose and XadA expression was not induced by any sugar in this basal medium (Fig. 5.5a).

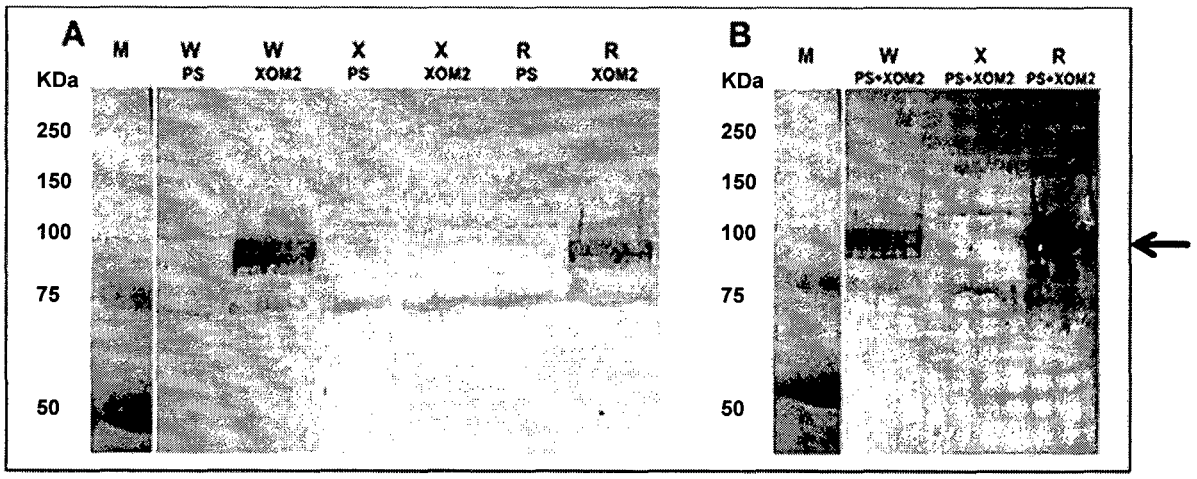


Figure 5.4. Western blot analysis of XadA expression from outer-membrane protein pool of Xoo. **(A)** XadA is expressed in the plant-mimic medium XOM2 but not in the peptone-sucrose (PS) medium. Lanes WPS: Wild-type grown in PS; WXOM2: Wild-type in XOM2 pH 5.5; XPS: *xadA*$^-$ in PS; XXOM2. *xadA*$^-$ in XOM2 pH5.5; RPS: *xadA*$^+$ revertant in PS; RXOM2. *xadA*$^+$ Revertant in XOM2 pH5.5. **(B)** XadA expression is also observed in a mixture of PS and XOM2; Lanes WPS+XOM2: Wild-type in a 1:1 mixture of PS and XOM2 pH5.5; XPS+XOM2: *xadA*$^-$ in 1:1 mixture of PS and XOM2 pH5.5; RPS+XOM2: *xadA*$^+$ revertant in a 1:1 mixture of PS and XOM2 pH5.5. The arrow indicates the expected 110KDa XadA band.

Each XOM2 component was added to XB pH 5.5 in different experiments and assessed for XadA expression by Western analysis. Strikingly, only when glutamate was added to XB pH 5.5, XadA expression could be regained (Fig. 5.5a). This glutamate containing XB medium was named XBGlu. Glutamate, along with other amino acids such as glutamine, asparagine and aspartate, is reported from the vasculature of rice and the guttation fluid of some grasses (Kamachi et al., 1991; Goatley and Lewis, 1966). Guttation fluid is the hydathodal exudate that appears as drops on leaf tips and apparently, serves as a chemoattractant for Xoo (Feng & Guo, 1975). XadA expression was also seen in XBGln medium containing glutamine instead of glutamate, albeit at a lower level, while similar replacement with aspartate and asparagine caused no induction (Fig. 5.5b).

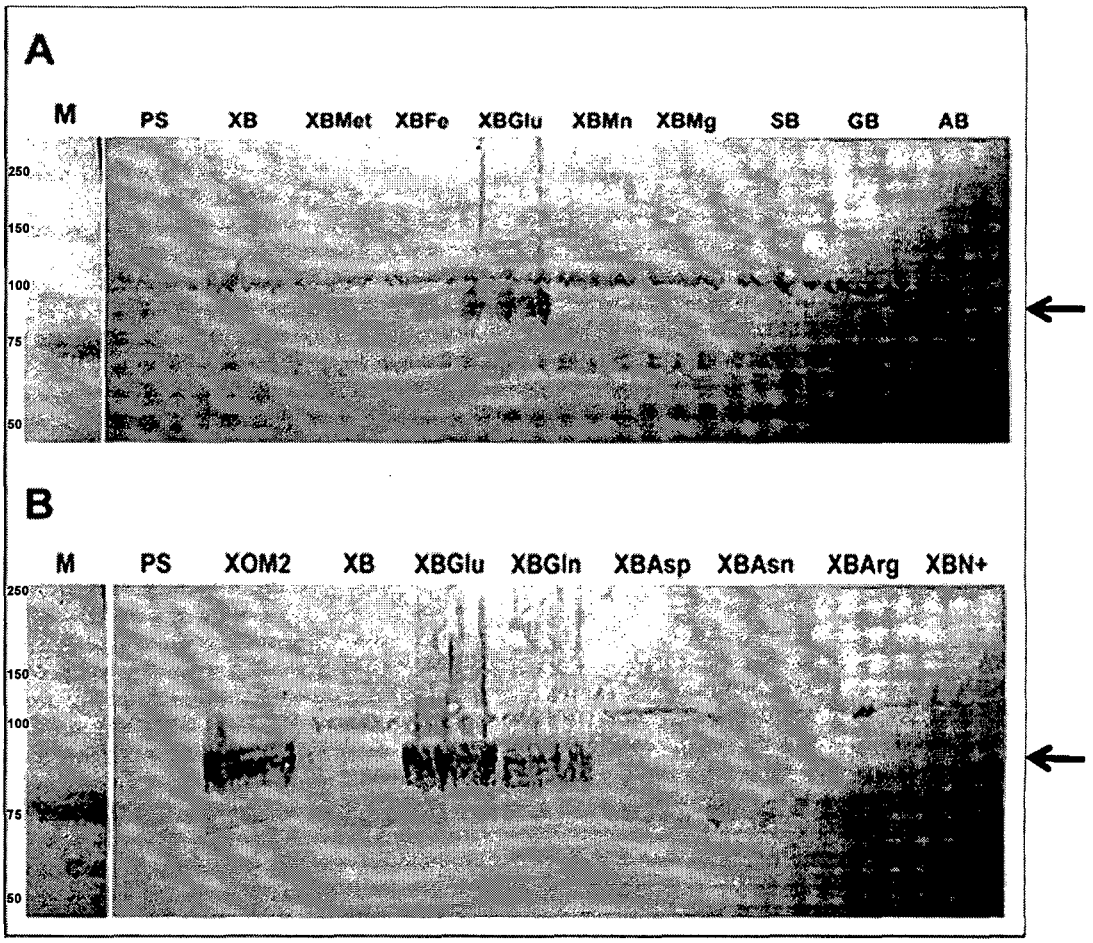


Figure 5.5. XadA expression is glutamate dependent. Western blot analysis of the outer membrane protein preparation of Xoo using anti-XadA antibodies was performed. Xylose and KH_2PO_4 containing medium was treated as the basal medium XB. All the other media components were tested in XB background. **(A)** XadA expression of wild-type Xoo in various media components of XOM2. Lanes XBMet: XB + methionine; XBFe: XB + Fe-EDTA; XBGlu: XB + sodium glutamate; XBMn: XB + $MnSO_4$; XBMg: XB + $MgCl_2$; SB: Sucrose instead of Xylose in XB; GB: Glucose instead of Xylose in XB; AB: Arabinose instead of Xylose in XB. **(B)** XadA expression in BXO43 in the presence of nitrogenous amino acids. Lanes XBGlu: XB + sodium glutamate; XBGln: XB + Glutamine; XBAsp: XB + Aspartate; XBAsn: XB + asparagine; XBArg: XB + arginine; XBN+: XOM2B + NH_4Cl. The arrow indicates the expected 110KDa XadA band.

5.4.4. Glutamate induces a rapid and stable XadA expression

Addition of glutamate to PS medium at pH 5.5 also induced XadA expression (Fig. 5.6a). A time course of Xoo induction in XBGlu and XOM2 revealed that the XadA expression begins within 1hr (Fig. 5.6b). The expression increases upto 24hr, where it remains stable even upto 48-60hr after induction (data not shown). XadA protein was partially purified (section 5.3.11)

from a relatively large (250ml) culture of BXO43 strain in XOM2 medium in order to assess the stability of XadA isolated from native source. XadA protein, thus isolated, was found to be relatively enriched several fold (Fig.5.7) and very stable at 4°C for several months (data not shown). The final XadA preparation might retain traces of SDS, even after thorough dialysis and the presence of detergent micelles might have a role in the long-term stability. However, it is important to note that with this method of purification, XadA can be relatively enriched in a stable form. This method, when scaled up, might be possibly used for obtaining full-length XadA protein for crystallization studies.

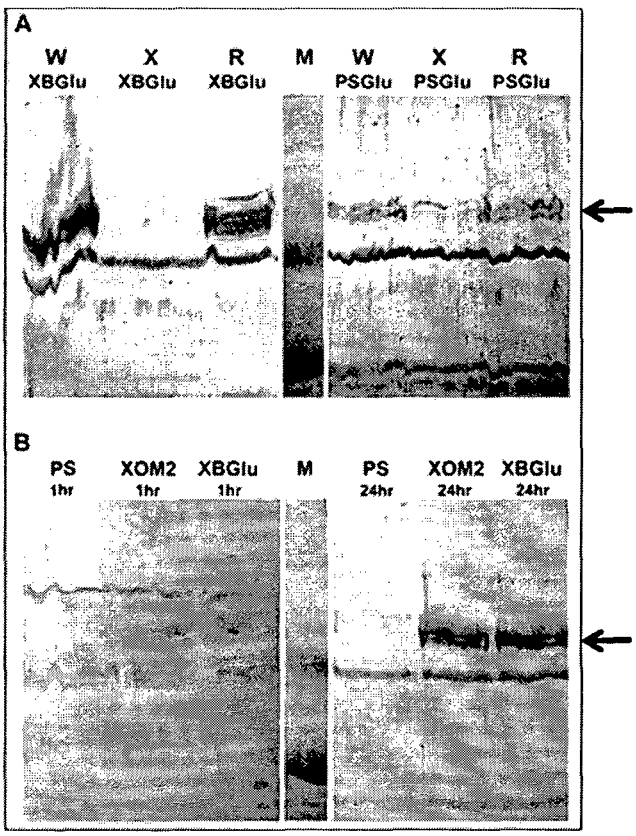


Figure 5.6. Characteristics of Glutamate-dependent XadA expression. Western blot analysis of the outer membrane protein preparation of Xoo using anti-XadA antibodies was performed. **(A)** Addition of glutamate into PS can induce XadA expression. Lanes WXBGlu: Wild-type Xoo grown in XB + Glu; 2. XXBGlu: *xadA*$^-$ Xoo grown in XB + Glu; RXBGlu: *xadA*$^+$ Revertant Xoo grown in XB + Glu; M: Marker; WPSGlu: Wild-type Xoo grown in PS + Glu; 2. XPSGlu: *xadA*$^-$ Xoo grown in PS + Glu; RPSGlu: *xadA*$^+$ Revertant Xoo grown in PS + Glu. **(B)** One hour of induction in the presence of glutamate is enough to induce XadA expression. All treatments in this figure are performed on wild-type Xoo cells grown in PS medium for 24 hours and induced for the requisite time-periods. Lanes PS1hr: Xoo induced in PS for 1hr; XOM21hr: Xoo induced in XOM2 for 1hr; XBGlu1hr: Xoo induced in XBGlu for 1hr; M: marker; PS24hr: Xoo induced in PS for 24hr; XOM224hr: Xoo induced in XOM2 for 24hr; XBGlu24hr: Xoo induced in XBGlu for 24hr. The arrow indicates the expected 110KDa XadA band.

There is also a possibility of post-translational modification of XadA in Xoo in its native form, for example glycosylation, that might be absent when expressed in *E. coli*. Although the event of glycosylation is not very well studied in bacteria, certain Gram negative bacterial autotransporters are reported to be N-glycosylated (reviewed in Benz & Schmidt, 2002). Another plausible explanation for stability of XadA in Xoo could be that the lack of trimerization and/or proper folding in *E. coli* causes instability in the T1 loops that recur after every β-roll structure (Chapter 4; section 4.4.1), which might not be the case in XadA expressed in Xoo. The role of glutamate, if any, in conferring stability to XadA is also an interesting premise. This can be assessed by over-expressing XadA in *E. coli* in the presence of added glutamate.

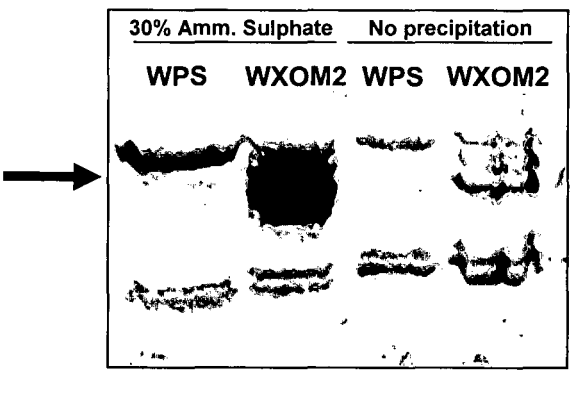


Figure 5.7. Relative enrichment of XadA obtained by the partial purification protocol. Western blot analysis of outer membrane protein preparations of wild type Xoo cells (WPS: PS grown Xoo or WXOM2: XOM2 grown Xoo cells). Step-wise ammonium sulphate precipitation was performed with one set of cultures (lanes marked '30% Amm. Sulphate') while the control set was not precipitated (lanes marked 'No precipitation'). The final volumes of the experiment and control preparations were equalized and finally equal volumes were loaded on the SDS-PAGE gel. The arrow indicates the expected 110KDa XadA band.

5.4.5. XadA expression is independent of the characterized response regulators of Xoo

The results clearly show that XadA is induced in the plant-mimic medium XOM2. Expression of bacterial genes in either plant-mimic medium or in planta is dependent on the HrpG and HrpX proteins (Mudgett & Staskawicz, 1998). *hrpG* and *hrpX* mutants of Xoo were generated in this study in order to determine if mutations in these genes would affected XadA

expression. Suprisingly, it was found that was found the expression of XadA in XOM2 is not affected by the either HrpG or HrpX mutants of Xoo (Fig. 5.8).

The role of another two-component transcriptional regulatory system, composed of the sensor protein ColS and the regulatory protein ColR, in XadA expression was assessed. *colS* and *colR* mutant strains (S. Sujatha, M.R. Vishnupiya, H. Patel, A. Pandey & R. Sonti, unpublished results) too had no effect on XadA expression in XOM2 (Fig. 5.8).

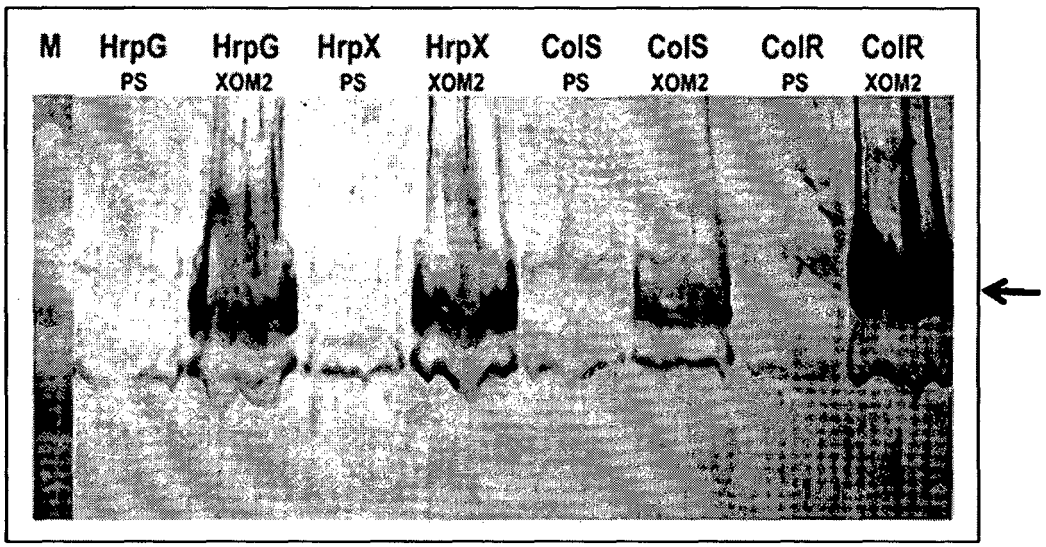


Figure 5.8. Western blot analysis of XadA expression in transcriptional regulator mutants of Xoo. HrpGPS: *hrpG* mutant grown in PS; HrpGXOM2: *hrpG*⁻ mutant grown in XOM2 pH5.5; HrpXPS: *hrpX*⁻ mutant grown in PS; HrpXXOM2: *hrpX*⁻ mutant grown in XOM2 pH5.5; ColSPS: *colS*⁻ mutant grown in PS; ColSXOM2: *colS*⁻ mutant grown in XOM2 pH5.5; ColRPS: *colR*⁻ mutant grown in PS; ColRXOM2: *colR*⁻ mutant grown in XOM2 pH5.5. The arrow indicates the expected 110KDa XadA band.

5.4.6. Glutamate-dependent XadA expression is post-transcriptionally regulated

The presence of *xadA* transcripts in PS, XOM2 and XBGlu media was assessed by a PCR using whole cDNA samples as templates. Total RNA was isolated from Xoo cultures grown in different media and converted into cDNA using random hexamers for first strand synthesis. Strikingly, *xadA* transcripts are present under all media conditions (Fig. 5.9a). Real Time PCR

(RT-PCR) confirmed this result and indicated that PS-grown Xoo cells showed more transcripts of *xadA* (Fig. 5.9b). The overall amount of *xadA* RNA in comparison to the control 16s rRNA in all samples was found be very less, indicating a low expression of XadA in Xoo cells. Since *xadA*⁻ strain BX0837 has a Tn5 insertion in the middle of the gene, it is known to express XadA protein in a truncated form (Ray et al., 2001). This is the reason for the presence of *xadA* transcript in the BXO837 strain. Nevertheless, RT-PCR was performed with a different set of primers specific to the 3' end of *xadA* and as expected, no transcript was seen in BXO837 samples (data not shown). The presence of *xadA* transcript under all nutrient conditions and the induction of the protein expression in plant mimic-medium suggest a post-transcriptional regulation of XadA.

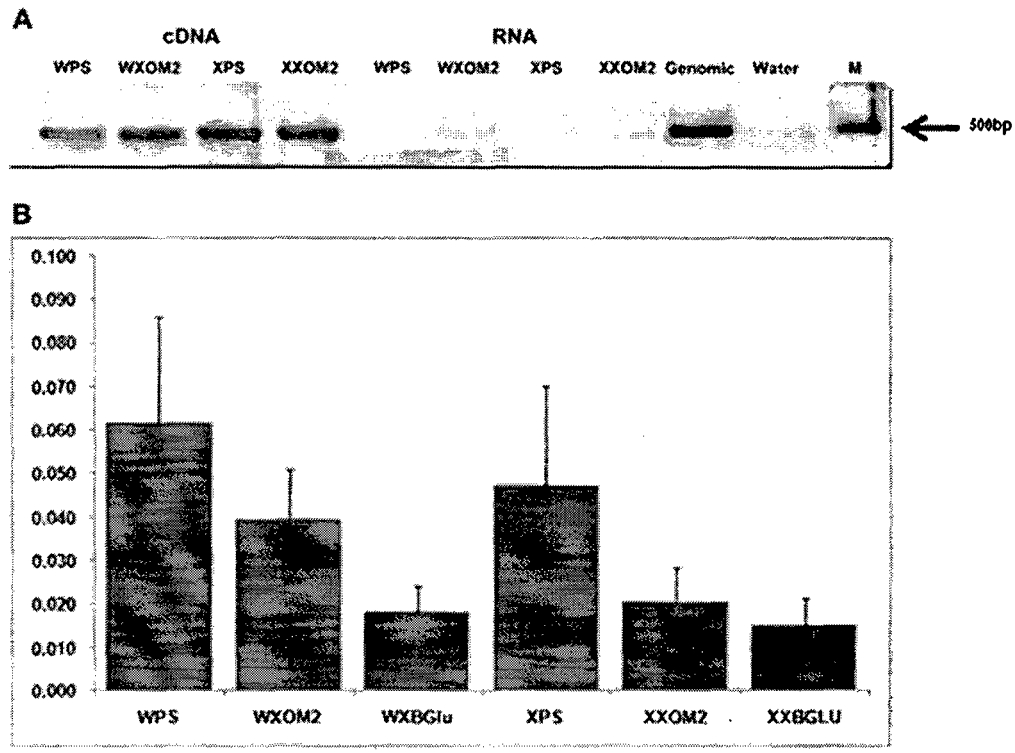


Figure 5.9. *xadA* transcripts are present in all media conditions. (A) PCR using cDNA templates prepared using total RNA from Xoo cells. RNA from WPS: Wild-type grown in PS; WXOM2: Wild-type in XOM2 pH 5.5; XPS: *xadA*⁻ in PS; XXOM2. *xadA*⁻ in XOM2 pH5.5 was isolated and converted into cDNA using first strand synthesis. PCR with corresponding RNA samples was performed to assess the genomic DNA contamination. PCR with wild-type genomic DNA and water were used as positive and negative controls, respectively. **(B)** Real Time PCR analysis of *xadA* transcripts in various media conditions: WPS: Wild-type grown in PS; WXOM2: Wild-type in XOM2 pH 5.5; WXBGlu: Wild-type grown in XB + Glutamate; XPS: *xadA*⁻ mutant grown in PS; XXOM2. *xadA*⁻ mutant in XOM2 pH5.5; XXBGlu: *xadA*⁻ mutant grown in XB + Glutamate. Y-axis represents the relative enrichment of *xadA* transcript in each media condition (X-axis) with respect to the 16s rRNA internal control.

5.4.7. The *xadA* transcript has a long 5' untranslated region

Another interesting feature that was observed in the *xadA* sequence was the lack of -10 and -35 promoter elements immediately upstream of the translational start site of *xadA*. Promoter prediction softwares (Neural Network Promoter Prediction software; BPROM software) predicted promoter elements atleast 150bp upstream of the translational start site (data not shown). However, the accuracy of these softwares is known to be low for bacterial sequences. Therefore, a primer extension analysis was performed using the total RNA isolated from wild-type Xoo cells grown in PS, XOM2 and in XB + glutamate medium. The primer extension product demonstrated that *xadA* has a long 5' untranslated region (UTR) and the transcriptional start site is -178bp upstream of the translational start site in Xoo cells grown in any of the three media (Fig. 5.10).

5.5. Inferences from the study

Pathogenic bacteria respond to the highly heterogeneous environment of their eukaryotic host by optimally regulating the expression of their genes. In order to understand the host-pathogen interactions better, it is crucial to identify the host cues that are used by the pathogen to regulate expression of virulence factors. Although some 'host-like' conditions can be mimicked *in vitro* (e.g., low iron, acidic pH, antimicrobial peptides and oxidative stress), the host environment remains too complex and dynamic to be accurately modelled in the laboratory.

In the case of rice-Xoo interactions, a reasonably good model of the rice xylem sap, wherein Xoo grows, has been constituted in the form of XOM2 medium in which, the expression of several plant-inducible promoters is observed. The expression of the trimeric autotransporter adhesin XadA is induced in this medium. Moreover, the presence of glutamate or glutamine only and no other amino acid including aspartate and asparagine could induce XadA expression. However, this induction was found to be independent of several known response regulators like HrpG and HrpX that transcriptionally regulate the expression of genes in response to either plant mimic media or growth *in planta*. Real time PCR analysis indicates that *xadA* transcripts are present in the PS medium and that expression of XadA protein in XOM2 appears to be regulated at the post-transcriptional level.

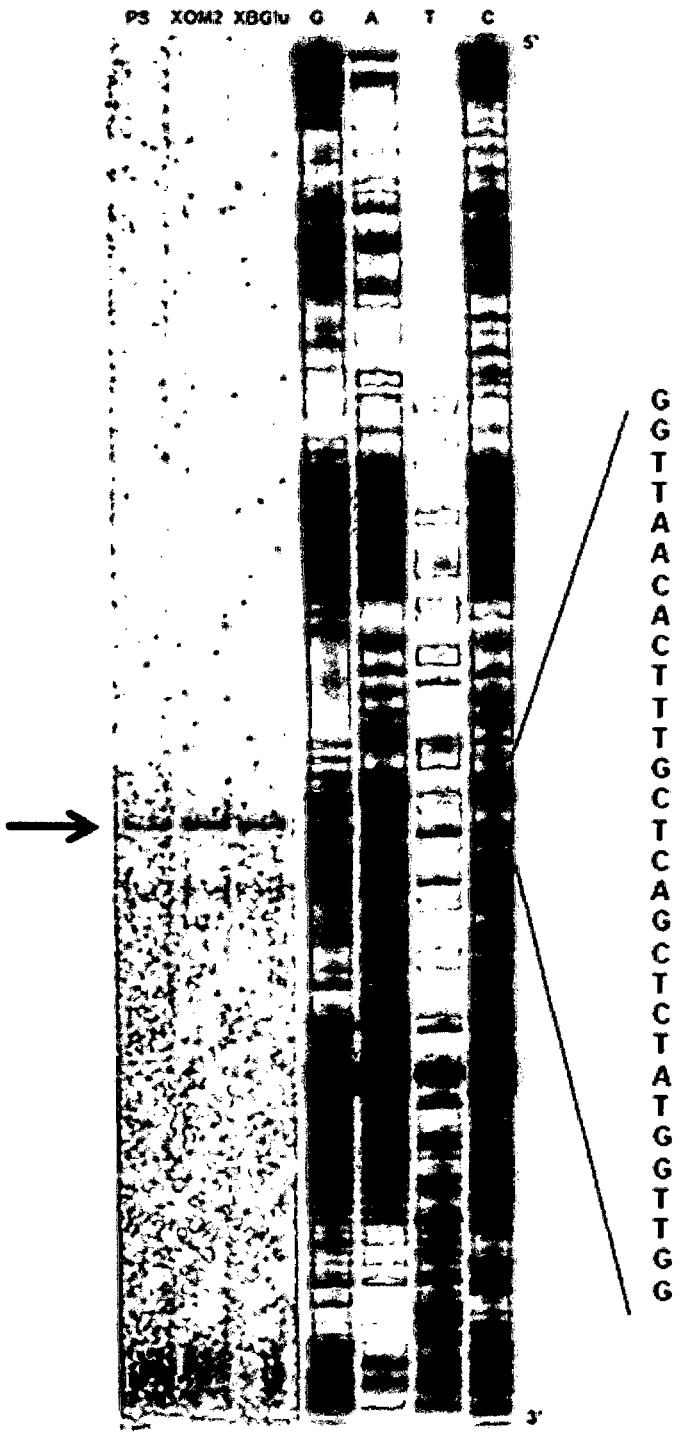


Figure 5.10. *xadA* transcript has a long 5'UTR. Primer extension analysis of *xadA* transcript, using a PCR primer that anneals to the xadA RNA around the translational start site, reveals that *xadA* has a long 5'UTR with the transcriptional start site at -178 with respect to the translational start site. All the experiments were performed with wild-type Xoo cells. RNA was isolated from WPS: Wild-type grown in PS; WXOM2: Wild-type in XOM2 pH 5.5; WXBGlu: Wild-type induced in XB + Glutamate and used for first strand synthesis using the α-P^{32} end-labelled primer. The lanes labelled as G, A, T and C represent manual sequencing using the corresponding labelled dNTPs.

A post-transcriptional control can be mediated at the translation step or after translation. In order to assess whether the long leader sequence of *xadA* has a role in the glutamate-dependent post-transcriptional regulation of XadA expression, a β-glucuronidase (GusA) reporter-tagged construct of *xadA* has been generated. This construct is a translational fusion of *xadA* with *gusA* ~150bp downstream to the translational start site. GusA enzyme assay experiments with this Xoo construct are in progress. Very preliminary results with a *xadA::gusA* reporter suggest that XadA expression is controlled atleast in part, at the translational level. A common mechanism of translational regulation of protein expression in bacteria is brought about either by a *cis*-acting RNA secondary structure or by non-coding small RNAs. Bacterial protein translation, in general, is coupled to the gene transcription process and small RNAs provide an alternative to this tight coupling by providing a means of stabilizing the mRNA for a later use in translation (Gottesman, 2005). These RNAs are typically 40-400bp in length, transcribed from the opposite strand from their targets and therefore, able to base pair extensively with their target mRNA (Majdalani et al., 2005). Apart from these *cis*-acting RNAs, several *trans*-acting small RNAs are also known to regulate bacterial gene expression, including genes involved in bacterial pathogenesis. Genome-wide searches for these small RNAs in *E. coli*, and, more recently, in other bacteria as well have led to the identification of more than 200 predicted small RNAs and of these, the expression of 60 small RNAs has been confirmed by Northern blotting in *E. coli* K12 (Storz et al., 2004). *In silico* studies show the presence of small regulatory RNAs in the genomes of all the sequenced Xanthomonas strains (Luban & Kihara, 2007). In general, small RNAs are retained in their unstructured form to facilitate proper binding to their template by chaperones such as Hfq (host factor for the replication of the Qβ phage RNA) and CsrA (carbon storage regulator) (Majdalani et al., 2005). Sequence homology-based identification of Hfq protein from Xanthomonas species is reported (Sauter et al., 2003). The expression of XadA was found to remain unaltered in an *hfq* mutant of Xoo under the various media conditions (data not shown). However, since there are non-coding RNAs in the Xoo transcriptome that function independent of Hfq, this mode of translational control of XadA expression cannot be ruled out.

The post-translational control of XadA expression is also plausible as XadA might be produced and rapidly degraded in PS medium while being stabilized in XOM2. It is interesting to note that XadA protein is intrinsically unstable in *E. coli*. The degradation of

XadA, as soon as it is translated, in nutrient-rich condition might atleast partially be responsible for the absence of XadA under these growth conditions. The possibility of a post-translational modification, such as glycosylation, in minimal plant-mimic conditions but not in PS or when overexpressed in *E. coli* needs to be explored. A search for putative N-glycosylation sites in XadA sequence yields atleast seven sites (residues 63, 134, 140, 209, 556, 717 & 986) at very high confidence levels (data not shown) in the software NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/), which are located on the surface of XadA model presented in Chapter 4 (Fig.4.7). In the absence of these post-translational modifications, XadA might be unstable. However, this mechanism of regulation of XadA expression might be less likely since the translation process is energetically demanding and it would be wasteful to synthesize a protein of 110KDa and then degrade it.

The data presented in the current study indicates the translation of XadA from a pre-existing pool of *xadA* transcript by Xoo upon sensing the presence of glutamate. It is interesting to note that there is a report stating that compounds in the rice hydathodal exudates act as chemoattractants for Xoo and that glutamate is the chemoattractant (Feng & Guo, 1975). It is also noteworthy that glutamate is one of the free amino acids found in hydathodal exudates from several grasses (Goatley & Lewis, 1966). It is interesting to note that XadA expression is also induced by glutamine, which is also present in the rice xylem sap (Oaks, 1992). The results suggest that rapid expression of XadA is achieved by keeping the transcripts ready under all conditions but expressing XadA only after sensing the glutamate that is exuded into the hydathodal pores on the rice leaf tips. In this manner, Xoo may be using glutamate as a cue that it is in the presence of rice hydathodes and that it is time to express XadA. Previous work in our lab has shown that XadA protein is involved in promoting plant colonization (Das et al., 2009).

The results discussed in this chapter also indicate that XadA is stably expressed in the glutamate-containing XOM2 medium. This feature might prove useful while designing a strategy for large-scale XadA purification. The problems encountered during the overexpression and purification of XadA from *E. coli* can be overcome by scaling up the XadA enrichment method described here.

Future Plans

FUTURE PLANS

The pathogenic interaction of Xoo with rice provides a good platform for studying plant host factors, bacterial virulence factors, their interplay and the causation of disease. Moreover, the Xoo-rice pathosystem demands a careful study because the consequential disease, bacterial blight of rice, is a devastating disease causing severe yield losses across the rice-growing regions of the world. Several years of research and a lot of recent interest in this field has succeeded in identifying many Xoo virulence factors that are crucial for its pathogenesis (Subramoni et al., 2006; Jha & Sonti, 2009). The work presented in this thesis is a structural and functional characterization of two such virulence factors.

High-resolution crystal structure of the Xoo esterase LipA led to the identification of a unique mode of plant substrate recognition by a bacterial enzyme. This finding has opened several avenues for further exploration. Ongoing work in this direction includes characterization of a natural ligand of LipA from the rice cell wall extracts. Preliminary results show that upon treatment of crude rice cell wall material with LipA, a small soluble molecule 'elicitor' is released that can induce defense responses, including programmed cell death, in the rice cells. Presumably, this molecule is a degradation product of the plant cell wall that is sensed by the host upon LipA action. This 'product' of LipA action might provide a good estimation of the natural substrate of LipA and a combinatorial approach with NMR, MS-MS and other analytical chemistry techniques can be employed to characterize it. Another interesting aspect of the study can be the crystal structures of LipA homologs from other plant-pathogenic Xanthomonads. *Xanthomonas campestris* pv. campestris has two LipA homologs while *Xylella fastidiosa* has three. Since only one of the LipA-like proteins in both species have a carbohydrate-binding pocket, it would be very interesting to identify 'gain-of-function' mutation that would endow sugar binding in LipA homologs in these species that do not possess a LipA-like ligand binding. Experiments to alter the sugar specificity in Xoo LipA by site-directed mutagenesis would also be interesting. Finally, the possibility of use of LipA in plant cell wall degrading enzyme cocktails for degradation of agricultural wastes towards biofuel production can also be explored.

Another major aspect of study in this thesis is the expression analysis of the Xanthomonas adhesin-like factor A (XadA) wherein the conditional expression of this virulence factor was

found to be glutamate-dependent and post-transcriptionally regulated. The exciting observation that glutamate, a key component of plant hydathodal exudates, has a role in XadA expression needs to be substantiated with experiments to explore the role of glutamate, if any, in the post-transcriptional control of XadA. The *xadA::gusA* reporter fusion construct generated in this study can be used to perform whole genome screen for regulators of XadA expression, which might reveal a glutamate-sensing regulatory RNA or a completely novel pathway involved in sensing the plant milieu that might also be responsible for the interesting expression profile of XadA. The stable expression of XadA in Xoo is also important knowledge since XadA protein was found to be intrinsically unstable upon overexpression in *E. coli*. The purification strategy of XadA from Xoo and solubilisation of this membrane-anchored protein in SDS can be used for crystallization purposes.

Xoo is under a high selection pressure to survive and efficiently cause disease in the rice host. It is clear from the current study that both LipA and XadA show interesting structural features that confer evolutionary advantage towards proficient entry and infection of the host. Considering the limited available information on structures of Xoo virulence factors, it is essential to generate a repertoire of structures to analyse the adaptive features of the whole Xoo proteome. This would aid in our efforts to understand the Xoo-rice pathosystem intricately, thus adding to the basic knowledge and possibly, to the innovation of new methods of disease control through generation of molecules that block the action of virulence factors or act as novel bactericides.

Bibliography

BIBLIOGRAPHY

Abramovitch, R. B., Anderson, J. C., and Martin, G. B. (2006) Bacterial elicitation and evasion of plant innate immunity. Nat. Rev. Mol. Cell Biol. 7: 601-611.

Albersheim, P., Jones, T.M., and English, P.D. (1969). Biochemistry of the cell wall in relation to the infective processes. Ann. Rev. Phytopathol. 7: 171-194.

Alfano J. R. and Collmer A. (2004) Type III secretion system effector proteins: double agents in bacterial disease and plant defense. Ann. Rev. Phytopathol. 42: 385-414.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Mifler, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST; a new generation of protein database search programs. Nucleic Acids Res. 25: 3389-3402.

Andrews, J. H., and Harris, R. F. (2000) The ecology and biogeography of microorganisms on plant surfaces. Ann. Rev. Phytopathol. 38:145–180.

Aparna, G., Chatterjee, A., Jha, G., Sonti, R. V., and Sankaranarayanan, R. (2007) Crystallization and preliminary crystallographic studies of LipA, a secretory lipase/esterase from *Xanthomonas oryzae* pv. oryzae. Acta Cryst. F63: 708-710.

Arlat, M., Gough, C. L., Zischek, C., Barberis, P. A., Trigalet, A. and Boucher, C. A. (1992) Transcriptional organization and expression of the large hrp gene cluster of *Pseudomonas solanacearum*. Mol. Plant Microbe Interact. 5: 187-193.

Arthan, D., Kittakoop, P., Esen, A., and Svasti, J. (2006) Furostanol glycoside 26-O-β-glucosidase from the leaves of *Solanum torvum*. Phytochemistry. 67: 27-33.

Barnard, J., Dautin, N., Lukacik, P., Bernstein, H. D., Buchanan, S. K. (2007) Autotransporter structure reveals intra-barrel cleavage followed by conformational changes. Nat. Struct. Mol. Biol. 14: 1214-1220.

Bartlam, M., Wang, G., Yang, H., Gao, R., Zhao, X., Xie, G., Cao, S., Feng, Y., and Rao, Z. (2004). Crystal structure of an acylpeptide hydrolase/esterase from *Aeropyrum pernix* K1. Structure 12: 1481-1488.

Barton, G. J. (1993) Alscript- a tool to format multiple sequence alignments Prot. Eng. 6: 37-40.

Bauer, S., Vasu, P., Persson, S., Mort, A.J., and Somerville, C.R. (2006). Development and application of a suite of polysaccharide-degrading enzymes for analyzing plant cell walls. Proc. Natl. Acad. Sci. USA 103: 11417-11422.

Beattie, G. A. and Lindow, S. E. (1999) Bacterial colonization of leaves: a spectrum of strategies. Phytopathol. 89: 353-359.

Beattie, G. A., and Lindow, S. E. (1995) The secret life of foliar bacterial pathogens on leaves. Ann. Rev. Phytopathol. 33: 145-172.

Beaulieu, C. and van Gijsegem, F. (1990) Identification of plant-inducible genes in *Erwinia chrysanthem.i* J. Bacteriol. 172: 1569-1575.

Benz, I. and Schmidt, M. A. (2002) Never say never again: protein glycosylation in pathogenic bacteria. Mol. Microbiol. 45: 267-276.

Bernstein, H. D. (2007) Are bacterial 'autotransporters' really transporters? Trends Microbiol. 15: 441-447.

Bestwick, C. S., Bennett, M. H., and Mansfield, J. W. (1995) Hrp mutant of *Pseudomonas syringae* pv. phaseolicola induces cell wall alteration but not membrane damage leading to the hypersensitive reaction in lettuce. Plant Physiol. 108:, 503-516.

Bianchi, G., Lupotto, E. and Russo, S. (1979) Composition of epicuticular wax of rice, *Oryza sativa.* 35: 1417-1420.

Biely, P. (2003) Xylanolytic enzymes. Handbook Food Enzymol., Marcel Dekker, New York and Basel. 879-915.

Blakeman, J. P. (1991) Foliar plant pathogens: epiphytic growth and interactions on leaves. J. Appl. Bacteriol. 70(Suppl.):49S-59S.

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., et al. (2007) Plant carbohydrate scavenging through TonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. Plos One 2: e224.

Blow, D. (2002) Outline of Crystallography for Biologists. Oxford: Oxford University Press.

Blundell, T. L. and Johnson, L. N. (1976) *Protein Crystallography.* London: Academic Press.

Boller, T. (1995) Chemoperception of microbial signals in plant cells. Ann. Rev. Plant Physiol. Plant Mol. Biol. 46: 189–214.

Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382: 769-781.

Bradbury, J. F. (1970a) *Xanthomonas oryzae* CMI descriptions of pathogenic fungi and bacteria no. 239. CAB International, Wallingford, UK.

Brencic, A., and Winans, S. C. (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. Microbiol. Mol. Biol. Rev. 69: 155-194.

Brito, B., Marenda, M., Barberis, P., Boucher, C. and Genin, S. (1999) *prhJ* and *hrpG*, two new components of the plant signal-dependent regulatory cascade controlled by PrhA in *Ralstonia solanacearum*. Mol Microbiol. 31: 237-251.

Buranov, A. U., and Mazza, G. (2008) Lignin in straw of herbaceous crops. Industrial Crops and Products 28: 237-259.

Burkholder, W. H. (1948) Bacteria as plant pathogens. Ann. Rev. Microbiol. 2: 389-412.

Cantarel, B. L. (2008) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucl. Acids Res. 37: D233-238.

Carpita, N. C (1996) Structure and biogenesis of the cell walls of grasses. Ann. Rev. Plant Physiol. Plant Mol. Biol. 47: 445-476.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B.G., Parkhill, J. (2005) ACT: the Artemis Comparison Tool. Bioinformatics. 21: 3422-3423.

Casadevall A., and Pirofski L. A. (1999) Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infect Immun.67:3703-3713.

Cassab, G. L and Varner, J. E. (1988) Cell wall proteins. Ann. Rev. Plant Physiol. Mol. Biol., 39: 321-353.

Chan, J. W and Goodwin, P. H. (1999) A physical map of the chromosome of *Xanthomonas campestris* pv. phaseoli var. fuscans BXPF65. FEMS microbiol. lett. 180: 85-90.

Chao, N. X., Wei, K., Chen, Q., Meng, Q. L., Tang, D. J., He, Y. Q., Lu, G. T., Jiang, B. L., et al., (2008) The rsmA-like gene r*smA*(Xcc) of *Xanthomonas campestris* pv. campestris is involved in the control of various cellular processes, including pathogenesis. Mol Plant Microbe Interact. 21: 411-423.

Chen, J. C. H., Miercke, L. J. W., Krucinski, J., Starr, J. R., Saenz, G., Wang, X., et al., (1998) Structure of Bovine Pancreatic Cholesterol Esterase at 1.6 Å: Novel Structural Features Involved in Lipase Activation. Biochemistry 37: 5107-5117.

Chisholm, S. T., Coaker, G., Day, B. and Staskawicz, B. J. (2006) Host-Microbe Interactions: Shaping the Evolution of the Plant Immune Response. Cell 124: 803-814.

Chothia, C. (1993). Proteins: one thousand families for the molecular biologist. Nature 357: 543-544.

Cianciotto, N. P. (2005) Type II secretion: a protein secretion system for all seasons. Trends Microbiol. 13: 581-588.

Clamp, M., Cuff, J., Searle, S. M. and Barton, G. J. (2004) The Jalview Java Alignment Editor. Bioinformatics 20: 426-7

Computational Collaborative Project, Number 4 (1994). The CCP4 Suite: Programs for Protein Crystallography. Acta Cryst. D50: 760–763.

Cornelis, G. (2006) The type III secretion injectisome. Nat. Rev. Microbiol. 4: 811-825.

Cowtan, K. (1994) DM: an automated procedure for phase improvement by density modification. Joint CCP4 & ESF-EACBM Newsletter on Protein Crystallography. 31: 34-38.

da Silva, A. C., Ferro, J. A., Reinach, F. C., Farah, C. S., Furlan, L. R., et al. (2002) Comparison of the genomes of two Xanthomonas pathogens with differing host specificities. Nature 417: 459-463.

Dangl, J. L. and Jones, J. D. G (2001) Plant pathogens and integrated defence responses to infection. Nature 411: 826-833.

Darvill, A. G., and Albersheim, P. (1984) Phytoalexins and their elicitors-a defense against microbial infection in plants. Ann. Rev. Plant Physiol. 35: 243-275.

Das A., Rangaraj N. and Sonti R. V. (2009) Multiple adhesin-like functions of *Xanthomonas oryzae* pv. oryzae are involved in promoting leaf attachment, entry, and virulence on rice. Mol Plant Microbe Interact. 22: 73-85.

Dautin N. and Bernstein H. D. (2007) Protein secretion in gram-negative bacteria via the autotransporter pathway. Ann. Rev. Microbiol. 61: 89-112.

Dawkins, R. (1982) The Extended Phenotype, Oxford University Press, UK.

de Graaf, F. K. and Mooi, F. R. (1986) The fimbrial adhesins of Escherichia coli. Adv Microb Physiol. 65-143.

Deeds, E. J., Shakhnovich, E. I. (2006) A structure-centric view of protein evolution, design, and adaptation. Adv. Enzymol. Relat. Areas. Mol. Biol. 75: 133-191.

Desvaux, M., Parham, N. J. and Henderson I. R. (2004) Type V protein secretion: simplicity gone awry? Curr Issues Mol Biol. 6: 111-24.

Desvaux, M., Parham, N. J. and Henderson, I. R. (2004) The autotransporter secretion system. Res Microbiol. 155: 53-60.

DeWit, P. J. G. M. (2007) How plants recognize pathogens and defend themselves. Cell. Mol. Life Sci. 64: 2726-2732.

Ding, X., Cao, Y., Huang, L., Zhao, J., Xu, C., Li, X., and Wang, S. (2008) Activation of the indole-3-acetic acid amido synthetase GH3-8 suppresses expansin expression and promotes salicylate- and jasmonate-independent basal immunity in rice. Plant Cell 20: 228-240.

Dirk, L., Tanja, R., Ingo, B. A., Andrei, L., Volkhard, A. J. K. (2006) Trimeric autotransporter adhesins: variable structure, common function. Trends Microbiol. 14: 264-270.

Dodson G, Wlodawer A. (1998) Catalytic triads and their relatives. Trends Biochem Sci. 23:347-52.

Dowson, W. J. (1939) On the systematic position and generic names of the gram-negative bacterial plant pathogens. Zentralbl Baklerio Parasilenkti. II, 100: 177-193.

Drenth, J. (1999) Principles of Protein X-Ray Crystallography. New York: Springer-Verlag.

Dunn, A. K. and Handelsman, J. (2002) Toward an understanding of microbial communities through analysis of communication networks. Antonie Van Leeuwenhoek. 81: 565-574.

Ealick, S. E. (2000) Advances in multiple wavelength anomalous diffraction crystallography. Curr. op. chem.biol. 4: 495-9.

El Tahir, Y., Kuusela, P. and Skurnik, M. (2000) Functional mapping of the *Yersinia enterocolica* adhesin YadA: identification of eight NSVAIG-S motifs in the amino-terminal half of the protein involved in collagen binding. Mol Microbiol 37: 192-206.

Ericsson, D.J., Kasrayan, A., Johansson, P., Bergfors, T., Sandstrom, A.G., Backvall, J.E., and Mowbray, S.L. (2008). X-ray structure of *Candida antarctica* lipase A shows a novel lid structure and a likely mode of interfacial activation. J. Mol. Biol. 376: 109-119.

Esquerre-Tugaye, M.T., Boudart G., and Dumas, B. (2000). Cell wall degrading enzymes, inhibitory proteins, and oligosaccharides participate in the molecular dialogue between plants pathogens. Plant Physiol. Biochem. 38: 157–163.

Esseaberg, M., Cason, E. T., Hamilton, B., Brinkerhoff, L. A., Gholson, R. K., Richardson, P. E. (1979) Single cell colonies of *Xanthomonas malvacearum* in susceptible and immune cotton leaves and the local resistant response to colonies in immune leaves. Physiol. Plant Pathol. 15: 53-56

Expert, D., Enard, C. and Masclaux, C. (1996) The role of iron in plant host-pathogen interactions. Trends Microbiol. 4: 232-237.

Fang, C. T., Ren, H. C., Chen, T. Y., Chu, Y. K., Faan, H. C. and Wu, S. C. (1957) A comparison of the rice bacterial leaf blight organism with the bacterial leaf streak organism of rice. Acta Phytopathol. Sin. 3: 99-124.

Feng, T. Y. and Guo, T. T. (1975) Bacterial leaf blight of rice plant. VI. Chemotactic response of *Xanthomonas oryzae* pv. *oryzae* to water droplets exudated form water pores on the leaf of rice plants. Bot. Bull. Acad. Sin. (Taipei) 16: 126–136.

Finlay, B. B. and Cossart, P. (1997) Exploitation of mammalian host cell functions by bacterial pathogens. Science. 276: 718-725.

Finn, R. D., Tate, J., Mistry, J., Coggill, P.C., Sammut, J.S., Hotz, H.R., Ceric, G., Forslund, K., et al., (2008) The Pfam protein families database Nucleic Acids Res. Database Issue 36: D281-D288.

Flor, H. H. (1956) The complementary genetic systems in flax and flax rust. Adv. Genet. 8: 29-54.

Flor, H. H. (1971) Current status of the gene-for-gene concept. Ann. Rev. Phytopathol. 9:275–96.

Frank, S. A. and Schmid-Hempel, P. (2008) Mechanisms of pathogenesis and the evolution of parasite virulence. J. Evol. Biol. 21: 396-404.

Furutani, A., Takaoka, M., Sanada, H., Noguchi, Y., Oku, T., Tsuno, K., Ochiai, H. and Tsuge S. (2009) Identification of Novel Type III Secretion Effectors in *Xanthomonas oryzae* pv. oryzae. Mol. Plant Microbe Interact.22: 96-106.

Furutani, A., Tsuge, S., Oku, T., Tsuno, K., Ochiai, H., Inoue, Y., Kaku, H. and Kubo, Y. (2003) Hpa1 secretion via type III secretion system in *Xanthomonas oryzae* pv. oryzae. J.Gen. Plant Pathol. 69: 271-275.

Ghosh, D., Erman, M., Sawicki, M., Lala, P., Weeks, D. R., Li, N. Y., Pangborn, W., et al., (1999) Determination of a protein structure by iodination: the structure of iodinated acetylxylan esterase. Acta Cryst. D55: 779-784.

Gilbert, H.J., Stalbrand, H., and Brumer, H. (2008) How the walls come crumbling down: recent structural biochemistry of plant polysaccharide degradation. Curr. Op. Plant Biol. 11: 338-348.

Goatley, J. L. and Lewis, R. W. (1966) Composition of Guttation Fluid from Rye, Wheat, and Barley Seedlings. Plant Physiol. 41: 373-375.

Goldstein, R. A. (2008) The structure of protein evolution and the evolution of protein structure. Curr. Opin. Struct. Biol. 18: 170-177.

Goodman, R. N. and Novacky, A. J. (1994) The hypersensitive reaction in plants to pathogens. St Paul: APS Press.

Gopalan, S., Bauer, D. W., Alfano, J. A., Loniello, A. O., He, S. Y. and Collmer, A. (1996) Expression of the *Pseudomonas syringae* avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive cell death. Plant Cell 8, 1095–1105.

Gottesman, S. (2005) Micros for microbes: non-coding regulatory RNAs in bacteria. Trends Genetics 21: 399-404.

Grant, S. R., Fisher, E. J., Chang, J. H., Mole, B. M. and Dangl, J. L. (2006) Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. Ann. Rev. Microbiol. 60: 425-449.

Hauck, P., Thilmony, R., and He, S. Y. (2003) A *Pseudomonas syringe* type III effector suppresses cell wall-based extracellular defense in susceptible Arabidopsis plants. Proc. Natl. Acad. Sci. U.S.A. 100: 8577-8582.

He, P., Shan, L. and Sheen, J. (2007) Elicitation and suppression of microbe-associated molecular pattern-triggered immunity in plant–microbe interactions. Cellular Microbiol. 9: 1385-1396.

Hegyi, H., and Gerstein, M. (1999) The relationship between protein structure and function: a comprehensive survey with application to the yeast genome J. Mol. Biol. 288: 147-164.

Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. and Ala'Aldeen, D. (2004) Type V protein secretion pathway: the autotransporter story. Microbiol. Mol. Biol. Rev. 68:692-744.

Hermoso, J.A., Sanz-Aparicio, J., Molina, R., Juge, N., Gonzalez, R., and Faulds, C.B. (2004) The crystal structure of feruloyl esterase A from *Aspergillus niger* suggests evolutive functional convergence in feruloyl esterase family. J. Mol. Biol. 338: 495-506.

Hilaire, E., Young, S. A, Willard, L. H., McGee, J. D., Sweat, T., Chittoor, J. M., Guikema, J. A. and Leach, J, E. (2001) Vascular defense responses in rice: peroxidase accumulation in xylem parenchyma cells and xylem wall thickening. Mol. Plant Microbe Int. 14: 1411-9.

Hoiczyk, E., Roggenkamp, A., Reichenbecher, M., Lupas, A. and Heesemann J. (2000) Structure and sequence analysis of YersiniaYadA and Moraxella UspAs reveal a novel class of adhesins. EMBO J.19: 5989-5999.

Holm, L., and Sander, C. (1998). Dictionary of recurrent domains in protein structures. Proteins 33: 88-96.

Holmquist, M. (2000) Alpha/beta-hydrolase fold enzymes: structures, functions and mechanisms. Curr. Protein. Pept. Sci. 1: 209-235.

Horino, O. (1981) Ultrastructural histopathology of rice leaves infected with *Xanthomonas campestris* pv. oryzae on Kogyok group rice varieties with different levels of resistance at the seedling stage. Ann. Phytopathol. Soc. Jpn. 47: 501-509.

Howard, G. C. and Brown, W. E., eds. (2002) Modern Protein Chemistry: Practical Aspects. CRC Press, Boca Raton.

Howard, R. L., Abotsi, E., Jansen van Rensburg, E. L. and Howard, S. (2003) Lignocellulose biotechnology: issues of bioconversion and enzyme production, Afr. J. Biotechnol. 2: 602–619.

Hu, J., Qian, W. & He, C. (2007) The *Xanthomonas oryzae* pv. oryzae eglXoB endoglucanase gene is required for virulence to rice. FEMS Microbiol. Lett. 269, 273-279.

Hultgren, S. J., Abraham, S., Caparon, M., Falk, P., St. Geme, J. W. and Normark, S. (1993) pilus and nonpilus bacterial adhesins: assembly and function in cell recognition. Cell 73: 887-901.

Hussain, M., Becker, K., von Eiff, C., Peters, G. and Herrmann M. (2001) Analogs of Eap Protein Are Conserved and Prevalent in Clinical *Staphylococcus aureus* Isolates. Clin Diagn Lab Immunol. 8: 1271-1276.

Huynh, T. V., Dahlbeck, D. and Staskawicz, B. J. (1989) Bacterial blight of soybean: regulation of a pathogen gene determining host cultivar specificity. Science. 245, 1374–1377.

Ito, K., Nakajima, Y., Xu, Y., Yamada, N., Onohara, Y., Ito, T., Matsubara, F., et al., (2006). Crystal structure and mechanism of tripeptidyl activity of prolyl tripeptidyl aminopeptidase from *Porphyromonas gingivalis*. J. Mol. Biol. 362: 228-240.

Jaeger, K.-E., Dijkstra, B.W., and Reetz, M.T. (1999). Bacterial biocatalysts: Molecular biology, structure and biotechnological applications of lipases. Ann. Rev. Microbiol. 53: 315-351.

Jha, G. and Sonti, R. V. (2009) Attach and defense in Xanthomonas-rice interactions. Proc. Indian Natl. Sc. Acad. 75: 49-68.

Jha, G., Rajeshwari, R. and Sonti, R. V. (2007) Functional interplay between two *Xanthomonas oryzae* pv. oryzae secretion systems in modulating virulence on rice. Mol. Plant Microbe Interact. 20: 31-40.

Jones, J. D. G. and Dangl, J. L. (2006) The plant immune system. Nature. 444: 323-329.

Jones, T.A., Zou, J-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for the building of protein models in electron density maps and the location of errors in these models. Acta Cryst. A47: 110-119.

Kamachi, K., Yamaya, T., Mae, T. and Ojima, K. (1991) A role for glutamine synthetase in the remobilization of leaf nitrogen during natural senescence in rice leaves. Plant Physiology 96, 411-417.

Kauffman, H. E. and Reddy, A. P. K. (1975) Seed transmission studies of *Xanthomonas oryzae* in rice. Phytopathol. 65: 663-666.

Kauffman, H. E., Reddy, A. P. K., Hsieh, S. P. Y., and Merca, S. D. (1973) An improved technique for evaluation of resistance of rice varieties to *Xanthomonas oryzae*. Plant Dis. Rep. 57: 537-541.

Knox, J. P. (2008) Mapping the walls of the kingdom: the view from the horsetails. New Phytologist 179: 1-3.

Kouker, G., and Jaeger, K. (1987) Specific and sensitive plate assay for bacterial lipases. Appl. Environ. Microbiol. 53: 211-213.

Lafontaine, E. R., Wagner, N. J., Hansen, E. J. (2001) Expression of the *Moraxella catarrhalis* UspA1 protein undergoes phase variation and is regulated at the transcriptional level. J. Bacteriol. 183: 1540-1551.

Lambrechts, L., Fellous, S. and Koella, J. C. (2006) Coevolutionary interactions between host and parasite genotypes. Trends Parasitol. 12-16.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, et al., (2007) ClustalW and ClustalX version 2. Bioinformatics 23: 2947-2948.

Lee, B. M., Park, Y. J., Park, D. S., Kang, H. W., Kim, J. G., et al. (2005) The genome sequence of *Xanthomonas oryzae* pv. oryzae KACC10331, the bacterial blight pathogen of rice. Nucleic Acids Res. 33: 577-586.

Lee, D., Grant, A., Buchan, D., and Orengo, C. (2003) A structural perspective on genome evolution. Curr. Op. Struc. Biol. 13: 359-369

Lindgren, P. B., Peet, R. C. and Panopoulos, N. J. (1986) Gene cluster of Pseudomonas syringae pv. phaseolicola controls pathogenicity of bean plants and hypersensitivity of non-host plants. J Bacteriol 168: 512-522.

Linke, D., Riess, T., Autenrieth, I. B., Lupas, A. and Kempf, V. A. (2006) Trimeric autotransporter adhesins: variable structure, common function. Trends Microbiol. 14: 264-70.

Lu, H., Patil, P., Van Sluys, M., White, F. F., Ryan, R. P., et al. (2008) Acquisition and evolution of plant pathogenesis–associated gene clusters and candidate determinants of tissue-specificity in xanthomonas. Plos one 3: e3828.

Luban, S, and Kihara, D. (2007) Comparative genomics of small RNAs in bacterial genomes. Omics. 11: 58-73.

Majdalani, N., Vanderpool, C. K. and Gottesman, S. (2005) Bacterial small RNA regulators. Crit Rev Biochem Mol Biol. 40: 93-113.

Marco, M. L., Legac, J. and Lindow, S. E. (2003) Conditional Survival as a Selection Strategy To Identify Plant-Inducible Genes of *Pseudomonas syringae*. Appl Environ Microbiol. 69: 5793-5801.

Marschner, H. (1995) Mineral nutrition of higher plants, 2nd ed. Academic Press, Ltd., London, United Kingdom.

Martin, P., van de Ven, T., Mouchel, N., Jeffries, A. C., Hood, D. W. and Moxon, E. R. (2003) Experimentally revised repertoire of putative contingency loci in *Neisseria meningitidis* strain MC58: evidence for a novel mechanism of phase variation. Mol. Microbiol. 50: 245-257.

Martinelle, M., Holmquist, M., and Hult, K. (1995) On the interfacial activation of *Candida antarctica* lipase A and B as compared with *Humicola lanuginosa* lipase Biochimica et Biophysica Acta 1258: 272-276.

Martinez, C., de Geus, P., Lauwereys, M., Matthyssens, M., and Cambillau, C. (1992) *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. Nature 356: 615-618.

Marti-Renom, M.A., Stuart, A., Fiser, A., Sanchez, R., Melo, F., and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. Ann. Rev. Biophys. Biomol. Struct. 29: 291-325.

McPherson, A. (1999) Crystallization of Biological Macromolecules. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

McRee, D. E. (1993) Practical Protein Crystallography. San Diego: Academic Press.

Meng G., Surana N. K., St Geme J. W. 3rd, Waksman G. (2006) Structure of the outer membrane translocator domain of the *Haemophilus influenza* Hia trimeric autotransporter. EMBO J. 25: 2297-304.

Mew, T. W. (1987) Current status and future prospects of research on bacterial blight of rice. Ann. Rev. Phytopathol. 25: 359-382.

Mew, T.W. (1989) Overview of the world bacterial blight situation. Proc. Int. Workshop Bact. Blight Rice, 14-18 March 1988, IRRI, Manila, Philippines.

Mizukami, T. and Wakimoto, S. (1969) Epidemiology and control of bacterial leaf blight of rice. Ann. Rev. Phytopathol. 7: 51-72.

Montesinos, E., Bonaterra, A., Badosa, E., Frances, J., Alemany, J., Llorente, I. and Moragrega, C. (2002) Plant-microbe interactions and the new biotechnological methods of plant disease control. Int. Microbiol. 5: 169-175.

Moore, A. D., Bjorklund, A. K., Ekman, D., Bornberg-Bauer, E. and Elofsson A. (2008) Arrangements in the modular evolution of proteins. Trends Biochem. Sc. 33: 444-451.

Mudgett, M. B. and Staskawicz, B. J. (1998) Protein signaling via type III secretion pathways in phytopathogenic bacteria. Curr Opin Microbiol 1: 109-114.

Nardini, M., and Dijkstra B.W. (1999). α/β Hydrolase fold enzymes: the family keeps growing. Curr. Op. Struct. Biol. 9: 732-737.

Nicholls, A., Sharp, K. A. and Honig, B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 11: 281-296.

Nino-Liu, D. O., Ronald, P. C. and Bogdanove, A. J. (2006) *Xanthomonas oryzae* pathovars: model pathogens of a model crop. Mol. Plant Pathol. 7:303-324.

Nummelin, H., Merckel, M. C., Leo, J. C., Lankinen, H., Skurnik, M. and Goldman, A. (2004) The Yersinia adhesin YadA collagen-binding domain structure is a novel left-handed parallel beta-roll. EMBO J. 23:701-11.

Nurnberger, T., Brunner, F., Kemmerling, B., and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. Immunol. Rev. 198, 249-266.

Ochiai, H., Inoue, Y., Takeya, M., Sasaki, A. and Kaku, H. (2005) Genome sequence of *Xanthomonas oryzae* pv. oryzae suggests contribution of large numbers of effector genes and insertion sequences to its race diversity. Jpn. Agric. Res. Q 39: 275-287.

Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., et al., (1992) The α/β hydrolase fold. Protein Engg. 5: 197-211.

Oomen, C. J., van Ulsen, P., van Gelder, P., Feijen, M., Tommassen, J. and Gros P. (2004) Structure of the translocator domain of a bacterial autotransporter. EMBO J. 23: 1257-1266.

Orengo, C. A., Michie, A. D., Jones, D. T., Swindells, M. B., and Thornton, J. M. (1997) CATH: a hierarchic classification of protein domain structures. Structure 5: 1093-1108.

Osbourn, A. E., Barber, C. E. and Daniels, M. J. (1987) Identification of plant-induced genes of the bacterial pathogen *Xanthomonas campestris* pv. campestris using a promoter-probe plasmid. EMBO J. 6: 23-28.

Otwinowski, Z. and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276: 307-326.

Pallen, M. J. and Wren, B. W. (2007) Bacterial pathogenomics. Nature 449: 835-842.

Pennell, R. I., and Lamb, C. (1997) Programmed cell death in plants. Plant Cell 9: 1157-1168.

Petrek, M., Kosinova, P., Koca, J., and Otyepka, M. (2007) MOLE: a Voronoi diagram-based explorer of molecular channels pores and tunnels. Structure 15: 1357-1363.

Pizarro-Cerda, J., Cossart, P. (2006) Bacterial adhesion and entry into host cells. 124: 715-727.

Pohlner, J., Halter, R., Beyreuther, K. and Meyer, T. F. (1987) Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. Nature 325: 458-462.

Qian, W., Jia, Y., Ren, S. X., He, Y. Q., Feng, J. X., et al. (2005) Comparative and functional genomics analyses of the pathogenicity of phytopathogen *Xanthomonas campestris* pv. campestris. Genome Res. 15: 757-767.

Rademaker, J. L. W., Hoste, B., Louws, F. J., Kersters, K., Swings, J., Vauterin, L., Vauterin, P., De Bruijn, F. J. (2000) Comparison of AFLP and rep-PCR genomic fingerprinting with DNA-DNA homology studies: Xanthomonas as a model system.. Int J Syst Evol Microbiol. 2: 665-77.

Rahme, L. G., Mindrinos, M. N. and Panopoulos, N. J. (1992) Plant and environmental sensory signals control the expression of *hrp* genes in *Pseudomonas syrinagae* pv. phaseolicola. J Bacteriol 174: 3499-3507.

Rajeshwari, R., Jha, G. and Sonti, R. V. (2005) Role of an *in planta* expressed xylanase of *Xanthomonas oryzae* pv. oryzae in promoting virulence on rice. Mol. Plant Microbe Interact. 18: 830-837.

Raven, J. A. (1984) Phytophages of xylem and phloem: a comparison of animal and plant sap-feeders. Adv. Ecol. Res. 13:135-234.

Ray, S. K., Rajeshwari, R., & Sonti, R. V. (2000) Mutants of *Xanthomonas oryzae* pv. oryzae deficient in general secretary pathway are virulence deficient and unable to secrete xylanase. Mol. Plant Microbe Interact. 13, 394-401.

Ray, S. K., Rajeshwari, R., Sharma, Y. and Sonti, R. V. (2002) A high-molecular-weight outer membrane protein of *Xanthomonas oryzae* pv. oryzae exhibits similarity to non-fimbrial adhesins of animal pathogenic bacteria and is required for optimum virulence. Mol Microbiol. 46: 637-47.

Read R. J. (2001) Pushing the boundaries of molecular replacement with maximum likelihood Acta Cryst. D57, 1373-1382.

Reddy, P. R. (1984) Kresek phase of bacterial blight of rice. Oryza. 21: 179-187.

Remaut, H. and Waksman, G. (2004) Structural biology of bacterial pathogenesis. Curr. Op. Struc. Biol. 14:161-170.

Rhodes, G. (2000) Crystallography Made Crystal Clear. San Diego: Academic Press.

Riess, T., Raddatz, G., Linke, D., Schäfer, A. and Kempf, V. A. (2007) Analysis of Bartonella adhesin A expression reveals differences between various *B. henselae* strains. Infect Immun. 75: 35-43.

Robinson, R. A. (1986) New concepts in breeding for disease resistance. Ann. Rev. Phytopathol. 18: 189-210.

Roggenkamp, A., Ackermann, N., Jacobi, C. A., Truelzsch, K., Hoffmann, H. and Heesemann, J. (2003) Molecular analysis of transport and oligomerization of the *Yersinia enterocolitica* adhesin YadA. J. Bacteriol. 185: 3735–3744.

Roggenkamp, A., Neuberger, H-R., Flugel, A., Schmoll, T. and Heesemann, J. (1995) Substitution of two histidine residues in YadA protein of *Yersinia enterocolitica* abrogates collagen binding, cell adherence and mouse virulence. Mol Microbiol 16: 1207-1219.

Romantschuk M. (1992) Attachment of plant pathogenic bacteria to plant surfaces. Ann. Rev. Phytopathol. 30: 225-243.

Ron, M., and Avni, A. (2004). The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. Plant Cell 10: 1604-1615.

Rossier, O., Wengelnik, K., Hahn K. and Bonas U. (1999) The Xanthomonas Hrp type III system secretes proteins from plant and mammalian bacterial pathogens. Proc. Natl. Acad. Sc. USA 96: 9368-9373

Ryan, C.A., and Farmer, E.E. (1991). Oligosaccharide signals in plants: A current assessment. Ann. Rev. Plant Physiol. Mol. Biol. 42:651-674.

Sacristan, S. and Garcia-Arenal, F. (2008) The evolution of virulence and pathogenicity in plant pathogen populations. Mol. Plant Pathol. 9: 369-84.

Saier, M. H. (2006) Protein secretion and membrane insertion systems in gram-negative bacteria. J. Membrane Biol. 214: 75-90.

Sali, A. & Blundell,T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234: 779-815.

Salnier, L. and Thibault, J.-F. (1999) Ferulic acid and diferulic acid as components of sugar-beet pectins and maize bran heteroxylans, J. Sci. Food. Agric. 79: 396-402.

Salzberg, S. L., Sommer, D. D., Schatz, M. C., Phillippy, A. M., Rabinowicz, P. D., et al.(2008) Genome sequence and rapid evolution of the rice pathogen *Xanthomonas oryzae* pv. oryzae PXO99A. BMC Genomics 9: 204.

Sauter, C., Basquin, J., Suck, D. (2003) Sm-like proteins in Eubacteria: The crystal structure of the Hfq protein from Escherichia coli . Nucleic Acids Res. 31: 4091-4098.

Schuster, M. L. and Coyne, D. P. (1974) Survival mechanisms of phytopathogenic bacteria. Ann. Rev. Phytopathol. 12: 199-221.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M. C. (2003) Swiss-Model: an automated protein homology-modeling server. Nucleic Acids Res. 31: 3381-3385.

Setubal, J. C., Moreira, L. M. and da Silva, A. R. (2005) Bacterial phytopathogens and genome science. Curr. Opinion Microbiol. 8: 595-600.

Shepherd, R. W. and Wagner, G. J. (2007) Phylloplane proteins: emerging defenses at the aerial frontline? Trends Plant Sc. 12: 51-56.

Shinohara, M., Nakajima, N., and Uehara, Y. (2007) Purification and characterization of a novel esterase (β-hydroxypalmitate methyl ester hydrolase) and prevention of the expression of virulence by *Ralstonia solanacearum*. J. Appl. Microbiol. 103: 152-162.

Showalter, A. M. (1993) Structure and function of plant cell wall proteins. Plant Cell 5: 9-23.

Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga. R., et al. (2000) The genome sequence of the plant pathogen *Xylella fastidiosa*. Nature 406: 151-157.

Smeltzer, M. S., Hart, M., and Iandolo, J. J. (1992) Quantitative spectrophotometric assay for staphylococcal lipase. Appl. Env. Microbiol. 58: 2815-2819.

Smyth, C. J., Marron, M. B., Twohig, J. and Smith, S. J. G. (2006) Fimbrial adhesins: similarities and variations in structure and biogenesis. FEMS Immunol. Med.l Microbiol. 16: 127-139.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., Osborne, E., et al., (2004) Toward a systems approach to understanding plant cell walls. Science 306: 2206-2211.

Stachel, S. E., Messens, E., van Montagu, M. and Zambryski, P. (1985) Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. Nature 318: 624–629.

Starr, M. P., and Stephens, W. L. (1964) Pigmentation and taxonomy of the genus Xanthomonas. Journal of bacteriology 87: 293-302.

Staskawicz, B. J., Mudgett, M. B., Dangl, J. L. and Galan, J. E. (2001) Common and contrasting themes of plant and animal diseases. Science 292: 2285-2289.

Stock, A. M., Robinson, V. L. and Goudreau, P. N. (2000) Two-component signal transduction. Ann. Rev. Biochem. 69: 183-215.

Storoni, L. C., McCoy, A. J. and Read, R. J. (2004) Likelihood-enhanced fast rotation functions Acta Cryst. D60: 432-438.

Storz, G., Opdyke, J. A. and Zhang, A. (2004) Controlling mRNA stability and translation with small, noncoding RNAs. Curr. Opin. Microbiol. 7: 140-144.

Struyve, M., Moons, M., and Tommassen, J. (1991) Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. J. Mol. Biol. 218: 141-148.

Sun, Q. H., Hu, J., Huang, G. X., Ge, C., Fang, R.X., and He, C. Z. (2005) Type-II secretion pathway structural gene *xpsE* xylanase and cellulase secretion and virulence in Xanthomonas *oryzae* pv. oryzae. Plant Pathol. 54: 15-21.

Surana, N. K., Cutter, D., Barenkamp, S. J. and St Geme, J. W. 3rd. (2004) The *Haemophilus influenzae* Hia autotransporter contains an unusually short trimeric translocator domain. J. Biol. Chem. 279: 14679-14685.

Swings, J., Van den Mooter, M., Vauterin, L., Hoste, B., Gillis, M., Mew, T.W. and Kersters, K. (1990) Reclassification of the causal agents of bacterial blight (*Xanthomonas campestris* pv. oryzae) and bacterial leaf streak (*Xanthomonas campestris* pv. oryzicola) of rice as pathovars of *Xanthomonas oryzae* (ex Ishiyama 1922) sp. nov., nom. rev. Intl. J. Systematic Bacteriol. 40: 309-311.

Szeltner, Z. and Polgar, L. (2008) Structure function and biological relevance of prolyl oligopeptidase. Curr. Prot. Pept. Sci 9: 96-107.

Tamm, A., Tarkkanen, A-M., Korhonen, T. K., Kuusela, P., Toivanen, P. and Skurnik, M. (1993) Hydrophobic domains affect the collagen-binding specificity and surface polymerization as well as the virulence potential of the YadA protein of *Yersinia enterocolitica.* Mol Microbiol 10: 995-1011.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24: 1596-1599.

Tang X., Xiao Y. and Zhou J. M. (2006) Regulation of the type III secretion system in phytopathogenic bacteria. Mol Plant Microbe Interact. 19:1159-66.

Taylor, G. (2003) The phase problem. Acta Crystallogr. D Biol. Crystallogr. 59: 1881-90.

Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension. Acta Cryst. D59: 38-44.

Terwilliger, T. C., and Berendzen, J (1999). Automated MAD and MIR structure solution. Acta Cryst. D55: 849-861.

Thieme, F., Koebnik, R., Bekel, T., Berger, C., Boch, J., et al. (2005) Insights into genome plasticity and pathogenicity of the plant pathogenic bacterium *Xanthomonas campestris* pv. vesicatoria revealed by the complete genome sequence. J Bacteriol 187: 7254–7266.

Thornton, J. M., Orengo, C. A., Todd, A. E., and Pearl, F. M. G. (1999) Protein folds functions and evolution. J. Mol. Biol. 293: 333-342.

Tseng, T., Tyler B. M. and Setubal J. C. (2009) Protein secretion systems in bacterial-host associations and their description in the Gene Ontology. BMC Microbiol. 9 (Suppl 1):S2.

Tsuchiya, K., Mew, T.W., and Wakimoto, S. (1982) Bacteriological and pathological characteristics of wild-type and induced mutants of Xanthomonas *campestris* pv. oryzae. Phytopathology, 72: 43-46.

Tsuge, S., Furutani, A., Fukunaka, R., Oku, T., Tsuno, K., Ochiai, H., Inoue, Y., Kaku, H. and Kubo, Y. (2002) Expression of *Xanthomonas oryzae* pv. oryzae *hrp* genes in XOM2, a novel synthetic medium. J. Gen. Plant Pathol. 68: 363-371.

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

van Gijsegem, F. (1997) In planta regulation of phytopathogenic bacteria virulence genes: relevance of plant-derived signals. Eur. J. Plant Pathol. 103: 291-301.

van Overbeek, L. S. and van Elsas, J. D. (1995) *Pseudomonas fluorescens* mutants in the wheat rhizosphere. Appl. Environ. Microbiol. 61: 890-898.

van Pouderoyen, G., Eggert, T., Jaeger, K., and Dijkstra, B. W. (2001) The crystal structure of *Bacillus subtilis* lipase: a minimal α/β hydrolase fold enzyme. J. Mol. Biol. 309: 215-226.

Van Valen, L. (1973) A new evolutionary law. Evol. Theory 1: 1-30.

Verger, R. (1997) Interfacial activation of lipases: facts and artifacts. Trends Biotech. 15: 32-38.

Vorholter, F. J., Schneiker, S., Goesmann, A., Krause, L., Bekel, T., et al. (2008) The genome of *Xanthomonas campestris* pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. J. Biotechnol. 134: 33–45.

Vorwerk, S., Somerville, S., and Somerville, C. (2004). The role of plant cell wall polysaccharide composition in disease resistance. Trends Plant Sci. 9: 203-209.

Wei, K., Tang, D.-J., He, Y.-Q., Feng, J.-X., Jiang, B.-L., Lu, G.-T., Chen, B., and Tang, J.-L. (2007) *hpaR*, a putative marR family transcriptional regulator, is positively controlled by HrpG and HrpX and involved in the pathogenesis, hypersensitive response and extracellular protease production of *Xanthomonas campestris* pv. campestris. J. Bacteriol. 189: 2055-2062.

Wei, Y., Contreras, J. A., Sheffield, P., Osterlund, T., Derewenda, U., Kneusel, R. E., et al., (1999) Crystal structure of brefeldin A esterase, a bacterial homolog of the mammalian hormone-sensitive lipase. Nat. Struct. Biol. 6: 340-345.

Wei, Z. M., Sneath, B. J. and Beer, S. V. (1992) Expression of *Erwinia amylovora hrp* genes in response to environmental stimuli. J. Bacteriol. 174: 1875-1882.

Wengelnik K., Ackerveken, G. and Bonas,U. (1996) HrpG, a key hrp regulatory protein of *Xanthomonas campestris* pv. *vesicatoria* is homologous to two-component response regulators. Mol Plant-Microbe Interact. 9: 704-712.

Wengelnik, K. and Bonas, U. (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the hrp cluster of *Xanthomonas campestris* pv. vesicatoria. J Bacteriol. 178: 3462-3469.

Wengelnik, K. and Bonas, U. (1996) HrpXv, an AraC-type regulator, activates expression of five of the six loci in the hrp cluster of *Xanthomonas campestris* pv. vesicatoria. J Bacteriol. 178: 3462–3469.

Whitehead, N. A., Byers, J. T., Commander, P., Corbett, M. J., Coulthurst, S. J., Everson, L., et al., (2002) The regulation of virulence in phytopathogenic *Erwinia* species: quorum sensing, antibiotics and ecological considerations. Antonie Van Leeuwenhoek. 81: 223-31.

Wilson, A. C., Carlson, S. S., White, T. J. (1977) Biochemical evolution. Ann. Rev. Biochem. 46: 573-639.

Wilson, M., Hirano, S. S. and Lindow, S. E. (1999) Location and survival of leaf-associated bacteria in relation to pathogenicity and potential for growth within the leaf. Appl. Environ. Microbiol. 65: 1435-1443.

Windgassen, M., Urban, A. and Jaeger, K. E. (2000) Rapid gene inactivation in *Pseudomonas aeruginosa*. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Lett. 193:201-205.

Yang, B., Sugio, A. and White, F. F. (2006) Os8N3 is a host disease-susceptibility gene for bacterial blight of rice. Proc Natl Acad Sci USA 103: 10503-10508.

Yeo, H., Cotter, S. E., Laarmann, S., Juehne, T., St.Geme III J. and Waksman G. (2004) Structural basis for host recognition by the *Haemophilus influenzae* Hia autotransporter. EMBO J. 23: 1245-1256.

Young, S. A., Guo, A., Guikema, J. A., White, F. F., and Leach, J. E. (1995) Rice cationic peroxidase accumulates in xylem vessels during incompatible interactions with *Xanthomonas oryzae* pv. oryzae. Plant Physiol. 107: 1333-1341.

Zhang Y. (2008) I-Tasser server for protein 3D structure prediction. BMC Bioinformatics, 9: 40-43.

Zhu, W., Macbanua, M. M., and White, F. F. (2000) Identification of two novel *hrp* associated genes in the *hrp* gene cluster of *Xanthomonas oryzae* pv. oryzae. J. Bacteriol. 182: 1844-1853.