PA-IIL as a Model Lectin for the Structural and Functional Characterisation of Related Lectins

Thesis submitted for the degree of
Doctor of Philosophy

by
Aileen Creavin, B.Sc.

Supervised by
Dr. Brendan O'Connor B.Sc., Ph.D
and
Dr. Michael O'Connell B.Sc., Ph.D

School of Biotechnology,
Dublin City University,
Dublin 9, Ireland

November 2009

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philoscophy is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: _____ (Candidate) ID No.: 51732625 Date: _____
Aileen Creavin

Acknowledgements

I'd like to take this opportunity to thank……

My family – for their endless support and encouragement, this thesis would not have been possible without them.

My supervisors, Mick and Brendan – for the opportunity to pursue this PhD and for your guidance, my deepest thanks.

To Paul and Roisin – for your enthusiasm, helping hand and for knowing sometimes that it really was my proteins fault.

To Susan and the coffee gang – I would never have finished this PhD without the kind words, friendly ear and the well timed injections of caffeine.

To my lab mates – it's been an experience (mostly good).

To Vinny – for helping me survive this experience with what little sanity I have left intact.

Abbreviations

3D	Three dimensional
a.a	Amino acid
AHL	Acylhomoserine lactones
b.p	Base pair
BCA	Bicinchoninic acid
Bcc	*Burkholderia cepacia* complex
BSA	Bovine serum albumin
CBL	Calcium binding loop
CF	Cystic fibrosis
DMSA	Dimethyl suphoxide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ELLA	Enzyme linked lectin assay
FAC	Frontal affinity chromatography
FPLC	Fast protein liquid chromatography
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LB	Luria bertani
Log	Logarithm
MCS	Multiple cloning site
MD	Molecular dynamic
MS	Mass spectrometry
MW	Molecular weight
NPF	p-Nitrophenyl-alpha-L-fucose
NTA	Nitrilotriacetic acid
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PDB	Protein data bank
PTM	Post-translation modification

Rpm	Revolutions per minute
SEC	Size exclusion chromatography
SDS	Sodium dodecyl sulphate
spp	Species
SPR	Surface plasmon resonance
TEMED	N,N,N,N'-tetramethyl ethylenediamine
Tris	Tris (hyroxymethyl) amino methane
tRNA	transfer ribonucleic acid
v/v	Volume per volume
w/v	Weight per volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Table of contents

List of Figures

Chapter 1

Chapter 4

Chapter 6

List of Tables

Chapter 1

Chpater 2

Chapter 3

Chapter 4

List of Equations

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Abstract

Glycosylation is one of the most abundant post-translational modifications of proteins and plays a diverse role in biological functions. The role of glycosylation in biological processes is a rapidly growing area of research. The growing development of glycoproteins as biopharmaceuticals has added to the interest in the characterisation of carbohydrate composition of glycoproteins as glycosylation-dependant therapeutic effects have been noted. Lectins are proteins of non-immune origin that recognise and bind specific carbohydrate structures with a relatively high degree of affinity. They represent a convenient biochemical tool to probe the carbohydrate composition of glycoproteins. Lectins occur ubiquitously in nature including in *Pseudomonas aeruginosa,* an opportunistic pathogen that produces two carbohydrate binding lectins, PA-IL and PA-IIL. The present study is dedicated to one of these lectins, the highly studied PA-IIL lectin (also known as LecB). A number of highly conserved homologous to PA-IIL have also been identified. PA-IIL was recombinantly expressed in *Escherichia coli* with the addition of a 6HIS affinity tag, separately to the N and C termini of the protein, to facilitate purification. Size exclusion chromatography established that the addition of an affinity tag to the C termini of PA-IIL affected the quaternary structure of the protein and was subsequently shown by HRP assay to affect the sugar binding. Optimized expression and purification protocols for the production of highly purified N-terminally tagged PA-IIL were then established. Once established the cloning, expression and purification protocols of PA-IIL were utilized to express the lectin domain of previously uncharacterized homologues, Photopexin A and Photopexin B from *Photorhabdus luminescens.* Once purified, methods for the structural and functional characterization of these novel lectins were investigated.

1.0 Introduction

1.1 Lectins

Lectins are proteins of non-immune origin capable of binding saccharide structures with a relatively high degree of specificity and affinity. Originally lectins were identified as proteins possessing the ability to agglutinate erythrocytes and other types of cells, which in later years was shown to be sugar specific binding. They were often termed Hemagglutinins because of this ability. The earliest identified lectin, at the turn of the 19th century, is believed to be ricin, isolated from seeds of the castor tree (*Ricinus communis*), presented by Peter Hermann Stillmark in his doctoral thesis in 1888. Hemagglutinins were reported throughout the following years from numerous sources, primarily plants, but very few were purified. The pace of lectin isolation increased significantly during the 1970s with the development of affinity chromatography for improved lectin isolation (Agrawal, Goldstein 1965). Interest in lectins was stimulated when lectins were demonstrated to be potentially valuable tools for (i) the detection and characterisation of glycoconjugates (Fisher *et al.* 1984) (ii) for the examination of changes that occur on cell surfaces during physiological and pathological processes (Aub, Sanford and Cote 1965) and (iii) for histochemistry of cells and tissues (Katsuyama, Spicer 1978). Although reported earlier, the first sugar specific mammalian lectin, the galactose specific hepatic asialoglycoprotein receptor was not isolated until 1974 (Ashwell, Morell 1974).

1.2 Lectins of *Pseudomonas aeruginosa*

Pseudomonas aeruginosa (*P. aeruginosa*) is a gram negative aerobic rod belonging to the bacterial family *Pseudomonadaceae.* It is a free living ubiquitous bacterium commonly found in soil and water, it also regularly occurs on the surface of plants and occasionally on the surface of animals. *P. aeruginosa* is of interest as it is an opportunistic human pathogen with intrinsic resistance to antibiotics and disinfectants. It is associated with chronic airway infections and is regarded as a primary cause of death in immunocompromised patients and those with cystic fibrosis. Extracts of certain strains of *P. aeruginosa* were found to contain D-galactose binding hemagglutinins (Gilboa-Garber 1972) and mannose binding hemagglutinins (Gilboa-Garber, Mizrahi and Garber 1977). These hemagglutinins were identified as lectins and named, respectively, PA-IL and PA-IIL (also known as LecA and LecB respectively). The

carbohydrate binding specificity of these lectins was determined using a variety of assays;

(i) Hemagglutination assay: Based on the lectins ability to agglutinate papin-treated erythrocytes. The precision of the assay is limited.

(ii) Peroxidase-binding assay: Based on the lectins ability to interact with both cells and the carbohydrate bearing enzyme horse radish peroxidase. This is a semi-quantitative assay (Huet, Bernadac 1974).

(iii) Mitogenic assay: A measurement of mitogenic stimulation in cultures of neuraminidase treated lymphocytes.

The preferential affinity for PA-IL was shown to be: α-D-Gal > D-Gal, β-D-Gal > D-GalNAc; and PA-IIL was shown to be, L-Fuc > D-Man (and mannan), L-Gal> D-Fru (Gilboa-Garber 1982). This present study will concentrate on one of these lectins, namely the PA-IIL and its highly conserved orthologs.

1.2.1 *PA-IIL* gene

In a study carried out by Gilboa-Garber, N. *et al.*, (2000) PA-IIL was purified by affinity chromatography using D-mannose bearing Sepharose (as outlined in Gilboa-Garber 1982) and its N-terminal amino acid sequence was obtained by Edman degradation. A T-blast search against the *P. aeruginosa* genome (PA01) coupled with genetic analysis identified a single open reading frame (ORF) encoding a 114-amino acid protein. The putative protein predicted from this ORF was found to perfectly match the query sequence at its N-terminus which was then further confirmed as being the ORF of PA-IIL by mass spectrometric analysis. This study found that PA-IIL is a 114 amino acid protein with a molecular weight of approximately 11.7 kDa. The C-terminus of the protein appears to be particularly hydrophobic while methionine is largely absent from the N-terminus. A Shine-Dalgarno sequence was located 9 b.p upstream of the translation start site (Gilboa-Garber, Katcoff and Garber 2000).

Identification of the *PA-IIL* gene facilitated the recombinant expression of PA-IIL. The protein was successfully cloned into a number of pET vectors and expressed in *E. coli* (DeB3) expression system in a number of studies (Loris *et al.* 2003, Sabin *et al.* 2006, Tielker *et al.* 2005). Protein was purified over mannose agarose as described by Gilboa-Garber, N. *et al.* (1977).

1.3 The crystal structure of PA-IIL

The PA-IIL lectin has been subject to four crystallographic studies (Loris *et al.* 2003, Mitchell *et al.* 2002, Sabin *et al.* 2006, Mitchell *et al.* 2005) and a molecular dynamic study (Mishra *et al.* 2008).These studies yielded several structures: native and calcium-free lectin and complexes with monosaccharides.

Mitchell, E. *et al.* (2002) were first to report the high resolution crystal structure of PA-IIL in a complex with fucose. Single wavelength anomalous diffraction (SAD) technique, with the protein crystal soaked in a holmium-containing solution, was used to determine the 3-Dimensional (3D) protein structure. The refined crystal structure of PA-IIL in (i) its carbohydrate-free form, (ii) as a complex with mannose, mannotriose, fucose and fructose**,** and (iii) in its inactive calcium-free form were reported by Loris, R. *et al.* (2003). A study by Mitchell, E. *et al.* (2005) further refined the crystal structure to a resolution of 1.0Å and compared structural results to thermodynamics to investigate the fine monosaccharide specificity of PA-IIL. Sabin, C. *et al.* (2006) elaborated further on the work of Mitchell, E. *et al.* (2005) by expansion of the study to include various monosaccharides at high resolution.

The PA-IIL crystal structure was found to consist of a tetramer of four independent subunits. Each subunit contains two Ca^{2+} ions and binds a fucose ligand (see Figure.1.2). PA-IIL was observed to fold as a nine-stranded anti-parallel β-sandwich with strands 1-5 forming a greek-key structural motif, schematics representing these protein formations can be seen in Figure 1.1.

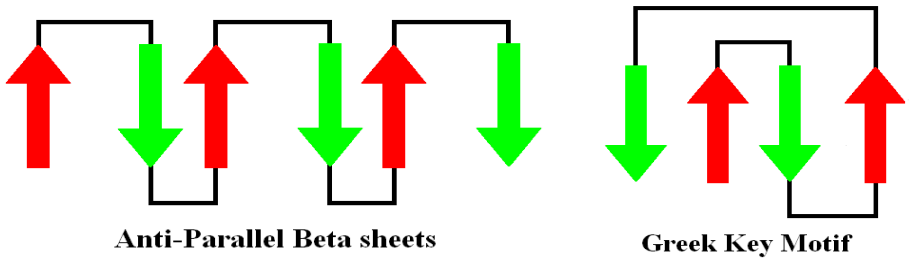


Figure 1.1: Schematic of protein folding.
The red and green arrows represent the β-sheets of the protein. The formations taken by anti-parallel beta sheets and anti-parallel sheets in a Greek key motif are represented in a simplified format above.

The greek-key motif of PA-IIL is extended by strands 6-8, which associate with strands 1 and 4 to form a 5-stranded curved β-sheet. The N-terminal region forms a β-strand, referred to as strand 0, that inserts itself between strand 5 and 6 (Figure 1.2). The PA-IIL tetramer forms a quaternary structure arranged around a pseudo C222 axis of

symmetry. The discovery of the involvement of the C-terminal carboxyl group (Gly 114) of each monomer in the ligand binding site of the neighbouring monomer is the most striking feature of this lectin (Mitchell *et al.* 2005). The C-terminal glycine group (Gly114) is involved in calcium co-ordination through hydrogen bonding and using experimental and *in silico* studies results suggest that the glycine also has a significant role in carbohydrate binding (Wimmerová *et al.* 2009).

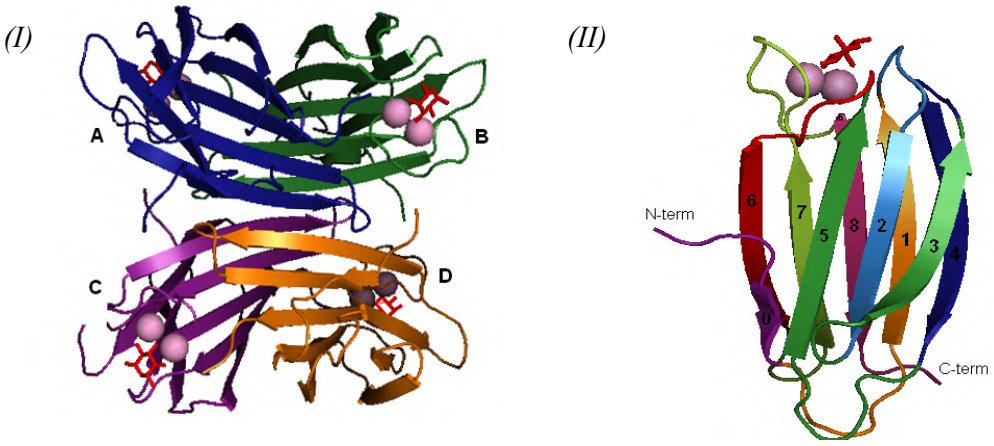


Figure 1.2: Overall structure of *P. aeruginosa* lectin, PA-IIL.
(I) Tetrameric structure of the PA-IIL-fucose complex, each monomer shown in a different colour. Stick representation of fucose (red) and two calcium ions shown as pink spheres. Each monomer is labelled A, B, C and D according to the nomenclature used in the text. *(II)* Monomer of PA-IIL with numbering of β-strands according to the greek-key motif (strands 1-5). Ca^{2+} and fucose shown as (I*)* (Mitchell *et al.* 2002). Images generated using Pymol and PDB file 1UZV.

Three intermolecular interfaces can be discerned in the PA-IIL tetramer. Where A, B, C and D represent each respective monomeric subunit of the tetramer (see Figure1.2), the interfaces are termed AB, AC and AD (Loris *et al.* 2003);

(i) AB interface: two sheets packed onto each other. Primarily a hydrophobic interface also includes two phenylalanine side chains and several valine and leucine side chains.

(ii) AD interface: continuous ten-stranded β sheet created *via* sheet extension, hydrophilic in nature.

(iii) AC interface: seen as an extension of the AD interface. Limited contacts between hydrophilic side chains of A and the C promoter, creates the AC interface.

1.3.1 Binding pocket of PA-IIL

The major feature of the carbohydrate-binding pocket is a pair of calcium ions. These calcium ions directly interact with three hydroxyl groups when a monosaccharide

is bound. This mode of protein/carbohydrate interaction had not previously been observed. In addition to the calcium ion interaction, several residues of PA-IIL interact directly with the bound monosaccharide *via* hydrogen bonds (see Figure1.3.); namely Ser22 (S22), Asp96 (D96), Asp99 (D99), Asp101 (D101) and Gly114 (G114) from the C-terminus of the neighbouring subunit (Mitchell *et al.* 2002).

The calcium binding pocket is made up of two calcium binding loops (CBL). CBL-1 connects strands 1 and 2 with CBL-2 connecting strands 7 and 8. The C-terminus of the adjacent monomer is also involved in the calcium binding loop. The calcium ions are bound *via* five acid groups found within the calcium binding loops. Both calcium ions have a classical seven ligand coordination. However there is distinct asymmetry in ligand distribution;

(i) Calcium 1 (as depicted in Figure 1.3): - the calcium ion receives five carboxylate and two oxygen ions.

(ii) Calcium 2 (as depicted in Figure 1.3): - the calcium ion receives three carboxylate, two carbonyl and two ligand oxygen atoms.

The structural role of the calcium ions was investigated by determining the crystal structure of calcium-free PA-IIL at 1.2Å resolution (Loris *et al.* 2003). Loris, R *et al.* (2003) reported that the overall structure of the monomer and the integrity of the tetramer were unaffected by the loss of the calcium ions. The only observed structural changes are limited to the metal/carbohydrate binding loop Asn95-Asp104 and the C-terminal residue Gly114. The loop becomes less well ordered most likely due to the repulsion of negatively charged side chains that are close together (Loris *et al.* 2003). It has been suggested that calcium ions have a regulatory function in some lectins (Ng, Weis 1998, Brewer, Brown and Koenig 1983)**,** as of yet there is no information on any possible regulatory function of the calcium dependence of PA-IIL.

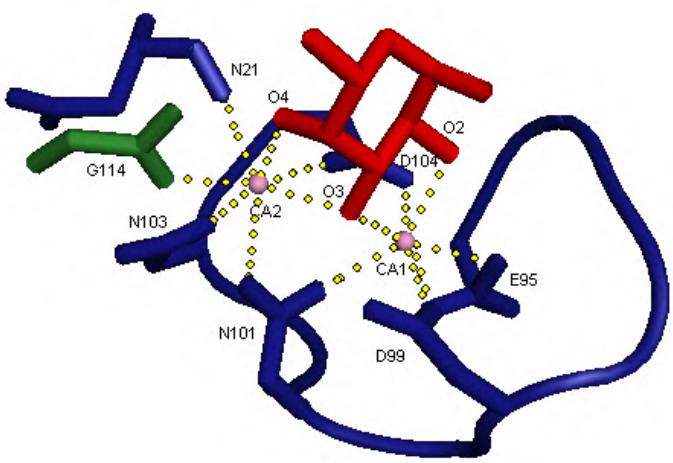


Figure 1.3: Stick representation of amino acids involved in interaction between PA-IIL calcium binding loop (blue), the C-terminus (green) of the adjacent subunit, the Ca^{2+} ions (pink spheres) and fucose (red).
Interaction with Ca^{2+} ions are shown as yellow dotted lines (Mitchell *et al.* 2002). Image generated using Pymol and PDB file 1UZV.

1.3.2 Fine specificity

While Mitchell, E. *et al.* (2002) were the first to report the quaternary structure of PA-IIL and the unique protein/carbohydrate binding mechanism, Loris, R. *et al.* (2003) investigated the fine specificity of PA-IIL to a number of monosaccharides (Figure 1.4).

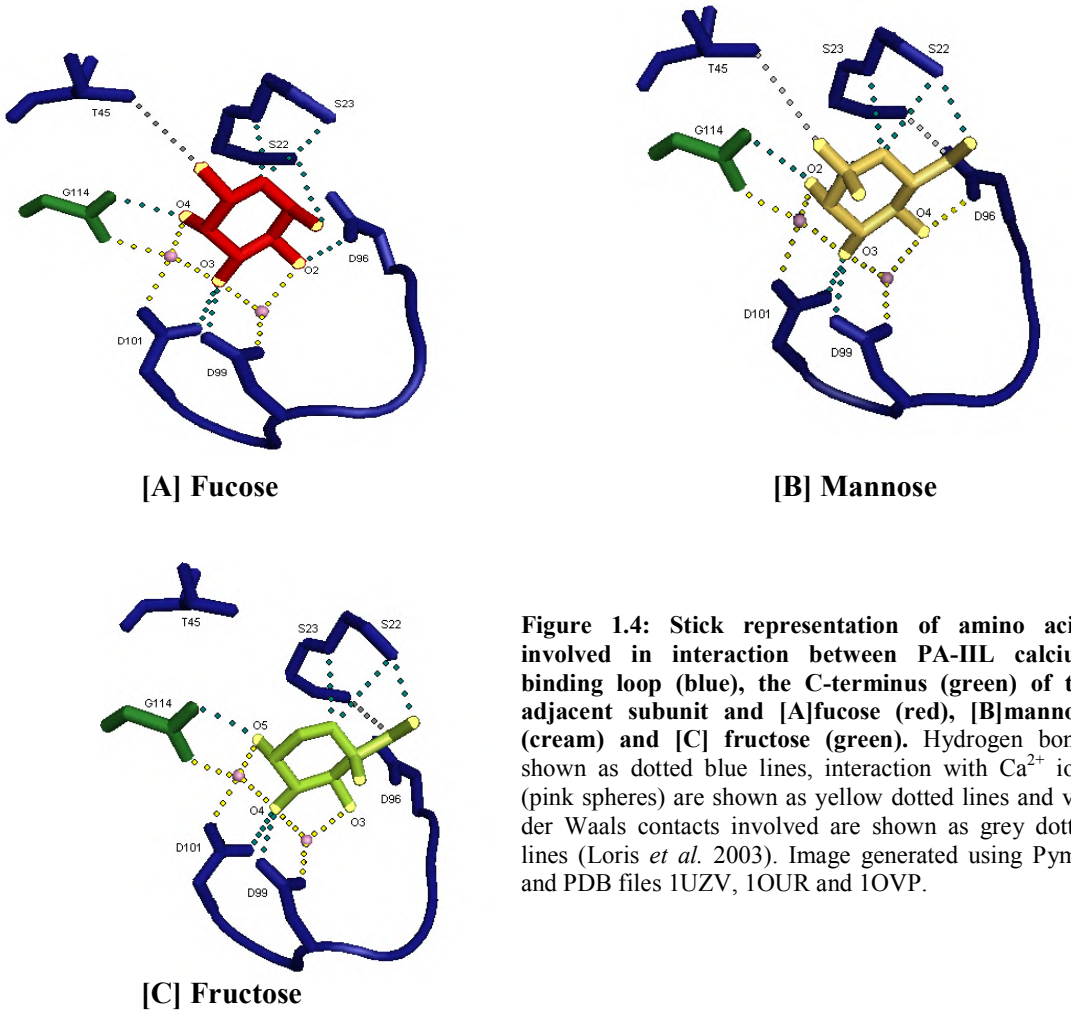


Figure 1.4: Stick representation of amino acids involved in interaction between PA-IIL calcium binding loop (blue), the C-terminus (green) of the adjacent subunit and [A]fucose (red), [B]mannose (cream) and [C] fructose (green). Hydrogen bonds shown as dotted blue lines, interaction with Ca^{2+} ions (pink spheres) are shown as yellow dotted lines and van der Waals contacts involved are shown as grey dotted lines (Loris *et al.* 2003). Image generated using Pymol and PDB files 1UZV, 1OUR and 1OVP.

Loris, R. *et al.* (2003) compared the crystal structure of the binding pocket of PA-IIL in complex with fucose, mannose and fructose. It was found that all three monosaccharides bind isosterically *via* identical interactions with the calcium ions and several side chains. Each of the monosaccharides are bound in their lowest-energy confirmation, namely, $^4C_1$ for fucose, $^4C_1$ for mannose and $^2C_6$ for fructose. In these confirmations, the hydroxyl O_2, O_3 and O_4 in fucose are sterically equivalent to O_4, O_3 and O_2 in mannose and O_3, O_4 and O_5 in fructose (Figure 1.4). The equivalent oxygen atoms form the direct ligands to the two structural calcium ions of PA-IIL. The remaining hydroxyl group and aliphatic substituents of the sugar groups determine the difference between the three structures (Loris *et al.* 2003). Differences were found around the C_6 methyl group of fucose, where a hydroxyl group that was found in the mannose complex was absent in the fructose complex. The C_6 methyl group of fucose forms favourable hydrophobic contacts with Thr45 and Ser23 as it is located within 4Å of these amino acids. Since no

equivalent hydroxyl or methyl group is found in fructose a small void is created which is not occupied by a water molecule. In mannose the O_1 in its β-configuration is the equivalent atom, resulting in a less favourable interaction. In fucose, O_1 in its α-configuration makes hydrogen bonds with Ser22 and Asp96. In mannose and fructose-protein complexes this interaction is replaced by the less favourable van der Waals interaction with C6 (mannose) or C1 (fructopyranose). In each instance mentioned in the mannose and fructose complexes a hydroxyl group is forced to interact with an aliphatic patch so that it cannot fulfil its hydrogen bond potential. Lastly an additional hydrogen bond is present in both the mannose and fructose-protein complexes, Ser23 to O_6 (mannose) and O_1 (fructose), for which there is no equivalent in the fucose-protein complex. The relatively weak substrate specificity associated with PA-IIL is therefore likely to be due to the requirement for the particular stereochemistry of two equatorial and one axial hydroxyl group leading to a number of sugars being able to bind (Loris *et al.* 2003).

PA-IIL exhibits strong affinity for fucose (Ka of 1.5 x $10^6 M^{-1}$) (Garber *et al.* 1987) and sparingly binds to α-D-mannose. The factors for the high affinity of PA-IIL are mostly likely due to the unique protein/carbohydrate bonding using Ca^{2+} ions. The high affinity was investigated by Mitchelle, E. *et al.* (2005) by (i) thermodynamic study of enthalpy and entropy contributions (ii) 1.0 Å resolution crystal structure and (iii) study of charge distribution by *ab initio* calculations.

Thermodynamic analysis confirmed the relatively high affinity binding of PA-IIL albeit with values of one order of magnitude lower (association constant (Ka) 1.6 x $10^6 M^{-1}$) then that previously published from equilibrium dialysis study (Garber *et al.* 1987). Variation of values may be due to differences between methodological approaches or between protein samples. There was a number of different conditions between the experimental approach of Mitchell, E. *et al.* (2005) and Garber, N. (1987), such as varying protein concentrations, 0.78mM compared to 1.66 – 3.1μM, differing experimental temperature, 25°C compared to 3°C and carried out in different buffering solutions, TBS pH 7.5 compared to PBS pH 7.3. This micromolar affinity range is very unusual among lectin-monosaccharide interactions. Lectin-protein interactions are usually observed in the millimolar range for binding unsubstituted monosaccharides (Dam, Brewer 2002). Previously, the highest reported value, K_a value of 9 x 10^3 M^{-1}, was for the GalNAc binder from soybean agglutinin (Gupta *et al.* 1996). The various thermodynamic contributions were analysed for PA-IIL/fucose and were characterized

by a strong enthalpy of binding, accompanied by a weakly favourable entropy term (Mitchell *et al.* 2005). When compared to classic lectin-monosaccharide interactions, the unusual affinity is explained by the relatively high enthalpy of binding and also by the absence of an entropy barrier. The favourable entropy of binding is very unusual but not unique in protein-carbohydrate complexes where an entropy barrier is generally observed (Dam, Brewer 2002). Favourable entropy has previously been observed for the binding of glucose by *Lens culinaris* lectin (Schwarz *et al.* 1993) and the binding of galactose by *Erythrina corallodendron* lectin (Surolia, Sharon and Schwarz 1996). Structural studies of *Erythrina corallodendron* indicate that favourable entropy may be due to the release of two tightly bound water molecules upon galactose binding (Elgavish, Shaanan 1998). A similar explanation could apply to the favourable entropy found for the PA-IIL/fucose interaction since the native lectin structure contains three water molecules tightly bound to the calcium ions that are released by fucose hydroxyl groups in complex (Loris *et al.* 2003).

The environment of the two close calcium ions was also analyzed by Mitchell, E. *et al.* (2005). There are three fucose hydroxyl groups directly involved in the coordination sphere of the calcium ions. Very strong hydrogen bounds are formed by two of these hydroxyl groups with acetate from protein side chains with an O-O distance of 2.55 (+/- 0.02) Å, a value significantly shorter then that currently observe in protein carbohydrate interactions. The acetates from protein side chains that are involved in the hydrogen bonds also coordinate the calcium ions. This forms two peculiar six-membered ring arrangements, which may be responsible for the strong charge delocalization shown by the *ab initio* calculations. These rings may therefore play a major role in the unusually high affinity observed between PA-IIL and the monosaccharides.

Sabin, C. *et al* (2006) expanded on this study by investigating PA-IIL complexes with a number of monosaccharides and their thermodynamics. In this study the thermodynamics and crystal structure of PA-IIL with Me-α-Fuc (as investigated by Mitchelle, E. *et al.* 2005), Me-α-L-Gal, Me-β-Ara and Me-α—Fru were investiagted. Methyl derivations of monosacchardies were used to avoid problems related to anomeric disorder in solution and also give a better potential mimic of real biological interactions. The order of preference of affinity of PA-IIL that was observed by inhibition of hemagglutination experiments (Garber *et al.* 1987) was confirmed by this study using isothermal titration calorimetry. Quantitatively, PA-IIL affinity for Me-α-

Fuc is threefold stronger then for Me-α-L-Gal and Me-β-Ara and 140- fold better than for Me-α-Man. All complexes formed are driven by enthalpy.

The role of the moiety at position 5 of the monosaccharide in binding was highlighted by Loris, R. *et al.* (2003). In this study comparison of the crystal structures allow the role of the moiety at position 5 of the ring to be explored. Since all other contact points between protein-monosaccharide complexes are identical in the different complexes, e.g. PA-IIL –Me-α-Fuc, PA-IIL- Me-α-L-Gal etc., the difference in binding affinity must therefore be correlated with this particular group. Interactions between proteins and carbohydrate generally present a balance of hydrogen bonds and hydrophobic contacts. In contrast, the interaction of PA-IIL with fucose is established mainly through coordination of calcium ions and hydrogen bonds. The only established hydrophobic interaction is between the methyl group at C5 of fucose and the side chain of Thr45. When this interaction is weakened, as in the complex with L-Gal, or absent, as with Me-β-Ara, the affinity is decreased by 3 fold (Sabin *et al.* 2006).

All factors explaining the relatively high affinity of PA-IIL have yet to be fully resolved. Mishra, N. *et al.* (2008) attempted to unravel what causes the PA-IIL high binding affinities by using molecular dynamics simulations (MD). MD has been shown to serve as a reliable computer simulation method to complement experimental observations.

Comparison of several MD trajectories of PA-IIL lectin either free or complexed with L-fucose, Me-D-Ara, D-mannose and D-fructose led to the observation that the ligand binding loop distance is shorter in the bound lectin compared with the free lectin. This indicates that the loops come closer after the binding of a monosaccharide. The geometry of the binding site in PA-IIL complexes show a stabilization effect of the monosaccharide site upon binding. The results suggest that the monosaccharide keeps the loops closer by the polar hydroxyl group interactions with the binding site *via* calcium ions. It has been demonstrated that hydroxyl group orientation of the monosaccharide and the ions within the vicinity of the binding site provide important means for the lectin binding (Mishra *et al.* 2008).

The root mean square deviations (RMSDs) fits from the MD simulations indicate that the water molecule network around the binding pocket in complex seems to be one of the important factors influencing the proper orientation of the sugar. There is one more water molecule present in the fucose complex compared with other complexes which increases the binding affinity. The additional oxygen contributes to

the higher binding energy compared with the other saccharides. Calcium ion simulation analysis proved that the ions presence and its coordination with protein side chains and the monosaccharide are important driving factors for the lectin/carbohydrate binding (Mishra *et al.* 2008).

To date, using 3-D crystal structures and molecular dynamic simulations, a number of factors have been associated with the relatively high affinity binding of PA-IIL. The protein crystal structures revealed the unusual protein/carbohydrate binding interaction formed by the two Ca^{2+} ions, the amino acid side chain that co-ordinates the Ca^{2+} ions and the sugar, as a basis for the relatively strong binding affinity. The moieties at position 5 and 6 of the monosaccharide ring were also demonstrated to affect the binding affinity. The relative contributions of thermodynamic parameters were also investigated. A strong entropy barrier is usually observed for classical protein/carbohydrate interactions. This strong entropy barrier is absent for PA-IIL due to the release of water molecules, which are found more strongly bound in PA-IIL then classical carbohydrate sites, upon binding. Favourably enthalpy and the lack of a strong entropy barrier are believed to contribute to the high affinity of PA-IIL. Finally molecular dynamic simulations indicate the role played by water molecules and confirm the role played by calcium ions in the binding pocket and their contribution to PA-IIL's higher binding energy.

1.4 Biological role of PA-IIL

1.4.1 Regulation of PA-IIL

P. aeruginosa produces a wide spectrum of virulence factors and secondary metabolites which include such products as alkaline protease, elastase, endotoxin A, LasA protease, phospholipase C, exoenzymeS, pyocyanin and hydrogen cyanide (Nicas, Iglewski 1985, Preston *et al.* 1995, Saiman *et al.* 1992, Hirakata *et al.* 1995). The expression of multiple virulence and survival genes in *P. aeruginosa* is cell density dependant and relies on a cell-cell communication system which is known as quorum sensing (Passador *et al.* 1995). Quorum sensing is the phenomenon whereby the accumulation of a low molecular weight, diffusible, signal molecule (also known as a pheromone or autoinducer) enables individual bacterial cells to sense when the minimum population unit or "quorum" of bacteria has been achieved for a concerted

population response to be initiated (Fuqua, Winans and Greenberg 1994). Quorum sensing is an example of multicellular behaviour and regulates a variety of physiological processes, such as bioluminescence, swarming, antibiotic biosynthesis, plasmid conjugal transfer and production of virulence determinants in fish, animal and plant pathogens (Hardman, Stewart and Williams 1998). Quorum sensing, due to the large number of regulatory systems associated with it, appears to constitute a global regulatory system in *P. aeruginosa*. Whitely, M. *et al.* (1999) have proposed that up to 4% of *P. aeruginosa* genes are regulated by quorum-sensing. *N*-acylhomoserine lactones (AHLs) have been reported to be employed as quorum sensing molecules in *P. aeruginosa*. *P. aeruginosa* has a sophisticated regulatory hierarchy linking quorum-sensing to virulence and survival in the stationary phase (Latifi *et al.* 1996, Pesci *et al.* 1997). A quorum-sensing circuit is composed of an AHL synthase enzyme and a sensor-regulator protein. The AHL synthase enzyme acts on a substrate to produce an AHL which then acts on a sensor regulator protein that modulates gene transcription. Two separate quorum-sensing circuits, termed the *las* and *rhl* systems, are found within *P. aeruginosa*. Each possesses an AHL synthase enzyme, *Las*I or *Rhl*I and a sensor regulator protein, LasR or RhlR. These modulate gene transcription in response to increasing AHL concentrations (las AHL: *N*-(3-oxododecanoyl)-L-homoserine lactone (3O-C12-HS2) and rhl AHL; *N*-butanoyl-L-homoserine lactone (C4-HSL))(Brint, Ohman 1995, Gambello, Iglewski 1991, Gambello, Kaye and Iglewski 1993, Latifi *et al.* 1995, Ochsner *et al.* 1994, Ochsner, Reiser 1995, Williams *et al.* 1992).

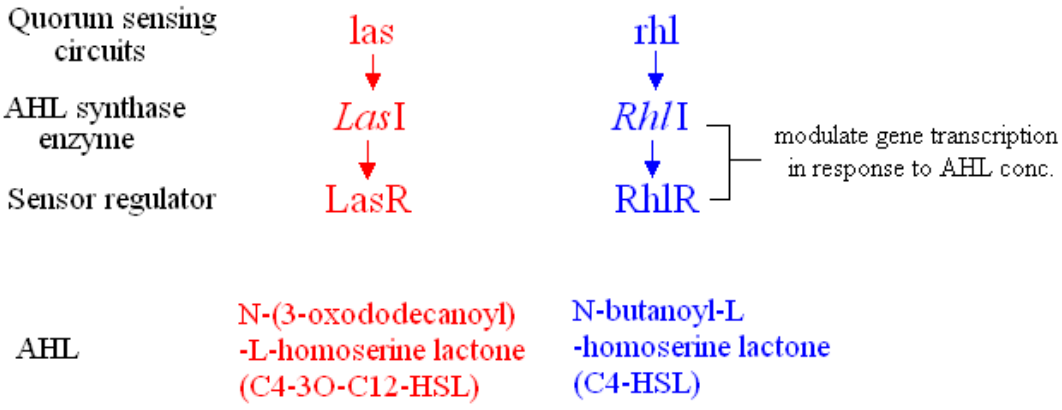


Figure 1.5: Quorum-sensing systems found in *P. aeruginosa* and AHL they respond to.

The two systems form a hierarchical cascade that coordinate virulence factors and stationary phase gene production via the alternative sigma factor RpoS (Latifi *et al.*

1996, Pesci *et al.* 1997). Each system regulates a regulon comprising of an overlapping set of genes. In *P. aeruginosa* the las system has been shown to directly regulate the rhl system. Therefore the las system provides overall coordination of quorum sensing and temporal gene expression in response to cell-to-cell communication (Hassett *et al.* 1999, Whiteley, Lee and Greenberg 1999).

Both lectins of *P. aeruginosa*, PA-IL and PA-IIL, are produced at high cell density, during the transition to stationary phase. In order to assess if quorum-sensing is involved in the production of these lectins, Winzer, K. *et al* (2000) carried out a study on PA-IL and PA-IIL production in various mutated strains of *P. aeruginosa* (Table 1.1).

Table 1.1: Strains of *P. aeruginosa* used in the study by Winzer, K. *et al.* (2000) and the effect mutation had on PA-IIL expression.

Strain		PA-IIL expression	Ref
P. aeruginosa			
PAO1	Wild type	Stationary phase production	Holloway collection
PANO67	NTG mutant derived from PAO1	Abolished	(Jones *et al.* 1993)
PAOR	*lasR* mutant derived from PAO1	Late stationary phase production	(Latifi *et al.* 1996)
PDO100	*rhlI* mutant derived from PAO1	Abolished	(Brint, Ohman 1995)
PDO111	*rhlR* mutant derived from PAO1	Abolished	(Brint, Ohman 1995)
PAO1 rpoS negative	*rpoS* mutant derived from PAO1	Abolished	(Jørgensen *et al.* 1999)

PA-IL and PA-IIL production was abolished in the *P. aeruginosa* rpoS mutant. LasRI and rhlRI quorum sensing have previously been shown to regulate the stationary phase sigma factor RpoS (Latifi *et al.* 1996). This study indicates that PA-IIL expression is RpoS dependant. Strain PANO67 lacks the ability to produce many virulence factors and does not produce PA-IL or PA-IIL. When the *rhlRI* locus was restored by transformation, growth phase dependant PA-IL and PA-IIL production was restored. This suggests the *rhl* locus is required for lectin biosynthesis. Further correlation was

found by the studying of *rhl*I negative and rhlR negative strains, PDO100 and PDO111 (Brint, Ohman 1995). In both strains lectin synthesis was abolished completely (see Table 1.1). The mutation of the lasR locus delayed but did not abolish lectin production. The LasR mutant strain's (PAOR) production of lectin was severally down regulated such that lectin production could only be detected in the late stationary phase. This suggests that RhlR/C4-HSL rather that LasR/3O-C12-HSL is directly responsible for controlling lectin expression. This result also suggests that while the las system regulates the rhl system, in late stationary phase, the rhlRI can be activated independently from LasR/3O-C12-HSL such that loss by mutation of lasR only delays but does not abolishe lectin biosynthesis. This study failed to exclude the possibility that lectin synthesis was not directly activated by the quorum sensing circuitry but by other, as yet unidentified, regulators which are themselves controlled by quorum-sensing (Winzer *et al.* 2000).

For LuxR-type regulators such as LasR and RhlR, *lux* box like elements constitute the DNA binding site. The presence of a putative *lux* box upstream of PA-IL therefore suggests that it is directly regulated by RhlR. This *lux*-box like element is located 42 b.p upstream of the transcription start codon. Very similar distances have been reported for the *lux* box elements upstream of *rhl*AB (Pearson, Pesci and Iglewski 1997), *las*B (OPl) (Rust, Pesci and Iglewski 1996) as well as elements upstream of *lux*I in *Vitrio fischeri* (Egland, Greenberg 1999). The focus of the Winzer, K. *et al.* (2000) study was the lectin PA-IL as the PA-IIL gene was yet to be identified. Gilboa-Garber, N. identified and characterized the PA-IIL lectin gene in 2000. Despite the close relationship between PA-IL and PA-IIL their genes are widely separated on the *P. aeruginosa* chromosome (approx. 867.5 k.b). Upstream of the PA-IIL translational start site three sequences with homology to the *lux* box, originally described in *V. fisheri* (Devine, Shadel and Baldwin 1989, Rust, Pesci and Iglewski 1996*)* and similar to that found for the PA-IL gene, were identified (Gilboa-Garber, Katcoff and Garber 2000).

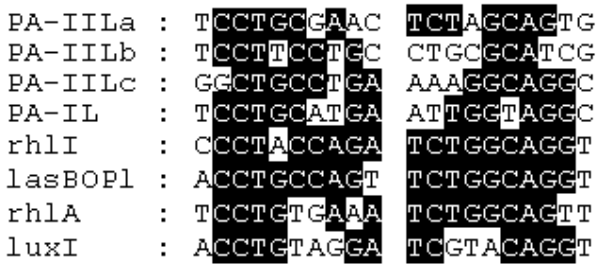


Figure 1.6: The three nucleotide sequences (a,b,c) in the upstream region of the *PA-IIL* gene resembling a lux box and compared to that of *PA-IL*, the lux box of *V. fischeri* (Gilboa-Garber, Katcoff and Garber 2000)**, and other genes of *P. aeruginosa*.** Image generated using ClustalW align and Genedoc. Identical residues highlighted in black.

Since PA-IL expression can be induced by AHLs and it has been shown that AHL of rhl locus effects production of PA-IIL (Winzer *et al.* 2000) it is possible these sequences may be involved in the PA-IIL autoinduction process. The relatively wide distance between the *PA-IL* and *PA-IIL* genes on the chromosome strongly indicates that the coordinate expression of both lectins may be only attributed to the common autoinduction process via the *lux*-like box (Williams *et al.* 1992).

1.4.2 Location of PA-IIL

Glick, J. *et al.* (1983) found both *P. aeruginosa* lectins to be located mainly in the cytoplasm of planktonic (single cells in suspension) cells. This finding makes it difficult to explain lectin mediated cytotoxic and adhesive properties (Adam, Schumacher and Schumacher 1997, Bajolet-Laudinat *et al.* 1994). However, the addition of fucose-branched chitosan was found to lead to specific cell aggregation which would suggest that PA-IIL may be exposed on the surface of sessile *P. aeruginosa* cells (Morimoto *et al.* 2001, Tielker *et al.* 2005). These researchers reported on the cellular localization in planktonic and sessile cells (cells not free to move about, i.e. attached to surfaces, involved in biofilm formation) and the role of PA-IIL in biofilm formation.

The biofilm-forming capability of wild type *P. aeruginosa* PAO1 and PA-IIL deficient strain PAI2 were compared over growth periods of 48 and 72 hours at both 30^{o}C and 37^{o}C. Biofilm formation was monitored by confocal laser scanning microscopy (CLSM). Biofilms of the mutant strain were found to be thinner then that of the wild type and the surface coverage was also found to be lower at both temperatures. Since the same growth rates and maximal cell densities were reached for PAO1 and PAI2 this effect was not caused by a general growth defect (Tielker *et al.* 2005). The

subcellular localisation of PA-IIL was determined by comparison of planktonic and sessile cells which had been grown in unsaturated biofilms of the type that is also found in the lungs of cystic fibrosis (CF) patients suffering from *P. aeruginosa* infection (Lyczak, Cannon and Pier 2002). Cellular compartments were isolated and probed with PA-IIL specific antiserum. PA-IIL was detected in the cytoplasm and total membrane fraction in both the planktonic and sessile cells. The precise location was further investigated by separation of the inner and outer bacterial membrane of sessile cells. PA-IIL was found exclusively in the bacterial outer membrane. The presence of PA-IIL on the outer membrane was suggested to be due to the lectin being bound via fucose-containing structures. This assumption was proven when PA-IIL was detected in the supernatant solution after the isolated outer-membrane was incubated with inhibitory sugar α-NPF (Garber *et al.* 1987). When incubated with galactose, for which PA-IIL has a low affinity, no lectin was found in the supernatant, i.e. it remains bound to membrane and found within the cell pellet. This result was confirmed by mutating PA-IIL which rendered it unable to bind (Tielker *et al.* 2005). In this case PA-IIL can be found in the cytoplasm similar to wild type PA-IIL but, unlike the wild type, it is not detected in the membrane fraction. Therefore the sugar binding capacity of PA-IIL is essential for outer membrane localization. Treatment of wild type PAO1 biofilms with yellow fluorescing protein (YFP) labelled PA-IIL demonstrated the binding of PA-IIL to the cells showing that carbohydrate receptors on the surface of *P. aeruginosa* biofilm cells are accessible to PA-IIL (Tielker *et al.* 2005).

This study by Tielker, D. *et al.* (2005) proves that PA-IIL plays an important role in biofilm formation and that PA-IIL is associated to the outer bacterial membrane *via* binding to carbohydrate-ligands. PA-IIL does not contain any of the presently known secretion signals and therefore is unlikely to be secreted *via* typeI pathway or the Sec or Tat pathways. As yet a secretion mechanism for PA-IIL has yet to be identified.

Since PA-IIL was implicated in biofilm formation it holds potential as a target for the prevention of *P. aeruginosa* biofilm formation. A previous study exploited the multivalency of lectins to produce a high affinity synthetic ligand, an octavalent dendrimer with galactose. This multivalent dendrimer was shown to inhibit hemagglutination induced by the bacterium *Streptococcus suis* revealing the potential of glycol-dendrimers as anti adhesions (Joosten *et al.* 2004). This precedent suggests that a multivalent fucose ligand should be a suitable compound to inhibit *P. aeruginosa* lectins.

The use of glycans as anti-adhesive agents can be found in nature. Immature animals (newborns, newly hatched vertebrates or larvae) are very sensitive to bacterial infection. In mammals infection is prevented by maternal milks (Newburg, Ruiz-Palacios and Morrow 2005). The majority of the oligosaccharides found in human milk are fucosylated and their production is dependant on enzymes which are encoded by the genes associated with the expression of the Lewis Blood groups (Newburg, Ruiz-Palacios and Morrow 2005). Studies have shown that complex oligosaccharides with terminal fucose residues of the Lewis a (Le^a) or Le^x series are the preferred targets for PA-IIL. The Le^a active pentasaccharide was found to have 120X higher activity then D-mannose (Mitchell *et al.* 2002). The order of preference for PA-IIL including Lewis groups was determined to be; Le^a active pentasaccharide ≥ sialy Le^a tetra- > methyl-α-Fuc > Fuc and Fucα1-2-Gal (human blood group H active disaccharide) > 2-fucosyllatose Le^x > H type I determinant > Le^x glycolope (Galβ1-4[Fucα1-3]GlaNAc > sialy Le^x tri- >>Man >>>Gal, GalNAc and Glc (Wu *et al.* 2006). It is unsurprising then that PA-IIL was found to display a unique high sensitivity to the fucosylated compounds of human milk (Lesman-Movshovich, Gilboa-Garber 2003). The maternal secretions contain saccharides in free form or carried on macromolecules (mainly glycoprotein) that act as decoys, competitively blocking the bacterial lectin mediated adhesion to cells. Avian embryos are protected by avian egg whites in the same manner. Glycans in quail eggs were shown to bear terminal mannose residues which are found to inhibit PA-IIL (Lerrer, Gilboa-Garber 2001).

Royal jelly and honey have been identified as other sources of inhibiting PA-IIL glycoproteins. When compared to the inhibition activity of human milk, honey and royal jelly would provide similar protection against PA-IIL mediated *P. aeruginosa* infection. The effectiveness of synthetic ligands to inhibit bacterium and the presence of fucosylated ligands in nature as decoys and competitive inhibitors suggest synthetic PA-IIL ligands may act as anti-microbial agents.

Johansson, E. *et al.* (2008) developed a library of multivalent fucosyl peptide dendrimers as ligands. The library was screened by functional selection. The ligand FD2 dendrimer (Figure 1.7), consisting of a structure with four fucosyl groups, was found to have the highest affinity toward PA-IIL (Figure 1.7, see Table 1.2 for chemical structure).

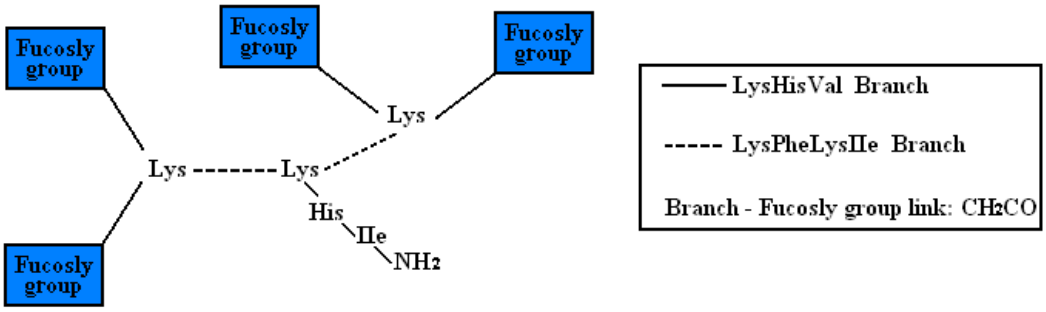


Figure 1.7: Schematic of the multivalent fucosyl-peptide dendrimer FD2

The effect of monovalent control ligands, such as fucose, on reduction of biofilm formation is weak. Conversely the addition of multivalent dendrimers to culture caused a strong reduction, in particular the high affinity dendrimer FD2, which showed complete inhibition of biofilm formation at 50μM. The effect of dendrimers on both biofilm inhibition and dispersion assays was much stronger then the effect observed with monovalent α-NPF. The observed effect was not due to potential toxicity of the dendrimers as bacterial growth was unaffected. FD2 was found to not only inhibit biofilm formation but in fact caused complete clearance of *P. aeruginosa* in dispersion assays. It was also found to inhibit biofilm formation in several clinical isolates derived from CF lungs, suggesting that the inhibition effect is general and supports the potential of these dendrimers in a clinical context.

Biofilm inhibition was proven to be due to inhibition of PA-IIL binding when it was shown that the monovalent control ligand α-NPF inhibited biofilm formation of wild type *P. aeruginosa* PAO1 and deletion mutation PAO1ΔPA-IL, but showed no effect on biofilm formation of deletion mutation PAO1ΔPA-IIL. Similar to α-NPF, no effect was seen on PAO1ΔPA-IIL biofilm formation by the dendrimer FD2. These results support the hypothesis that the effect of FD2 on biofilm formation is due to specific binding to PA-IIL (Johansson *et al.* 2008).

This study clearly suggests that multivalent fucosyl-peptide dendrimers may be used as ligands to prevent biofilm formation and disrupt existing biofilms of *P. aeruginosa*. The dendrimer created by this group was found to have a 46-fold higher affinity over fucose (estimated by isothermal titration calorimetry). Thus, these glycopeptide dendrimers represent strong multivalent PA-IIL ligands. The IC_{50} of FD2 is approx. 10μM for biofilm inhibition, a 50-fold affinity gain over α-NPF. FD2 was additionally tested for cytotoxicity against a human cell line by lactate dehydrogenase

(LDH) cytotoxicity assay where it did not show a significant increase in LDH release up to 0.1M concentration (Johansson *et al.* 2008).

Table 1.2: Synthetic yields of c-fucosyl peptide dendrimers produced by Johansson, E. (2008) and Kolomiets, E. (2009). [a] Number of fucosyl groups in the ligand [b] isolated by preparative HPLC [c] calculated molecular weight [d] mass obtained by MS ES(+).

Compound	n[a]	Sequence	Mass [mg][b]	Yield[b]	Mw[c]	MS EC(+)[d]
PA8	4	(cFuc-GlyThrVal)4(LysHisProThr) 2LysHisIleNH2	22	9	3104.4	3103.8
PA9	4	(cFuc-GluHisTyr)4(LysTyrGlyAsp) 2LysHisIleNH2	22	6	3791.6	3792
FD2*	4	(cFuc-LysProLeu)4(lysPheLyslle) 2LysHisIleNH2	38.5	14	3536.34	3536
FD10	4	(cFuc-GluThrAsp)4(LysTyrGlyAla) 2LysHisIleNH2	8	5	3369.38	3368
2G3	8	(cFuc-LysPro)8(LysLeuPhe)4(LysLyslle) 2LysHisIleNH2	11.7	2.5	5994.53	5994.25

Kolomiets, E. *et al.* (2009) further chemically refined the high affinity dendrimer FD2 to the 3rd generation octavalent dendrimer 2G3 (Table 1.2), with an IC_{50} value of 0.025μM, determined by enzyme linked lectin assay (ELLA). This N-terminal C-fucosyl residue represents a potent inhibitor for PA-IIL with a significantly increased potency up to 440-fold over fucose. The aqueous solubility of these fucosyl peptide dendrimers is generally good in contrast to the other synthetic multivalent fucosyl dendrimers reported to date where aqueous solubility is often low. This improves their applicability in a clinical setting. A number of observations with regard to the binding of PA-IIL were also made in this study. Although the best trivalent dendrimer FD2 is cationic, the neutral dendrimer, PA8 (Table 1.2), and anionic dendrimers, FD10 and PA9 (Table 1.2) bound almost as well as FD2 indicating that charge is not a primary determinant of the binding affinity to PA-IIL. Octavalent dendrimers bound better then the mono- or di- C-fucosyl moieties proving multivalency is an important factor for binding. Docking studies show that the C-fucosyl groups of the dendrimers adopt identical binding modes to fucose with peptide arms protruding form the binding pocket. It is these arms that establish specific contacts with the lectin (Kolomiets *et al.* 2009).

The suggestion that PA-IIL plays a role in bacterial adhesion to airway epithelia and in biofilm formation (Tielker *et al.* 2005), features attributed to surface exposure of PA-IIL, lead to the hypothesis that the lectins of *P. aeruginosa* exert some effect on

surface exposed proteins. Sonaware, A. *et al.* (2006) investigated the principle of this hypothesis.

A PA-IIL deficient mutated strain of *P. aeruginosa* PAK was created by insertional inactivation, whereby a gene is inactivated by the insertion of an unrelated fragment of DNA into the cloning sequence of the gene. Since all hypothesised actions of PA-IIL are surface related, e.g. biofilm formation and adhesion, any change in membrane protein levels was initially examined. Image analysis of 2D-gels revealed that a protein spot, identified as PilJ, was down regulated for the PA-IIL mutant. Analysis by MS indicate that there was a six fold reduction of this protein expressed in the PA-IIL mutant strain compared to the wild type strain while the PA-IL mutant strain was unaffected. The down regulation of PilJ was assessed to see if expression was effected at the transcriptional level. No significant difference was noted in transcriptional activity which suggests that the effect on expression of PilJ took place at the post-transcriptional level. No lectin was detected in the wild type membrane preparations used for the proteomic studies. This suggests that the lectins may be expressed at very low levels or are not found in significant amounts in the outer membrane. No carbohydrates were found attached to wild type PilJ protein which would imply that PA-IIL does not directly interact with the protein (Sonawane, Jyot and Ramphal 2006).

Twitching motility is mediated by type IV pili and is controlled by a complex chemosensory pathway which includes a number of proteins, PilJ being one of these (Whitchurch *et al.* 2004). Previously, it has been observed in a *P. aeruginosa* PilJ mutant, twitching motility was deficient and no type IV pili were expressed (Kearns, Robinson and Shimkets 2001). The twitching motility of PAK wild type was found to be normal while a PA-IIL mutant was found to be non-twitching. Twitching motility was restored when PA-IIL was complemented back into the mutant indicating that the loss of PA-IIL was responsible for loss of twitching motility. The twitching defect was also found not to be due to transcriptional change in protein synthesis but is most likely due to an inability to assemble surface pili (Sonawane, Jyot and Ramphal 2006).

Biofilm formation was compared between the PA-IIL mutant, complemented PA-IIL mutant and wild type *P. aeruginosa*. The PA-IIL was more then 50% defective in biofilm formation when compared to the wild type. Biofilm formation was fully restored for the PA-IIL compliment indicating the loss of PA-IIL was responsible for defective biofilm formation. It was proposed that the defect in biofilm formation may be

due to the absence of pili in the mutant. The PA-IIL mutant was found to be defective in type IV secretion, therefore the other secretion systems were examined for defects (Sonawane, Jyot and Ramphal 2006).

The type II secretory pathway is used to excrete a significant number of proteins in *P. aeruginosa* including alkaline protease, endotoxin A, Las A and Las B (Ball *et al.* 2002, Liu 1966, Pavlovskis, Gordon 1972). The PA-IIL mutant was found to have weak proteolytic activity when grown on casein-milk agar plates. When levels of LasB were examined both the intracellular and secreted LasB were found to be unaltered for the PA-IIL mutant. This suggests that the reduced caseinolytic activity is not due to a LasB defect. The deficit in caseinolytic activity is presumed to be due to modulation of protease IV activity after experiments showed that although secretion and size of LasB in the PA-IIL mutant was unaffected, no lactoferrin degradation by protease IV in the presence of EDTA was detected. In addition analysis showed that many extracellular proteins were absent from the PA-IIL mutant. However, phospholipase and endotoxin A secretion were unaffected. No effect was seen on the secretion of ExoS and ExoT which are secreted through the type II secretion system (Yahr, Goranson and Frank, Frank 1997).

Lectins have been suggested to be involved in adhesion to respiratory mucins in CF patients (Mitchell *et al.* 2002) but this hypothesis has yet to be demonstrated. The *P. aeruginosa* mutant lacking pili PAK-NP (Saiman *et al.* 1990) attaches to mucins as efficiently as the wild type strain (Ramphal *et al.* 1991). The PAK-NP mutant lacking functional *fliD* (PAK-NPD), which encodes the flagellar cap protein, was severely impaired in mucin adhesion (Arora *et al.* 1998). These strains were used as controls. Both lectin mutants were unimpaired in their ability to bind to CF mucin compared to the wild type strain. The PAK-NPD mutant did show a marked defect in its ability to bind to the CF mucin. This result conflicts somewhat with results obtained by Chemani, C. *et al.* (2009), who found reduced bacterial loads *in vivo* in a murine model with the loss of the PA-IIL gene.

Respiratory mucins contain a number of sugars including fucose, therefore, one would predict that if there is surface expression of PA-IIL then mucin binding would be a prominent function. This wasn't demonstrated in this study. The study concluded that PA-IIL is involved in multiple functions. It effects expression of PilJ, a member of the pilus biogenesis pathway which in turn affects pilus formation. Its role in pili generation

would explain its role in biofilm formation in vitro. This study also demonstrates a role for PA-IIL in proteolytic activity (Sonawane, Jyot and Ramphal 2006).

Chemani, C. *et al.* (2009) carried out a study to determine if lectins may contribute to the virulence of *P. aeruginosa*. The results demonstrated that there is a relationship between lectins and *P. aeruginosa* pathogenicity. PAO1::*PA-IL* and PAO1::*PA-IIL* mutant strains of wild type *P. aeruginosa*, PAO1, were produced by insertion of a transposon for interruption into the *PA-IL* and *PA-IIL* genes. Growth curves for the parental strain and mutants showed that there was no major growth defect for PAO1::PA-IL and PAO1::PA-IIL mutants. The effect of these mutations was examined under a number of conditions. The cytotoxicity of *P. aeruginosa* PAO1 and mutations was evaluated by measuring the LDH release at 4 and 6 hours after infection of A549 cells. When compared, the amount of LDH released was statistically larger for PAO1 infected cells then that for the mutated strains at 4 and 6 hours. PAO1 induced cytotoxicity was significantly reduced by the addition of specific inhibitors (GalNAc and Me-α-Gal for PA-IL and Me-α-Fuc for PA-IIL) after incubation for 6 hours compared to results for the non-specific lectin inhibitor Glc. Cytotoxicity obtained with purified lectin was also significantly decreased when specific inhibitors were added. Adherence of bacteria to A549 cells was also examined. The number of adherent bacteria to cells was approximately twofold greater for PAO1 infected cells then in comparison to PAO1::*PA-IL* and PAO1::*PA-IIL*. *In vitro* analysis in a murine model found that PAO1::*PA-IL* and PAO1::*PA-IIL* did not have a statistically significant affect on mortality rates (Chemani *et al.* 2009). Incubation of *P. aeruginosa* with specific PA-IL inhibitors did decrease mortality. However no difference was observed for the group which received the PA-IIL inhibitor, Me-α-Fuc, over the PAO1 infected group. Lung injury was also evaluated *in vitro*. Lung injury was measured by the efflux of the tracer protein ^{125}I-albumin from blood to lungs at 6 hours after intraperitoneal instillation or from blood into the lungs 16 hours after intraperitoneal injection. The concentration of the protein tracer was statistically greater for PAO1 then either mutated strain after both 6 hours and 16 hours. The concentration of protein tracer could be restored by trans-complementation of the gene into the mutated strain. Therefore the loss of PA-IL and PA-IIL can be said to reduce protein permeability of the lungs. Likewise purified lectins induced an increase in permeability. Co-instillation of specific inhibitors with purified PA-IIL significantly reduced lung injury when assed by alveolar barrier permeability. Bacterial load of lungs 16 hours after instillation of pathogen was also compared

between strains. Animals infected with PAO1 showed a significantly higher bacterial load then PAO1::*PA-IL* and PAO1::*PA-IIL*. Co-administration of lectin specific inhibitors with the pathogen reduced bacterial load. Blood culture analysis, after instillation of *P. aeruginosa,* found the number of positive blood cultures statistically lower for PAO1::*PA-IL* and PAO1::*PA-IIL* then PAO1. Co-instillation of lectin specific inhibitors decreased dissemination of bacteria in blood (Chemani *et al.* 2009). The parent strain PAO1 *P. aeruginosa* was found to have greater cytotoxicity then either of the mutants. Although cytotoxicity may be due to multiple factors the cytotoxic effect of purified lectins shows that the cytotoxicity is at least in part a result of the lectins. The *in vivo* relevance of PA-IL and PA-IIL suggests that the lectins are not direct or major determinants of mortality on their own but could act as important co-factors none the less.

Oligosaccharide-mediated recognition and adhesion are key stages in the early phase of *P. aeruginosa* infection. The reduced bacterial lung loads for PAO1::*PA-IL* and PAO1::*PA-IIL* would support the hypothesis that lectins act as virulence factors mainly in the early phase. Although PA-IIL may not play a direct role in adhesion as reported by Sonaware, A. *et al.* (2006) it does appear to be an indirect factor in bacterial adherence. Since PA-IL and PA-IIL have been found to play a major role in biofilm formation, modulation of PA-IL and PA-IIL could potentially inhibit initial infection of *P. aeruginosa* and then subsequent biofilm formation at a latter phase as demonstrated by the reported successful treatment of a *P. aeruginosa* airway infection (von Bismarck, Schneppenheim and Schumacher 2001)

1.6 Homologs of PA-IIL

The availability of the whole genome sequence of an ever growing number of bacteria has led to the identification of a number of homologs to PA-IIL (Table 1.3). Comparative studies of the molecular basis, properties and function reveal two highly homologous proteins to PA-IIL.

Ralstonia Solanacearum is a plant pathogen found in soil. It infects more than 200 plant species from 50 different families causing heavy agricultural crop loss (Hayward 1991, Denny 2000). Homology scans using the 114 amino acid sequence of PA-IIL as a query identified a 342 b.p open reading frame encoding a PA-IIL like

putative protein. This 113 a.a protein, RS-IIL, displays 69% identity and 85% similarity to PA-IIL (Sudakevitz *et al.* 2004).

Chromobacterium violaceum is a common inhabitant of soil and water in tropical and subtropical regions. It behaves as a saprophyte generally but can sporadically become an aggressive opportunistic animal pathogen causing serious infections with high mortality rates in immunodeficient human patients (Alves de Brito *et al.* 2004, Sirinavin *et al.* 2005). Comparison of the *C. violaceum* genome (Brazilian National Genome Project Consortium. 2003) with other organisms reveal *R. solanacearum* as having the highest percentage of ORF similarity (17.4%) followed by *P. aeruginosa* with 9.61%. A homologous ORF to PA-IIL and RS-IIL was found and named CV-IIL (Zinger-Yosovich *et al.* 2006).

	RS-IIL	CV-IIL	ppxA/L	ppxB/L
PA-IIL	69%	60%	31%	26%
	85%	80%	58%	57%
	0	0	2	1

Table 1.3: Homologous lectins; *P. aeruginosa* PA-IIL, *R. solanacearum* RS-IIL, *C. violaceum* CV-IIL, lectin region *P. luminescens* Photopexin A (ppxA/L) and lectin region *P. luminescens* PhotopexinB (ppxB/L). Percentage identity (black) refers to the percentage of amino aicds in the alignment of the above lectins which are identical between the proteins. Percentage similarity (red) refers to the percentage of amino acids in the alignment which are not identical between the proteins but which have similar biochemical properties and are considered interchangeable in many situations. Gap character (blue) is the number of padding or null characters to match homologous residues, indicated by – in the alignments. Statistic report created using ClustalW and GenDoc.

Sudakevitz, D. *et al.* (2004) carried out crystal structure analysis of RS-IIL. The topology and binding pocket of RS-IIL and PA-IIL were found to be very similar. Two calcium ions in a classical seven ligand co-ordination were found in the binding pocket as was found for PA-IIL. This is corroborated by the complete inhibition of RS-IIL in the presence of the metal chelator EDTA (as found for PA-IIL), as EDTA strips Ca^{2+} from the binding pocket. The C-terminus, Gly113, of a neighbouring monomer was found to be involved in the calcium and monosaccharide binding, another characteristic of PA-IIL binding. All the amino acids of the calcium binding site are strictly conserved between PA-IIL and RS-IIL (Figure 1.8) (Sudakevitz *et al.* 2004).

RS-IIL and PA-IIL were found to be inhibited by similar sugars in hemagglutinating inhibition assays but differed in their binding affinity ranges. The

order for RS-IIL was found to be, D-fructose and D-mannose, L-fucose, L-galactose and D-arabinose. The affinity range for RS-IIL to fructose and mannose is comparable to that of PA-IIL to fucose. The conserved calcium binding site is the basis of the high affinity as it allows three hydroxyl groups to be involved in the co-ordination sphere (as seen for PA-IIL, Mitchell *et al.* 2002). From the crystal studies of PA-IIL interacting with different ligands (Loris *et al.* 2003, Mitchell *et al.* 2002) it is known that specificity for monosaccharides is only partly dependant on the stereochemical arrangement required for interacting with the Ca^{2+} as this was shown to be met by other monosaccharides. The specificity difference between RS-IIL and PA-IIL is therefore directed by the second loop, the sugar binding loop. The Ser-22-Ser-23-Gly-24 in PA-IIL is replaced by Ala-22-Ala-23-Asn-24 in RS-IIL (Figure 1.8). Binding of Ser-22 to Asp-96 in PA-IIL results in a lack of space in the binding site for the O6 of mannose which is then directed towards the surface where it interacts with Ser-23. In RS-IIL the Asp-96 is available for hydrogen binding. RS-IIL preference for mannose is the result of these stronger hydrogen bonds between O6 and the acidic group of Asp-96 when compared with the weaker bond between O6 and the hydroxyl group of Ser-23 in PA-IIL (Sudakevitz *et al.* 2004). Small amino acid variations are therefore responsible for the differing binding affinities of PA-IIL and RS-IIL.

The crystal structure of CV-IIL in complex with Me-α-fuc and Me-α-man was produced by Pokorná, M. *et al.* (2006). The quaternary structure was found to be a tetramer very similar to those previously described for PA-IIL and RS-IIL (Sudakevitz *et al.* 2004). The binding mode of CV-IIL was also found to be very similar to both these lectins. The binding complex involves two Ca^{2+} ions and one sugar residue in each binding site. All the amino acids involved in calcium binding are strictly conserved among PA-IIL, RS-IIL and CV-IIL (see Figure 1.8) and the Ca^{2+} ions are bound in the same manner. As for PA-IIL and RS-IIL the acidic group of the C-terminal Gly113 of the neighbouring associated monomer is involved in the binding pocket. The preference order for CV-IIL was reported to be; L-fucose>D-arabinose>D-fructose>D-mannose. The preference of CV-IIL for fucose is only 6 fold stronger than mannose. Therefore this lectin is a fucose/mannose binding lectin, i.e. binding affinity appears somewhere between the PA-IIL and RS-IIL affinities. The amino acids of the sugar binding loop of CV-IIL, crucial for specificity, lies between those found for PA-IIL and RS-IIL (Figure 1.8) in terms of degree of sequence symmetry which correlates to the binding specificity (Pokorná *et al.* 2006). The variations in the sugar binding loop are; Ser22-Ser23-Gly24

in PA-IIL to Ser22-Ala23-Ala24 in CV-IIL (Zinger-Yosovich *et al.* 2006).The lower affinity binding of CV-IIL to fucose may be due to differences in thermodynamics. However at present it is difficult to correlate the thermodynamics and structural differences for binding Me-α-Fuc by CV-IIL and PA-IIL (Pokorná *et al.* 2006). Further studies are required to resolve this.

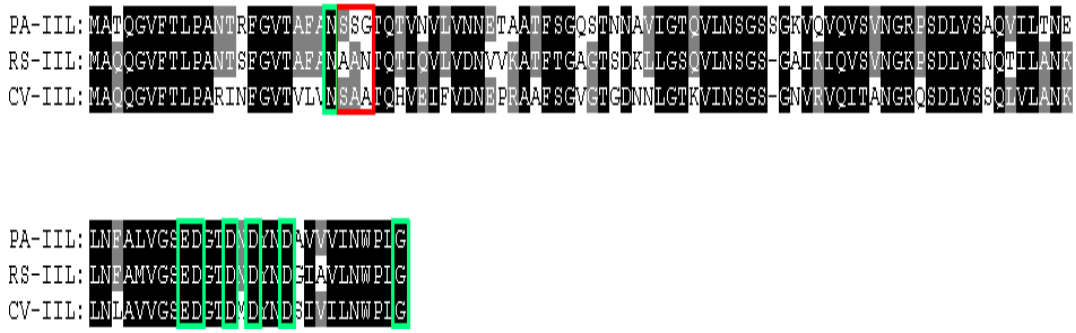


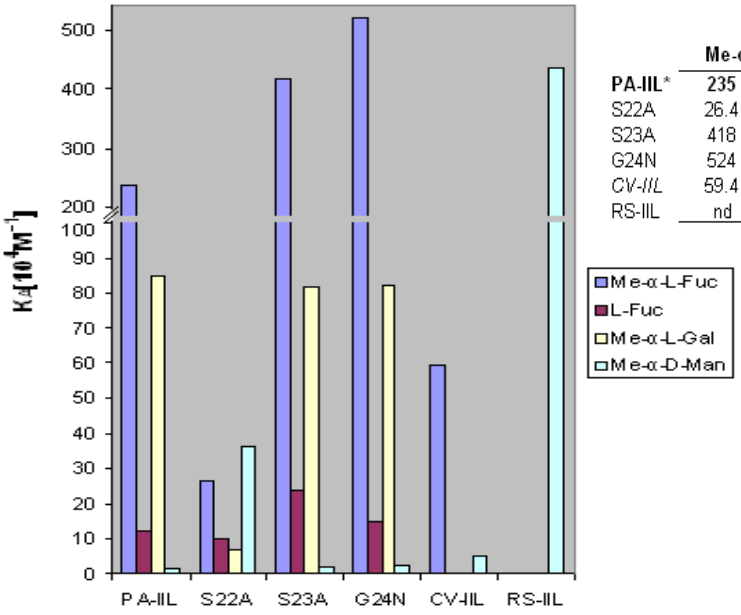
Figure 1.8: Amino acid alignment of protein sequences of *P. aeruginosa* lectin PA-IIL, *R. solanacearum* lectin RS-IIL and *C.violaceum* lectin CV-IIL.
Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey**.**

The structural comparison of PA-IIL and RS-IIL revealed the importance of the three amino acid motif for specificity (see above). A study by Adam, J. *et al.* (2007) confirmed this observation by constructing mutants replacing each of the amino acids in position 22, 23 and 24 of PA-IIL from *P. aeruginosa* PAO1. The point mutations were investigated both structurally by X-ray crystallography and functionally by isothermal titration calorimetry (ITC). The point mutations were as follows; Ser22→Ala22 (S22A), Ser23 →Ala23 (S23A) and Gly24→Asn24 (G24N).

The mutant S22A displays a reversed order of preference between mannose and fucose. The affinity of S22A to mannose was found to be 4 times higher compared to the wild type protein (Table 1.4., Figure 1.9.). This result confirms the hypothesis that S22 holds the main key to affinity determination. Without the hydroxyl group of S22 there is no steric hindrance blocking the interaction of the sugar with Asp96 (Adam *et al.* 2007). The Ser22A/Me-α-man complex closely resembles the crystal structure of RS-IIL lectin/mannoside complex (Sudakevitz *et al.* 2004).

The order of sugar preference for S23A and G24N is the same as that for wild type PA-IIL (fucose over mannose preference). The affinity of both mutants to Me-α-fuc is two fold higher then the wild type protein (Figure 1.9). Increases in affinity are due to slight conformational changes of the binding pocket leading to favourable

affinities (Adam *et al.* 2007). This study demonstrates the potential of mutagenesis of the sugar binding loop to determine sugar preference of lectins.



	KA[x104M-1] Values			
	Me-α-L-Fuc	L-Fuc	Me-α-L-Gal	Me-α-D-Man
PA-IIL*	**235 (±8.0)**	**12 (±1.0)**	**85 (±5.0)**	**1.42 (±0.05)**
S22A	26.4 (±4.9)	10.3 (±0.4)	6.75 (±0.31)	36 (±1.8)
S23A	418 (±72.8)	23.8 (±3.4)	81.6 (±14.4)	1.91 (±1.91)
G24N	524 (±36.7)	15.2 (±2.3)	82.2 (±18.8)	2.34 (±0.17)
CV-IIL	59.4 (±0.2)	nd	nd	5.4 (±0.6)
RS-IIL	nd	nd	nd	438 (±2.0)

Table 1.4 [Above]: Thermodynamics of lectin/sugar binding. Comparison of thermodynamic binding characteristics for different monosaccharides with PA-IIL, its mutants, RS-IIL from *R. solanacearum* and CV-IIL from *C. violaceum*. Determined by ITC at 25^0C. PA-IIL values obtained from Sabin, C *et al.* (2006) and Perret, S *et al.* (2005). Values obtained for CV-IIL taken from Pokorná, M. *et al.* (2006).

Figure 1.9 [Left]: Graphical representation of the binding affinities of the *P. aeruginosa* PA-IIL lectin and mutants, *C. violaceum* lectin CV-IIL and *R. Solanacearum* lectin RS-IIL, as determined by ITC.
Representation of the equilibrium association constants toward different monosaccharides for wild type PA-IIL, PA-LL mutants and the homologs RS-IIL from *R. solanacearum* and CV-IIL from *C. violaceum*.

Homologs to PA-IIL have also been located within the *Burkholderia cepacia* complex (Bcc). The Bcc is a collection of genetically distinct but phenotypically similar bacteria that can be divided into at least 9 species (Mahenthiralingam, Vandamme 2005). Bcc bacteria such as *B. cenocepacia* and *B. multivorans* are opportunistic pathogens that can cause lung infections in CF patients often resulting in death (Mahenthiralingam, Baldwin and Vandamme 2002). Bcc bacteria are found in a variety of environments like *P. aeruginosa* (e.g. soil, water and vegetation) and exchange of genetic material has been confirmed between both bacteria (Eberl, Tümmler 2004). It is therefore not surprising that several genes coding for lectins or lectin-like domains with similarity to PA-IIL have been identified in a number of strains of Bcc. Frequently an uncharacterised N-terminal domain is attached to the lectin domain. The genes identified as coding PA-IIL like lectins all belong to Bcc that were reported in C.F isolates (Lameignere *et al.* 2008). No single copy gene has been identified, all have coded for larger proteins then PA-IIL

Lameignere, E. *et al.* (2008) carried out a study on the lectin BclA (*B. cenocepia* lectin A), the product of the smallest PA-IIL-like gene from *B. cenocepia* J2315. The BclA gene was identified in five stains of *B. cenocepia* and one strain of *B. ambifaria.* All *Burkholderia* BclA protein sequences are very similar, with a 32% identity with PA-IIL. The calcium binding loop and C-terminal glycine amino acids are conserved between all the genes. The amino acids in the sugar binding loop are partially conserved; Ser22-Ser23-Gly24 of PA-IIL to Ala22-Ala23-Glu24 of BclA. The presence of alanine (22) in place of serine (22) for RS-IIL in the 1st position of the sugar binding loop trio is implicated in the higher binding affinity for mannose over fucose. BclA is therefore predicted to display a higher preference for mannose binding.

The Lameignere, E. *et al.* (2008) study cloned and expressed recombinant BclA. A crystal study of BclA revealed the topologies of BclA and PA-IIL to be similar including the presence of two Ca^{2+} ions in the binding pocket and the action of the C-terminal glycine of the neighbouring subunit in activity. Differences in the peptide sequence influence the tertiary arrangement, in particular the addition of N-terminal amino acids and a longer loop which would not allow the creation of a dimer-dimer interace. Therefore, BclA forms a dimer while PA-IIL forms a tetramer. Surface plasmon resonance (SPR) assays (based on the inhibiting power of a monosaccharide on binding of circulating BclA to a mannose surface) determined D-mannose to be an excellent inhibitor. L-fucose, D-arabinose and L-galactose also inhibit but two orders of magnitude weaker then D-mannose. D-galactose did not bind at all. These results were confirmed by broad spectrum glycan array analysis performed at the Consortium for Functional Genomics. These results show BclA has a strict specificity for oligomannose-type N-glycans, only binding to oligosaccharides with a α-mannoside-capped branch. Specificity for linkage does not appear to be strict (Lameignere *et al.* 2008).

Alignment of 19 of the highest similarity Bcc genes to PA-IIL (identified by protein Blast search) found 5 amino acid variations within the sugar binding loop. An example of each variation aligned with PA-IIL, RS-IIL and CV-IIL and the conserved calcium binding regions is shown in Figure 1.10. While the N-terminus and size of the protein varies the calcium binding amino acids remain largely conserved (Figure 1.10).

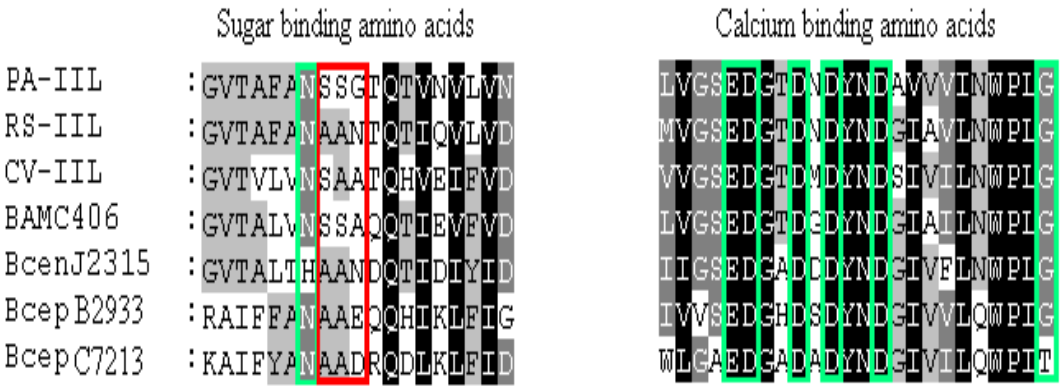


Figure 1.10: PA-IIL – fucose lectin *P. aeruginosa,* RS-IIL – mannose lectin *R. Solanacearum*, CV-IIL – fucose lectin *C. violaceum*, BC-IIL – fucose lectin II *B. cenocepacia* AU1054, Bcep1 – Hypothetical protein Bcep18194-B2934 *B. cenocepacia* spp.383, Bcep2 – Hypothetical protein Bcep18194-B2933 *B. cenocepacia* spp 383.
Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey**.**

Little variation of the sugar binding amino acids of Bcc bacteria indicates the most of these lectin-like domains are likely to fucose/mannose binders.

Homologs of greater diversity have also been identified in *Photorhabdus luminescens*. *P. luminescens* is an insect pathogenic bacterium that lives in symbiosis with nematodes that invade insects. On invasion of an insect host by a nematode the bacteria is released from the nematode gut and is thought to secrete a range of virulence factors that over come the insect while providing a suitable environment for both the bacteria and nematode. Among the virulence factors it produces are 'toxic complexes' which are encoded by several *toxic complex* (tc) loci or pathogenicity islands (Bowen *et al.* 1998). Surrounding these tc loci are many open reading frames. To formulate hypotheses for the putative biological role of such ORFs, work was carried out to model the structures on predicted amino acid sequences with conserved motifs (Crennell 2000). The locus discussed in the study by Crennell, S. *et al.* (2000) lie upstream of the *tcaABC2* locus of *P. luminescens* W14 (Figure 1.11) and was termed *Photopexin*.

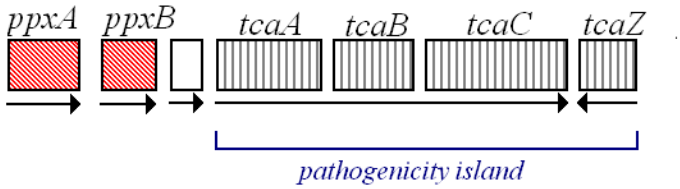


Figure 1.11: The Photopexin locus is found upstream of the toxic complex (tc) locus. PpxA and ppxB locus highlighted in red, remaining (grey) locus genes involved in toxic complexes.

This toxic complex and upstream ORFs have also been found for *P. luminescens* strain TT01. The locus has two ORFs, namely *Photopexin A (ppxA)* and *Photopexin B* (*ppxB)*. ppxA is the larger of the two, consisting of 564 amino acids to ppxB with 340 amino acids. ppxB has a 65% identity to ppxA. Analysis revealed ppxA to contain internal homology within itself. ppxA is composed of two 220 amino acid homologous domains and a less similar 100 amino acid C-terminus. The predicted ppxB sequence is homologous to the latter two domains of ppxA (one repeating domain and the C-terminal domain). ppxB could therefore be hypothesised to have a similar function to ppxA or may be a tandem (potentially non-functional) duplication of part of ppxA. The N-terminal domains of both ppxA and ppxB have significant protein sequence similarity to the phosphocholine binding protein limunectin from *Limulus amoebocytes*. Lesser similarities were found to a wide range of matrix metalloproteases (MMPs) and collagenases (Crennell 2000).

Crennell S. *et al.* (2000) carried out a number of modelling studies of the N-terminal domain of ppxA and ppxB. A four-blade β-propeller fold was assigned to these proteins (Figure 1.12).

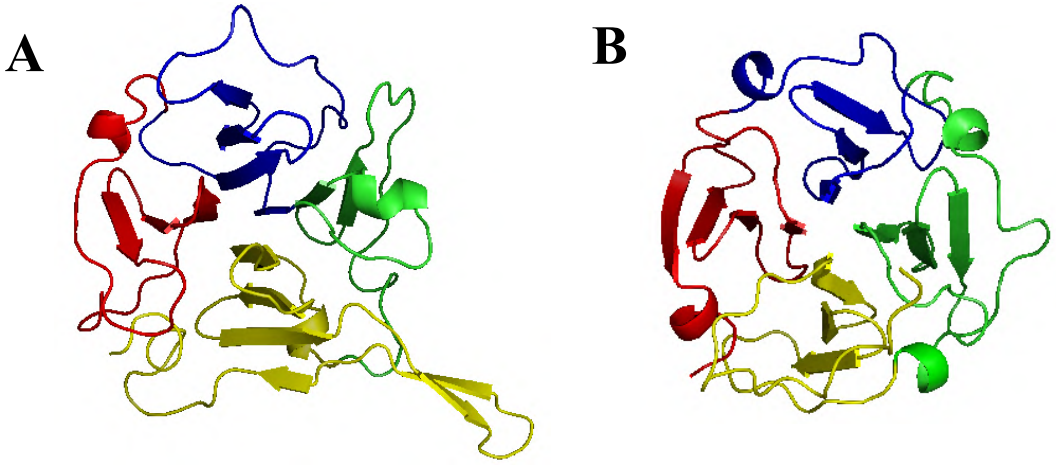


Figure 1.12: [A] Ribbon diagram of the N-terminal domain of ppxA, modelled on haemopexin-like domain of human matrix metalloproteinase-2 (MMP2), PDB file 1RTG. [B] Ribbon diagram of the four bladed β-topology of MMP2, PDB file 1RTG.
The perpendicular view illustrates the four bladed β-topology. ppxA protein model created using Swiss-Model and PDB file 1RTG as template. Images generated using Pymol.

This fold was first seen in the structure of the vertebrate haem-scavenging molecule haemopexin. This is the first 'haemopexin-like' motif identified outside of animal and plants. It was therefore speculated that in *Photorhabdus* ppxA and ppxB may be used for host surface attachment or for binding of haem like iron containing compounds

scavenged from the insect host. No predicted structure was assigned to the C-terminal domains of ppxA and ppxB in this study as a PSI-Blast search found no significantly similar sequences and was therefore excluded from the modelling study. The role and structure of these open reading frames are as yet unproven (Crennell 2000).

With the identification of the PA-IIL gene (Gilboa-Garber, Katcoff and Garber 2000) the C-terminal domain of ppxA and ppxB were found to have a significant similarity to PA-IIL (Table 1.3), in particular the main calcium-binding loop is highly conserved across all three proteins (Figure 1.13). A third protein (Q937N3) of 210 amino acids was identified from *Photorhabdus* which also displayed a lectin-like domain with homology to PA-IIL. The N-terminal domain does not display similarity with any protein of known function (Imberty *et al.* 2004).

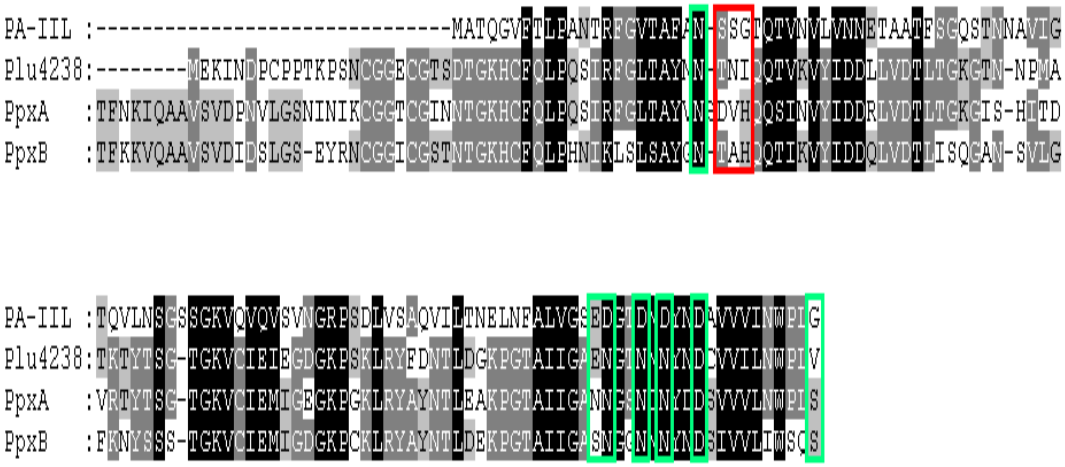


Figure 1.13: Sequence alignment of the lectin-like C-terminal domains of ppxA, ppxB and unknown protein Plu4238 from *P. luminescens* TT01.
Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

Although the amino acids involved in calcium binding are largely conserved, those involved in monosaccharide binding differ far greater then those seen in the Bcc species. Since it has previously been shown that a single amino acid change in the sugar binding loop can alter the specificity preference it can be hypothesised that the homologs found within *P. luminescens* may have differing sugar specificities to both PA-IIL and each other.

In summary the PA-IIL like lectin has been identified in a number of bacteria (*R. solanacearum*, *C. violaceum*, *Photorhabdus spp.*, *Burkholderia spp.*) and also in archaea, *Methanopyrus kandleri* (Majumder 2006). With the expanding number of bacterial genomes currently being sequenced, it is likely that more proteins of the PA-

IIL family will be identified. These homologs present a natural reservoir of structurally similar proteins with potentially a broad range of specificities. With more research these proteins could be exploited for protein array development (Hsu, Gildersleevec and Mahal 2008) and possible diagnostic tools.

1.7 Commercial applications of lectins

It is well recognised that most proteins are subject to post-translation modification (PTM), the most abundant and biologically significant PTM being glycosylation. Glycoslyation results in the addition of oligosaccharide moieties to proteins and cellular macromolecules such as lipids, bacterial cell membranes and eukaryotic cell membranes. Oligosaccharides have the potential to be assembled in branched structures as well as linearly (Figure 1.14), resulting in glycan (oligosaccharide) structures present on glycoconjugates exhibiting greater structural diversity than other celluar macromoleucels, e.g. there are 38,016 permutations of 3 monosaccharides that can theoretically produce a tri-saccharide compared to 64 possible permutations of four nucleotides in a three based codon. Adding to the complexity is the fact that a single protein can be glycosylated at a number of sites with potential glycan heterogeneity at each site. Glycans therefore have an inherent capacity and complexity for encoding enormous amounts of biological information.

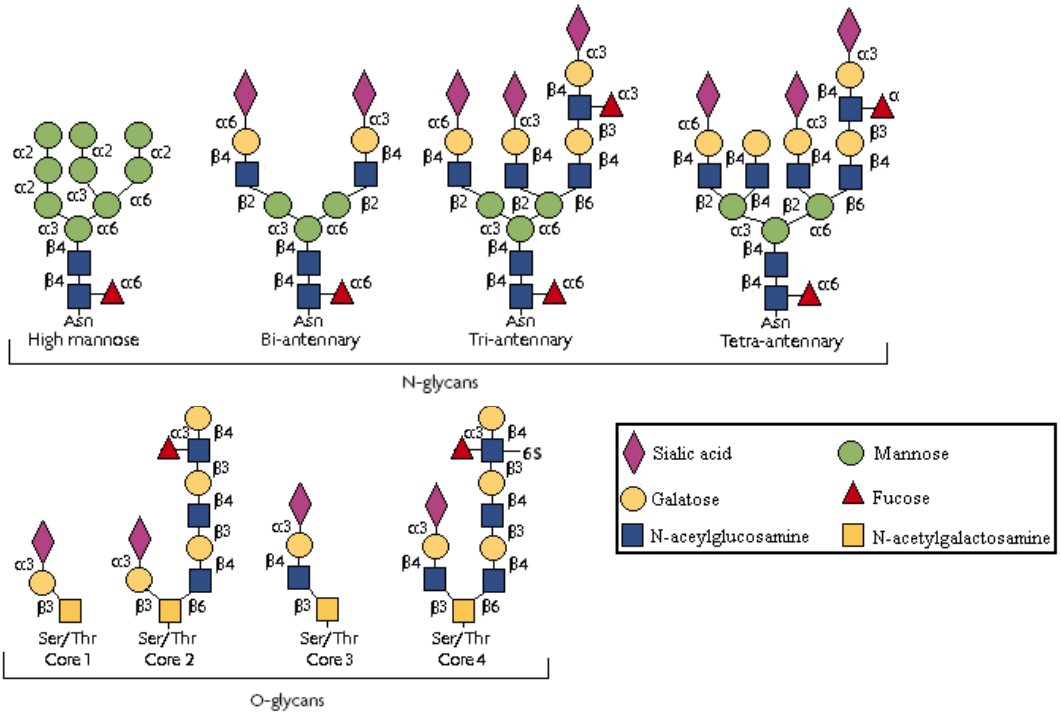


Figure 1.14: Examples of mammalian glycan structures
Schematics of representative examples of N-glycans and O-glycans indicated using the symbol nomenclature for monosaccharides (see key) to demonstrate the complexity of glycan structures. There are a variety of glycan structures not represented.

Glycosylation patterns have implications at a clinical level and at an industrial level in the biopharmaceutical industry.

Problems in the construction of glycan structures have been attributed to a number of disease states; Alzheimer's disease (Huang *et al.* 2004), auto immune disease (Hirschberg 2001), cystic fibrosis (Xia *et al.* 2005) and arthritis (Tomana *et al.* 1994). Therefore altered glycoform populations of a given glycoprotein may be diagnostic of the disease responsible. Cancer-specific glycan patterns have been identified and now account for tumour specific antigens used in diagnostics. Development of rapid and high throughput methodologies for the characterisation and detection of altered glycosylation patterns on glycoconjugates could greatly enhance clinical research and treatment of a variety of diseases.

Within the drugs industry protein based therapeutics are emerging as the largest class of new chemical entities being developed. In biological processes glycosylation affects the stability of protein confirmation, immunogenicity, clearance rate and more (Sinclair, Elliott 2005). Since different glycoforms can exhibit different biological properties it is essential to control glycosylation to maintain the quality of

biopharmaceutical molecules. Glycan heterogeneity also has serious implications for regulatory compliance and the FDA have indicated that sugar moieties on glycoconjugates will in future play a role in regards to license approval. To ensure carbohydrate heterogeneity the drugs industry must often implement additional expensive downstream purification. Since glycosylation is non-template driven it is not possible to predict glycan patterns. Analysis of glycans is therefore on the post-glycosylation stage. Current glycan monitoring techniques in the drugs industry are predominantly off-line and often require significant amounts of sample preparation. There has been an emphasis on techniques that require the removal of carbohydrate moieties from the molecule that has been glycosylated. The moieties are then analysed by one or more techniques such as chromatography, mass spectrometry or nuclear magnetic resonance spectrometry. These methods are expensive to establish and often required skilled operators to conduct.

The ability to characterise and monitor glycosylation of glycoconjugates is therefore of great interest for both clinical and commercial reasons. The methods currently used are reaching a bottle neck in terms of data analysis. The ability of lectins to distinguish between subtle variations of oligosaccharide structures make them suitable for the characterisation of glycans. Lectins have been utilised to some degree in glycobiology. Lectin Affinity Chromatography (LAC), where a lectin is immobilized, traditionally on agarose beads, is now commonly used for glycoprotein purification procedures. Multi-lectin affinity chromatography can be used to effect the fine fractionation and isolation of specific glycans and protein glycoforms (Wu *et al.* 2008, Yang, Hancock 2005, Geyer, Geyer 2006), often as an initial pre-concentration step prior to MS based glycoanalytical procedures. Labelled lectins are also widely used in glycoanalytical procedures such as lectin blotting (Zuber, Li and Roth 1998) and enzyme linked lectin assays (ELLA) which are analogous to the standard ELISA (Wu *et al.* 2008). The use of lectins in microarrays offers the greatest potential for indept, high throughput glycoanalysis and represents a powerful diagnostic tool potentially enabling a rapid screening method to detect alterations in glycosylation that may be indicative of specific infections or disease states (Pilobello *et al.* 2005). Lectin microarrays offer a number of advantages over traditional methods of glycan characterisation. Lectins can analyse *in situ* without the need for any prior release or sample preparation of glycans. Binding of lectins to a glycoprotein can reveal significant structural as well as compositional information about the glycan present due to the ability of lectins to

distinguish between subtle variations. Binding is dictated by the combination of a number of factors (Lis, Sharon 1998, Ambrosi, Cameron and Davis 2005);

- The overall composition
- The terminal monosaccharide present
- The specific positions and anomeric configurations of glycosidic bonds between the monosaccharide subunits of the glycan chain

Additional structural information may also be revealed due to the requirement of lectins for multivalent binding. This means binding is not only dependant on the presence of a single specific glycan but on the appropriate surface display of multiple glycans simultaneously at several independent glycosylation sites. The examination of the collective response of arrays of lectins with known carbohydrate binding specificity can therefore give detailed information about the overall glycosylation status of a glycoprotein. Identification and characterisation of novel lectins with a wide variety of carbohydrate binding specificities is critical to the development of highly accurate and reproducible lectin microarrays.

Most widely studied and characterised lectins to date have been those identified in plants, consequently most commercially available lectins are plant derived. Purification of plant lectins from source is time consuming and costly. Lectins purified in this manner often produce low yields of proteins and final preparation can be contaminated with a diverse range of biomolecules. For these reasons the purity, activity and specificity of commercial lectins available can vary from batch to batch from a single supplier or from supplier to supplier. Such inconsistencies complicate the use of lectins for glycoprotein analysis (Pilobello, Mahal 2007). Expression inconsistencies can be overcome via recombinant expression methodology. Eukaryotic proteins, such as plant lectins, often require PTM such as glycosylation. The requirement of PTM and the structural complexity of some plant lectins often leads to poor expression or inappropriate folding of plant lectins in the preferential host for recombinant protein production, *E. coli*. Expression of these lectins in *E. coli* usually results in the expressed proteins forming insoluble aggregates (Gemeiner *et al.* 2009). The application of lectins for glycoprotein analysis and selective purification has remained restricted to research and analytical scale activities due to these problems with the expression of plant lectins. Small scale application of lectin microarrays is not suitable for glycan analysis in the biopharmaceutical industry.

Microbial lectins have been largely unexploited for glycoprotein analysis. A huge potential reservoir of microbial lectins and lectin-like proteins has been revealed by advancement in genome sequencing and bioinformatics analysis. Microbial lectins, in particular, bacterial lectins are likely to be more amenable to expression from *E. coli* due to the more closely related genetic background. Additionally, lectin gene sequences can be genetically manipulated to incorporate specific amino acid sequence tags, e.g. six histidine codons (6HIS), which enable simple and rapid downstream purification via methods like Immobilised Metal Affinity Chromatography (IMAC). This facilitates the production of highly purified lectins and the products should therefore exhibit more consistent and reproducible specificity and activity. The ability to exploit the *E. coli* expression system to produce large quantities of these lectins, via cost effective and standardized protocols could finally lead to the application of lectins for industrial scale glycoprotein purification in the biopharmaceutical industry. Recombinant DNA techniques can be used to further enhance lectins and facilitate their implementation in glycoanalytical platforms, including;

- Addition of tags to facilitate specific immobilisation chemistries
- Addition of specific sequence tags and fusion proteins for rapid one-step purification and subsequent immobilisation via affinity based procedures
- Mutagenesis strategies with the intention of enhancing carbohydrate binding affinities or altering carbohydrate binding specificities (as discussed in Section 1.6)

Expansion of the number of lectins with a broad range of binding specificities, due to novel lectin identification or specific mutagenesis, amenable to large scale recombinant expression from *E. coli* could greatly enhance glycoanalytical techniques, in particular within the drugs industry.

1.8 Project aims and objectives.

The objective of this research project is to clone the gene for *PA-IIL* from *P. aeruginosa*. Oligonucleotide primers can be designed using the available sequence of *PA-IIL* from GenBank for PCR-based amplification of *PA-IIL* from chromosomal DNA. Upon cloning of the gene the project would focus on the development of a standardised method for expression from an *E. coli* expression system capable of producing active

recombinant PA-IIL. This approach for expression has been previously reported, as described in Section 1.2.1. If expression of active PA-IIL was possible in an *E. coli* system, large quantities of lectin could be rapidly and consistently produced for analysis.

Following successful expression of recombinant lectin, a standardised method of purification could be developed to enable production of yields of large quantities of highly purified protein. Recombinant fusion protein tagging system for one-step purification is the preferred method of purification. The effect of the addition of a fusion tag would be investigated. The optimal storage conditions for the long term storage of purified PA-IIL would be identified, facilitating availability of purified recombinant PA-IIL for the development of a highly sensitive ELLA assay for the determination of specificity.

Once standardised protocols for the expression, purification, stabilisation and characterisation of specificity are developed novel lectins with high homology to PA-IIL could be cloned and these standardised methods utilised to characterize their specificities.

2.0 Materials and Methods

2.1 Bacterial strains, primers sequences and plasmids

The bacterial strains, primers and plasmids used in this study are listed in Tables 2.1, 2.2 and 2.3 respectively. Selected plasmid maps are shown in Figure 2.1 to 2.5.

Table 2.1: Bacterial strains

Strain	Genotype	Features/Uses	Source
Escherichia coli			
XL10-Gold	Δ(*mcrA*),183Δ(*mcr*CB-*hsd*SMR-mrr), 173 endA1, *rec*A1, *rel*A1, *gyr*A96, *sup*E44, thi -1lac, Hte [F′*pro*AB *lac*I^{q}ZΔM15::Tn10 (tet^R)Amy (cam^R)]	High transformation efficiency (Tet^r, Cm^r) Expression host	Stratagene
BL21(DE3)	*F⁻, dcm lon, omp*T, *hsd*S$_B$(r$_{B-}$, m$_{B-}$)*gal* λ (DE3 [*lac*I^q lacUV5-T7 gene 1 indl sam7 nin5])	Protease deficient	Novagen
JM109	*end*A1, *rec*A1, *gyr*A96, *thi, hsd*R17 (r$_k^-$, m$_k^+$), *rel*A1, *sup*E44, λ–, Δ(*lac-pro*AB), [F′, *tra*D36, *pro*AB, *lac*I^{q}ZΔM15].	All purpose cloning strain, produced Plasmid DNA	Sigma
XL1-Blue	*rec*A1, *end*A1, *gyr*A96, *thi*-1, *hsd*R17(r$_K$-, m$_K$+), *sup*E44, *rel*A1, *lac*, [F', *proAB*, *lac*I^q ZΔM15::Tn10(tet^r)]		Promega
KRX	[F′, *tra*D36, Δ*omp*P, *pro*A$^+$B$^+$, *lac*I^q, Δ(*lac* Z)M15] Δ*omp*T, *end*A1, *rec*A1, *gyr*A96 (N al^r),*thi*-1, *hsd*R17 (r$_k^-$, m$_k^+$), e14$^-$ (McrA$^-$), *rel*A1, *sup*E44, Δ(*lac-pro*AB), Δ(*rha*BAD) ::T7 RNA polymerase	Protease deficient High transformation efficiency. Expression host	Promega
Rosetta	F$^-$ *ompT hsdS*$_B$(r$_B^-$ m$_B^-$) *gal dcm* pRARE2 (Cam^R)	Capable of expressing rare codons	Novagen
Pseudomonas aeruginosa			
PA01	Wild type		Dr. Keith Poole
Photorhabdus luminescens			
TT01	Wild type	Rif^R	Dr.David Clarke (U.C.C)

Table 2.2: Primers (Synthesised by Sigma-Aldrich, U.K.)
The restriction sites are underlined and non-binding tags are in bold type. The melting temperatures were calculated using the formula: $T_m = 4(GC) + 2(AT)$.

Name	Sequence (5' – 3')	T_m (^oC)
Cloning/Analysis		
PA-IIL-r	**AAAA**CTCGAGGCCGAGCGGCCAGTTGATCACC	73
PA-IIL-f	**AAAA**CCATGGCAACACAAGGAGTGTTCACCC	59
PA-IIL-r2	**AAAA**AGATCTGCCGAGCGGCCAGTTGATCACC	73
PA-IIL-f2	ATGGCAACACAAGGAGTGTTCACCC	68
QE60PAIIL-r	**GTACAT**AGATCTGTTAGCCGAGCGGCCAGTTGATCACC	73
QEPAIIL-f	GGATCCATGGCAACACAAGGAGTGTTC	58
QEPAIIL-r	AAGCTTTTAGCCGAGCGGCCAGTTG	65
PASY2-r	**CGATAG**AAGCTTTTATGTCAAAGGCCAATTTAAGAACACCAC	62
PASY2L-f	**GTACATC**GGATCCATGATACGGGTAAATATTGTTTCCAGTTACC	64
ppxA-r	AAGCTTTCATGATAATGGCCAATTCAAC	58
ppxA/L-f	GGATCCATGAATACCGGTAAACACTGC	55
ppxB-r	CTGCAGTCAGGACTGTGACCAAATTAACAC	59
ppxB/L-f	GGATCCATGGCAGCACAAATACCGG	60
ppxB/L.link.f*	GGTGGTGGTGGATCCATGACAAATACCGGTAAAC	55
ppxB/L-link.r*	ACCAGAGTGATGGTGATGGTGATGCGATCC	55
ppxB/L-linkfix.f*	GCCATTATCGGTGCCAGC	61
ppxB/L-linkfix.r*	TGTTCCCGGTTTTTCGTCAAGGG	69
Mutagenesis		
PAIILStrep-f*	CAGTTCGAAAAGGGATCCATGGCAACACAAGGAGTGTTCACCC	>75
PAIILStrep-r*	CGGGTGGCTCCACATAGTTAATTTCTCCTCTTTAATGAATTC	74
PAIILMU-f*	AACGTGCTGGTCAACAACGAGACGGCCGCGACC	>75
PAIILMU-r*	AACCGTCTGGGTGTTGGCCGCGTTGGCGAAGGCGGTGACG	>75

* denotes 5' phosphorylation of the primer

Table 2.3: Plasmids

Plasmid	Description	Source
pQE60	Expression vector for C-terminally tagged 6HIS proteins, T5 promoter/*lac* operon, *amp*^{R} ColE1 origin.	Qiagen
pQE30	Expression vector for N-terminally tagged 6HIS proteins, T5 promoter/*lac* operon, *amp*^{R}, ColE1 origin.	Qiagen
pQE30Xa	Expression vector for proteolytically removable N-terminally tagged 6HIS protein T5 promoter/*lac* operon, *amp*^{R}, ColE1 origin	Qiagen
pCR2.1	PCR cloning vector, Amp^R, Km^R, *lacZα*, ColE1 origin.	Invitrogen
pPC6	Expression vector for C-terminally tagged 6HIS proteins, P_{tac} promoter, *amp*^{R}, ColE1 origin	Invitrogen
Constructs		
pQE60PAIIL	pQE60 contain PA-IIL and stop codon prior to 6HIS tag	This work
pPAIIL3	pPC6 containing PA-IIL	This work
pQE30PAIIL	pQE30 containing PA-IIL	This work
pQEppxA/L	pQE30 containing lectin domain of ppxA gene	This work
pQEppxB/L	pQE30 containing lectin domain of ppxB gene	This work
pQEpASy2/L	pQE30 containing lectin domain of homolog 2 from *P.asymbiotica*	This work
Mutagenesis		
pQEStrepPAIIL	6His of pQE30.PAIIL replaced with StrepII tag by mutagenesis	This work
pQE30PAIILM1	Sugar binding loop of PA-IIl site specifically mutated	This work

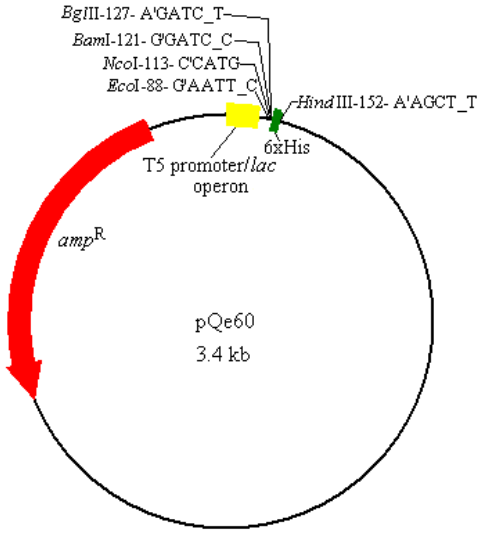


Figure 2.1: pQE60 vector (Quiagen)
Graphical representation of the 3431 b.p cloning vector pQE60 (Table 2.3) containing the following features; The multiple cloning site (MCS) located before the 6HIS amino acid sequence (green) which facilitates the translation of a fusion protein having a C-terminal 6HIS tag. Expression under the control of the T5 promoter/*lac* operon (yellow). The β-lactamase gene (red) conferring ampicillin resistance (amp^r) is located on the vector. The image was generated using BVTech (Section 2.12).

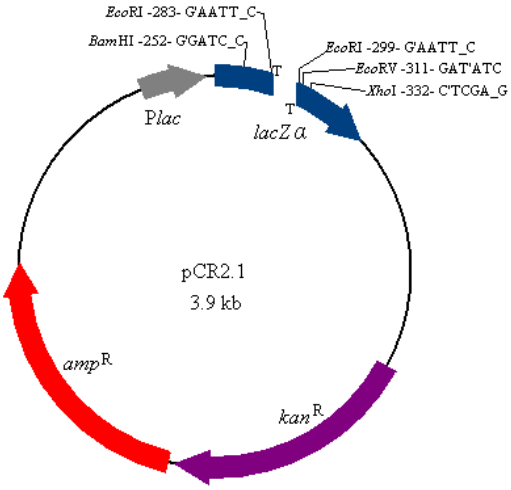


Figure 2.2: pCR2.1 Vector (Invitrogen)
Graphical representation of the TA cloning vector from Invitrogen, with the position of the 3' T overhands indicted with the letter T. These are located within the *lacZα* ORF, under the control of the P_{lac} promoter (shown in grey). Relavent restricition enzyme sites located within the MCS are displayed, as well as the position of the ampicillin (red) and kanamycin (purple) antibiotic resistance cassettes. The image was generated using BVTech (Section 2.12).

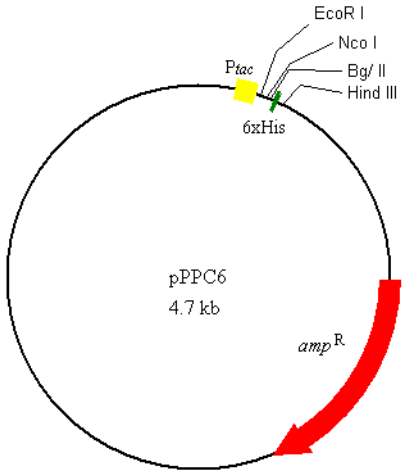


Figure 2.3: pPC6 vector
Graphical representation of the 4716 bp cloning vector pPC6 (Table 2.3) showing the relevant MCS restriction enzymes sites, the ampicillin resistance gene (*amp*^{r}) in red and the 6HIS coding sequencing in green. The cloning site is situated under the control of the P_{tac} promotor (yellow). The MCS is located before the 6HIS amino acid sequence facilitating the translation of a fusion protein having a C-terminal 6HIS tag. The image was generated using BVTech (Section 2.12).

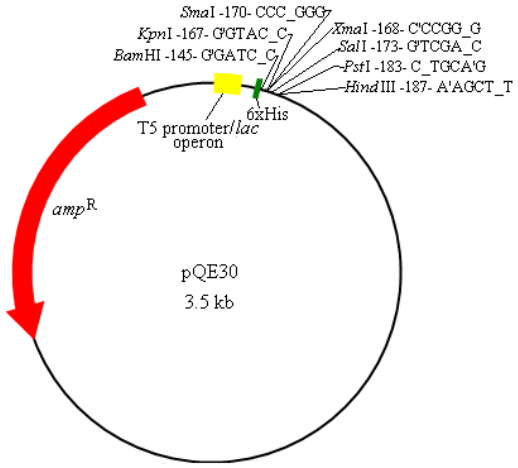


Figure 2.4: pQE30 vector (Quiagen)
Graphical representation of the 3461 b.p cloning vector pQE30 (Table 2.3) showing the relevant MCS restriction enzymes sites, the ampicillin resistance gene (*amp*^{r}) in red and the 6HIS coding sequence in green. The cloning site is situated under the control of the T5 promotor/lac operator. The MCS is located after the 6HIS amino acid sequence which facilitates the translation of a fusion protein having an N-terminal 6HIS tag. The image was generated using BVTech (Section 2.12).

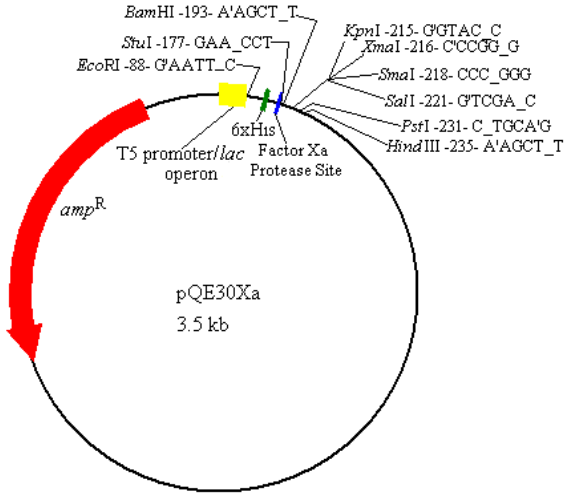


Figure 2.5: pQE30Xa vector (Quiagen)
Graphical representation of the 3509 b.p cloning vector pQE30Xa (Table 2.3) showing the relevant MCS restriction enzymes sites, the ampicillin resistance gene (*amp*^{r}) in red, the Factor Xa protease recognition site in blue and the 6HIS coding sequence in green. The cloning site is situated under the control of the T5 promotor/lac operator. The MCS is located after the 6HIS amino acid sequence and the protease recognition site which facilitates the translation of a fusion protein having an N-terminal 6HIS tag which can be cleaved by a FactorXa protease reaction. The image was generated using BVTech (Section 2.12).

2.2 Microbial media

All chemicals and reagents were obtained from Sigma-Aldrich unless otherwise stated. All chemicals were Analar grade. Microbiological media were obtained from Oxoid. Sterilisation was achieved by autoclaving at 121^{o}C and 15 lb/in^{2} for 20 min, unless otherwise stated.

- **Luria Bertani Broth (LB)**

Used for routine culturing of *E. coli.*

Tryptone	10g
Yeast Extract	5g
NaCl	10g

The solution was adjusted to pH 7.0 with NaOH, volume brought to 1L with dH_2O and sterilised by autoclaving. For solid LB, 15 g/L Technical Agar No.3 (Oxoid) was included.

- **SOB Medium**

Tryptone	20g
Yeast Extract	5g
NaCl	0.5g
KCL	2.5mM
dH_2O	1L
pH	7.0

The solution was sterilised by autoclaving and allowed to cool to 55^{o}C. Sterile solutions of $MgCl_2$ (1M) and Mg_2SO_4 (1M) were added to final concentrations of 10mM.

- **Terrific Broth**

Tryptone	12g
Yeast Extract	24g
Glycerol	4ml

The volume was brought to 1L with dH_2O and sterilised by autoclaving.

- **Super Broth**

Tryptone	35g

Yeast Extract	20g
NaCl	5g

The volume was brought to 1L with dH_2O and sterilised by autoclaving.

2.3 Solutions and buffers

- **Tris-EDTA (TE) Buffer**

Tris-HCL	10mM
Na_2-EDTA	1mM
pH	8.0

- **50x Tris Acetate EDTA (TAE) Buffer**

Tris	242 g/L
Glacial Acetic Acid	57.1ml/L
EDTA	100ml/L (of 0.5M stock)
pH	8.0

The solution was diluted to 1X with dH_20 before use.

- **Solutions for the 1, 2, 3 method**

Solution 1:

Glucose (0.5M)	1ml
Na_2-EDTA (0.1M)	1ml
Tris-HCl (1M)	0.25ml
dH_2O	to 10ml

Solution 2:

NaOH (1M)	2ml
SDS (10%)	1ml
dH_2O	to 10ml

The solution was made up fresh every month and stored at room temp.

Solution3:

Potassium acetate	3M
pH	4.8

To 60ml of 5M potassium acetate, 11.5ml of glacial acetic acid and 28.5ml of dH_2O was added. The resulting solution was 3M with respect to potassium and 5M with respect to acetate.

- **Gel Loading Dye (6X)**

Bromophenol Blue	0.25%
Xylene Cyanol	0.25%
Ficoll (Type 400)	15%

Made in dH_2O and sterilised by autoclaving. Bromophenol blue and/or xylene cyanol were used as appropriate. On a 1% agarose gel, bromophenol blue and xylene cyanol migrate approximately with the 300 b.p and 4000 b.p fragments respectively.

- **TBS buffer**

Tris-HCl	20mM
NaCl	150mM
$CaCl_2$	1mM
$MnCl_2$	1mM
$MgCl_2$	1mM
pH	7.5

- **TBST buffer**

TBST buffer with made by the addition 0.1% Triton X to TBS buffer.

- **TMB solution**

Citrate buffer:

Sodium citrate (0.1M)	3.63 ml
Citric acid (0.1M)	1.230 ml
dH_20	5 ml
pH	5.5

2mg of tetramethylbenzidine (TMB) was dissolved in 200 µl dimethyl sulfoxide (DMSO) and then added to 9.8 ml citrate buffer. Solution made up freshly 15 min before use. 3 µl 30% H_2O_2 was added directly before use.

- **Solutions for in-situ cell lysis**

SDS solution:

SDS	10 %

Denaturing solution:

NaOH	0.5M
NaCl	1.5M

Neutralization solution:

NaCl	1.5M
Tris-HCL	0.5M
pH	7.4

20x SSC:

NaCl	3M
Sodium.citrate	300mM

1X Tris Buffered Saline (TBS):

Tris-HCl	10mM
NaCl	150mM
pH	7.4

For TBST, the detergent Trition X-100 was added to a final concentration of 0.1% (v/v)

- **Lysis Buffer for Protein Purification**

NaH_2PO_4	50mM
NaCl	300mM
Imidazole	10mM
pH	8.0

- **5X Sample Buffer for SDS-PAGE**

Glycerol	50%l (v/v)
SDS (10%)	2% (w/v)
2-Mercaptoethanol	5% (v/v)
Bromophenol Blue	0.1% (w/v)
Tris-HCl, pH6.8	62.5mM
dH_2O	to 10ml

For native PAGE, SDS and 2-mercaptoethanol were omitted and replaced with the equivalent volume of dH_2O.

- **SDS-PAGE Running Buffer (5X)**

Tris-HCl	15g
Glycine	72g
SDS	5g
dH_2O	to 1L

For native PAGE, SDS was omitted.

- **Coomassie Blue Stain**

Coomassie Blue (R-250)	1g
Methanol	180ml
dH_20	180ml
Acetic Acid	40ml
dH_2O	to 400ml

- **Coomassie Blue De-Stain Solution**

Methanol	180ml
dH_20	180ml
Acetic Acid	40ml
dH_2O	to 400ml

- **Nickel Resin Binding Buffer**

Sodium Phosphate	20mM
NaCl	500mM
Imidazole	20-40mM
pH	7.4

The optimal imidazole concentration is protein dependent.

- **Nickel Resin Elution Buffer**

Sodium Phosphate	20mM
NaCl	500mM
Imidazole	40-250mM
pH	7.4

The optimal imidazole concentration for elution is protein dependent.

- **Stripping Buffer for Nickel IMAC Resins**

Sodium Phosphate	20mM
NaCl	500mM
EDTA	50mM
pH	7.4

- **Recharging Buffer for Nickel IMAC Resins**

$NiSO_4$	100mM

Phosphate Buffered Saline (PBS) 1X

NaCl	137mM
NaH_2PO_4	4.3mM
KCl	2.7mM
K_2HPO_4	1.5mM
pH	7.4

- **Buffers for the preparation of competent cells**

TB Buffer for Competent Cells

Pipes	10mM
$CaCl_2$	15mM
KCl	250mM
dH_2O	to 1L
pH with KOH	6.7

The pH of the solution was adjusted with KOH. $MnCl_2$ was then added to a final concentration of 55mM. The solution was then filter sterilised through a 0.45µm sterile filter and stored at 4^oC.

RF1 buffer for competent cells (Inoue *et al,* 1990)

RbCl	100mM
$CaCl_2$	10mM
Potassium acetate	30mM
Glycerol	15% (v/v)
pH (with HCl)	5.8

After the pH was adjusted $MnCl_2$ was added to a final concentration of 50mM. The solution was filter sterilised through a 0.45µm sterile filter and stored at 4^0C.

RF2 buffer for competent cells

RbCl	10mM
MOPS	10mM
$CaCl_2$	75mM
Glycerol	15% (v/v)
pH	6.8

The pH of the solution was adjusted with KOH. The solution was filter sterilised through a 0.45µM membrane and stored at 4^oC.

Ethidium bromide stain

A 10mg/ml stock solution of ethidium bromide in dH_2O was stored at 4^oC in the dark. For the staining of agarose gels, 100µl of the stock solution was mixed into 1L of dH_2O.

The staining solution was kept in a plastic tray and covered to protect against light. Used ethidium bromide stain was collected and left mixing overnight with a deactivating activated charcoal 'teabag' (GeneChoice). The clear liquid was disposed of routinely, and the ethidium waste was incinerated.

2.4 Antibiotic

Antibiotic used was from Sigma. Antibiotic stocks were prepared to a concentration of 100 mg/ml and stored at -20^{o}C.

- Ampicillin was prepared in dH_2O and was used at a final working concentration of 100 μg/ml for solid and liquid media for *E. coli*.

2.5 Storing and culturing bacteria

Strains were stored as glycerol stocks. Duplicate stocks were prepared for each strain and stored at -20^{o}C and -80^{o}C. A 1ml aliquot of a late log phase culture was added to 0.5ml of sterile 80% glycerol (v/v) and stored accordingly. Where hosts were harbouring plasmids, the appropriate antibiotic was added to the stock medium. Working stocks streaked on LB agar plate, containing antibiotics where appropriate, were stored on plates at 4^{o}C for a period up to one week.

2.6 Isolation and purification of DNA

The two procedures for the isolation of plasmid DNA were variably employed. The 1-2-3 Method (Section 2.6.1) yielded high quality DNA and was used routinely. The HiYield Plasmid miniprep Kit (ISIS) was used to prepare consistently pure and supercoiled plasmid DNA, mostly for the purpose of DNA sequencing.

2.6.1. Plasmid Preparation by the 1, 2, 3 Method

This method was adapted from the procedure described by (Birnboim, Doly 1979). The principle of this method is selective alkaline denaturation of high molecular weight chromosomal DNA while covalently closed circular DNA remains double stranded. A 1.5ml aliquot of a bacterial culture grown in selective media was pelleted by centrifugation at 13,000 rpm in a microfuge for 5 min. The supernatant was discarded and the pellet was re-suspended in 200 μl of Solution 1(Section 2.3). Alternatively,

bacterial growth was taken off an LB agar culture plate with a sterile loop and re-suspended in 200 µl of Solution 1. The re-suspension was left for 5min at room temperature. A 200 µl aliquot of Solution 2 (Section 2.3) was added, mixed by inversion and placed on ice for 5 min. To this mixture, 200 µl of Solution 3 (Section 2.3) was added, the tube mixed by inversion and placed on ice for 10 min. A colt of chromosomal DNA formed and was pelleted by centrifugation at 13,000 rpm for 10min. The supernatant was transferred to a fresh tube with 400 µl of phenol chloroform isoamylalcohol (25:24:1) and mixed by brief vortexing. Upon centrifugation at 13,000 rpm for 10 min the solution was divided into an upper aqueous layer and lower organic layer. The upper layer was removed to a fresh tube and an equal volume of isopropanol was added and mixed by inversion. After 10 min on ice the solution was centrifuged for 20 min at 13,000 rpm to pellet plasmid DNA. The pellet was washed with 70% ethanol and centrifuged at 13,000 rpm for 5 min. The ethanol was removed and the pellet was dried briefly in a SpeedVac (Savant) vacuum centrifuge. The plasmid DNA was re-suspended in 50 µl of TE buffer and 1 µl of Ribonuclease A was added to digest co-purified RNA. Plasmid DNA was stored at 4^{o}C.

2.6.2 Plasmid Preparation by HiYield plasmid miniprep kit

The kit was used according to the manufacturer's instructions: 1.5 ml of a bacterial culture was transferred to a microfuge tube and was centrifuged at 13,000 rpm for 5 min to collect the cells. The supernatant was discarded and the cell pellet was completely re-suspended in 200µl of re-suspension solution. Alternatively, bacterial growth was taken off an LB agar culture plate with a sterile loop and re-suspended in 200µl of re-suspension solution. 200µl of lysis solution was added and solution mixed by inversion to lyse the cells. 350µl neutralisation/binding buffer was added and mixed by inversion. The sample was centrifuged at 13,000 rpm for 10 min to precipitate cell debris, lipids, proteins and chromosomal DNA. The supernatant was removed and transferred to a spin column in a fresh microfuge tube and centrifuged at 13,000 rpm for 1 min to bind the plasmid DNA to the filter found within the spin column. The flow through was discarded and 750 µl of washing solution was added to the spin column. The column was centrifuge at 13,000 rpm for 2 min and the flow through discarded. This was followed by further centrifugation at 13,000 rpm for 2 min to dry the spin column. The

spin column was transferred to a fresh microfuge tube and 100µl elution buffer was added. The DNA was eluted by centrifugation at 13,000 rpm for 1 min.

2.6.3 Preparation of Total Genomic DNA Using the Wizard Genomic DNA Kit

The kit was used according to the manufacturer's instructions with some modifications. A 1 ml aliquot of a stationary phase culture was added to a 1.5 ml microfuge tube and the cells pelleted by centrifugation at 13000 rpm for 2 min. The supernatant was discarded and 600µl of the Nucleic Acid Lysis buffer was added to re-suspend the cells. The cells were lysed by incubation at 80^{o}C for 5 min and cooling to room temperature. 3µl of RNase solution was added to the cell lysate, inverted several times, and incubated at 37^{o}C for 15-60 min. To remove the protein in the sample, 200µl of the Protein Precipitation solution was added and vortexed vigorously. This was incubated on ice for 5 min to aid precipitation and centrifuged at 13000 rpm for 3 min. The supernatant was transferred to a fresh microfuge tube and 600µl of phenol chloroform isoamylalcohol (25:24:1) was added and the suspension was mixed slowly by inversion. The mixture was centrifugation at 13,000 rpm for 5 min and the upper aqueous phase was transferred to a fresh microfuge tube. 600 µl of phenol chloroform isoamylalcohol (25:24:1) was again added and mixed slowly by inversion. After centrifugation at 13,000 rpm for 5 min the upper aqueous phase was again transferred to a fresh microfuge tube. Phenol extraction was carried out by adding 700 µl of chloroform:isoamylalcohol (24:1), mixing by inversion and centrifugation at 13,000 rpm for 5 min. The upper aqueous layer was transferred to a fresh microfuge tube and 600 µl of room temperature isopropanol was added. The sample was mixed gently by inversion until thread-like strands of DNA formed a visible mass. The genomic DNA was then pelleted by centrifugation at 13000 rpm for 5 min. The pellet was washed twice with 100 µl 70% ethanol, air dried and re-suspended in 100 µl of DNA Rehydration Solution overnight at 4^{o}C. Genomic DNA was stored at 4^{o}C.

2.6.4 Purification and concentration of DNA samples

The sample containing the DNA to be precipitated was brought to 400 µl with dH_2O. 400 µl of phenol chloroform isoamyl alcohol (25:24:1) was added and mixed by brief vortexing. Upon centrifugation at 13,000 rpm for 5 min the mixture was divided into an upper aqueous and lower organic layer. The aqueous layer was removed to a fresh microfuge tube with an equal volume of chloroform and mixed by brief vortexing. The

tube was centrifuged at 13,000 rpm for 5 min and the upper aqueous layer was transferred to a new microfuge tube. A 1/10 volume of 3 M sodium acetate was added followed by an equal volume of isopropanol and mixed by inversion. The tube was left at room temperature for 60 min and then centrifuged at 13,000 rpm for 20 min to pellet the DNA. The pellet was washed with 70% ethanol and then dried briefly in a SpeedVac (Savant) vacuum centrifuge. The DNA was re-suspended in 20 – 50 µl of TE buffer.

2.7 Agarose Gel Electrophoresis for DNA Characterisation

DNA was analysed by electrophoress through agarose gels in a BioRad horizontal gel apparatus. Agarose was added to 1X TAE buffer to the required concentration (typically 0.7-1.2%) and dissolved by boiling. The agarose solution was poured into plastic trays and allowed to set with a plastic comb fitted to create sample wells. The 1X TAE buffer was also used as the running buffer. Loading dye was incorporated into DNA samples to facilitate loading of samples and to give indication of migration distance during electrophoresis. Mini-gels were run at 120 Volts for 25 min. Gels were stained by immersing in a bath of ethidium bromide for 15min. Gels were then visualised on a UV transilluminator (Syngen Gene Snap) coupled with an image analyser to capture the image to a PC. On every gel 7 µl of 1 Kb Plus DNA Ladder (Invitrogen) was run as a molecular size marker. When DNA was to be excised from the agarose gel (Section 2.8), SYBR safe DNA gel stain (Invitrogen) was used in preference to ethidium bromide. SYBR safe stained gels could be visualised on the Darkreader (blue light transilluminator), without the use of damaging UV, thereby increasing the efficiency of DNA clean-back.

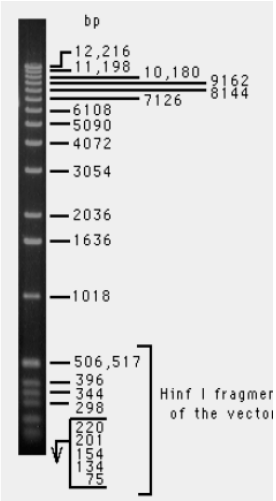


Figure 2.6: 1 kB DNA ladder
Representative imagine of the DNA ladder used in this study. Picture obtained from Invitrogen (www.invitogen.com).

2.8 Isolation of DNA from agarose gels/PCR reactions

DNA was isolated using HiYield Gel/PCR DNA fragment extraction kit (ISIS). The kit was used according to the manufacturer's instructions. DNA was visualize by the addition of SYBR safe (Sigma) to agarose gel and viewing on the Darkreader (blue light transilluminator). An agarose gel slice containing DNA was excised. Up to 300mg of the agarose gel slice was transferred into a microcentrifuge tube. 500µl of DF buffer was added to the sample and mixed by vortexing. The mixture was incubated at 55 ^{o}C for 10 -15 minutes until the gel slice was completely dissolved. During incubation, the tube was inverted every 2-3 mins. A fresh DF column was placed into a collection tube. 800 µl of the sample mixture from the previous step was applied into the DF column. The column was centrifuged at max. 13,000 xg for 30 sec. The flow-through was discarded. 500 µl of wash buffer was added and spun for 30 sec. The flow-through was discarded and the column was spun again for 2 min at max to dry the column. The column was transferred to a fresh microcentrifuge tube. 30 µl of elution buffer or water was added to the column. The column was allowed to stand for 2 min until buffer was absorbed and then centrifuged for 2 min at full speed to elute purified DNA.

2.9 Preparation of high efficiency competent cells

The two methods used to prepare competent cells were the TB method (Section 2.9.1) and the rubidium chloride method (Section 2.9.2). The latter method was favoured for its reliability.

2.9.1 TB method

This method was described by Inoue, H. *et al.* (1990). Sterile conditions were used throughout. 10 ml of LB broth containing the relevant antibiotics was inoculated with a single colony of the desired bacterial strain from a plate stock and cultured overnight at 37^{o}C. A 1L flask containing 250ml of SOB broth was inoculated with 250µl of the overnight culture and incubated at 37^{o}C shaking at 220 rpm. When the culture had reached an OD_{600} of 0.6, the flask was cooled in ice water to stop growth. All subsequent transactions took place at 4^{o}C. The culture was transferred to a sterile 250 ml centrifuge bottle and cells were collected by centrifugation using a Sorvall SLA-1500 rotor centrifuge at 3,000 rpm for 5 min. The supernatant was decanted and the pellet was gently re-suspended in 80 ml of ice cold TB buffer. The suspension was placed on ice for 10 min and the cells were again collected by centrifugation as before. The supernatant was decanted and the cell pellet was gently re-suspended in 15 ml of ice cold TB buffer. DMSO was added drop-wise with gentle swirling to a final concentration of 7% (v/v). The suspension was left on ice for 10 min. Aliquots of 800 µl of the cell suspension was dispensed into sterile 1.5 ml microfuge tubes. The cells were then, using – 80^{o}C ethanol, snap frozen and stored at -80^{o}C. Cells were routinely used within a few weeks.

2.9.2 Rubidium chloride method

This is an adaptation of the method outline by Hanahan, D. (1985). Sterile conditions were used throughout. 10 ml of LB broth containing the relevant antibiotics was inoculated with a single colony of the desired bacterial strain from a plate stock and cultured overnight at 37^{o}C. A 1 L flask with 250ml of LB broth was inoculated with 250 µl of the overnight culture and incubated at 37^{o}C shaking at 220 rpm. When the culture had reached an OD_{600} of ~0.5 (early-mid exponential phase) the flask was cooled in ice water. All subsequent transactions took place at 4^{o}C. The culture was transferred to a sterile 250 centrifuge bottle. The cells were collected by centrifugation at 3,000 rpm for 5 min (using a Sorvall SLA-1500 rotor). The supernatant was decanted and the cells

gently re-suspended in 60ml of chilled RF1 buffer. The suspension was left on ice for 90 min. The cells were again collected by centrifugation at 3,000 rpm for 5 min. The supernatant was decanted and the cells gently re-suspended in 8 ml of chilled RF2 buffer. Aliquots of 800 µl were prepared in sterile 1.5 ml microfuge tubes and the cells were then, using – 80^{o}C ethanol, snap frozen and stored at -80^{o}C. Cells were routinely used within a few weeks.

2.9.3 Transformation of competent cells

A microfuge tube of cells prepared according to the procedure outlined in Section 2.9.1 and 2.9.2 was used. An aliquot of competent cells was allowed to thaw on ice. 200 µl of the cell suspension was mixed gently with 1-5 µl of plasmid DNA in a sterile microfuge tube. The contents of the tube were incubated on ice for 30 min. The cells were heat shocked at 42^{o}C for 30 sec and then transferred back onto ice for 2 min. 0.8 ml of LB broth was added to the cells followed by incubation at 37^{o}C for 1 hour. A 100 µl aliquot of the resulting transformation mixture was plated on an appropriate selective media and the plates were incubated at 37^{o}C overnight.

2.9.4 Determination of competent cell efficiency

Competent cell efficiency is defined in terms of the number of colony forming units obtained per µg of transformed plasmid DNA. A 10 ng/µl stock of pUC18 plasmid DNA was diluted to 1 ng/µl, 100 pg/µl and 10 pg/µl. 1 µl of each dilution was transformed as described above. The cell efficiency was calculated from the number of colonies obtained, taking into account the dilution factor and the fraction of culture transferred to the spread plate.

2.10 TA cloning of PCR products

PCR products were cloned using Invitrogens Original TA Cloning Kit vector pCR2.1. The diagram below shows the concept behind the TA cloning method. The method is dependent on the fact that termostable polymerases like Red*Taq* (Invitrogen) DNA polymerase, which lacks 3'-5' exonuclease activity, leaves 3' A-overhangs (Figure 2.6) on amplified PCR products. PCR products generated with Red*Taq* DNA polymerase have a high efficiency of cloning in the TA cloning system.

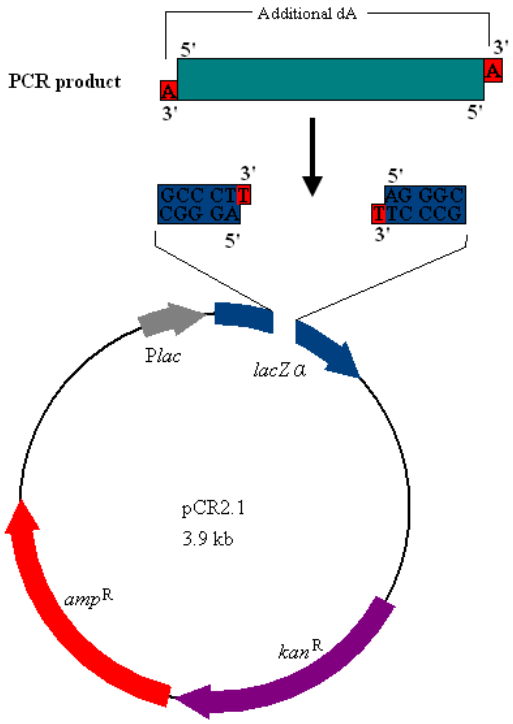


Figure 2.7: Principle of TA cloning
PCR products amplified by Red*Taq* polymerase have a deoxyriboadenosine (A) addition at the 3' end. The commercial pCR2.1 vector has been linearised within the *lacZα* gene and the cleavage site contains deoxyribothymidine (T) overhangs. This is complementary to the A overhang of the PCR product amplified by Red*Taq*, and can thus be ligated easily.

PCR products were amplified using a standard PCR reaction mixture (see Section 2.11.1) and using Red*Taq* DNA polymerase from Invitrogen, the PCR product was subsequently ligated with the TA pCR2.1 vector.

The ligation was set up as follows;

Fresh PCR product	2µl
pCR2.1 vector (25ng/µl)	2µl
10X ligation buffer	1µl
Sterile dH_2O	4µl
T4 DNA ligase	1µl
Total volume	10µl

The reaction was then incubated at room temperature overnight. Following incubation, 2-5µl of the ligation mixture was transformed into a suitable host strain and spread onto LB agar containing ampicillin (100ug/ml), 40µl of Xgal (40mg/ml) and 4µl of IPTG (100mM) to select for transformants by blue/white screening.

2.10.1 Blue/white screening of TA clones

The TA vector pCR2.1 (Figure 2.2) encodes the α-subunit of the LacZ protein. The chromosome of the host *E. coli* strain encodes the remaining Ω subunit to form a functional β-galactosidase. This enzyme metabolises the colourless galactose sugar X-gal to form 5-bromo-4-chloroindole which is bright blue. Transformants of the TA clones bearing plasmids with the PCR product were distinguished from transformants bearing re-circularised plasmids without inserts by plating on media containing IPTG and X-gal. Transformants harbouring plasmids with inserts did not express a functional *lacZα gene* product as the insertion point of the PCR product lies within the *lacZα* ORF, thus disrupting the gene. X-gal is therefore not hydrolysed due to the lack of the β-galactosidase enzyme production, resulting in white colonies. Transformants harbouring re-circularised plasmid DNA without any insert expressed functional *lacZα* gene product and thus X-gal was hydrolysed to produce colonies of a blue colour. The IPTG added to the media functions as the inducer to the Lac operon. Bacteria with no vector, which also appear white, are suppressed by the addition of an antibiotic in the growth media.

40µl of a stock solution of X-gal (40 mg/ml in dimethylformamide) and 4 µL of a stock solution of IPTG (100 mM in water) were spread on the surface of pre-made LB agar plates containing ampicillin and the plates were allowed to dry. The X-gal stock was stored at -20^{o}C and in the dark in order to prevent damage by light. The IPTG stock was stored at -20^{o}C. Transformations were plated out on the LB ampicillin/X-gal/IPTG plates and the plates were then incubated at 37^{o}C for 12-16 hours. The plates were then removed from the incubator and stored at 4^{o}C for several hours in order to allow colour development. Transformants harbouring plasmids with inserts appeared white on X-gal/IPTG plates. Transformants harbouring re-circularised plasmid DNA without any inserts appeared blue on X-gal/IPTG plates White colonies were thus picked from plates and screened for the presence of inserts.

2.11 Enzymatic Reactions

All enzymes and their relevant buffers were obtained from Invitrogen Life Technologies®, New England BioLabs® or Sigma Corporation and were used according to the manufacturer's instructions.

2.11.1 Polymerase chain reaction

PCR reactions (Mullis, Faloona 1987) were carried out using a Hybaid PCR Express Thermocycler.

- **Standard PCR reaction Mixture**

Template DNA	1µl
dNTPs (10mM of each)	1µl
Primers (25pmd/µl)	1µl of each
Buffer (10X)	5µl
dH_2O	40µl
Taq Polymerase	1µl

- **Standard PCR program cycle for Red*Taq* polymerase reactions**

Stage1:	Step 1: 95°C for 10 min
Stage 2:	Step1: 95°C for 1 min
	Step2: Annealing temperature for 30sec
	Step 3: 72°C for 1min for every Kb to be synthesised
Stage 3:	Step 1: 72°C for 10min

Stage 2 was typically carried out for 30 cycles as per manufacture guidelines.

- **Standard PCR program cycle for Phusion *Taq* polymerase reactions**

Stage 1:	Step 1**:** 98°C for 5 min
Stage 2:	Step1**:** 95°C for 10 sec
	Step 2**:** T_{ann} for 15 sec
	Step 3**:** 72°C for 20 sec per Kb to be synthesised.
Stage 3:	Step 1: 72°C for 10 min.

A 1 in 3 dilution of Phusion *Taq* polymerase was carried before use in PCR. All other polymerase used neat unless otherwise stated. Stage 2 was typically carried out for 50 cycles as per manufacture guidelines.

2.12 Gene manipulation

A variety of plasmid-based gene cloning approaches such as described by Maniatis, T. *et al.* (1982) were used, employing PCR techniques (Section 2.11.1), restriction endonucleases and DNA ligase.

2.12.1 Site-specific mutagenesis

Mutations were introduced into open reading frames on plasmid constructs by PCR amplification using complementary phosphorylated primers carrying the desired mutation. A standard PCR reaction mixture as shown in Section 2.11.1 was set up. The PCR was carried out using the high fidelity Phusion *Taq* DNA polymerase. The standard PCR program cycle for Phusion *Taq* polymerase reactions was used. Typically an extension time of 2 min and 30 sec was used to amplify plasmids of 3.5 kb to 4 kb. The template DNA was eliminated by digestion with *Dpn* I restriction endonucleases. *Dpn* I is biased toward a methylated recognition sequence. It selectively digests the template DNA from *dam*$^{+}$ *E. coli* strains over the newly synthesised DNA. The amplified plasmid, with the incorporated mutation, is circularisation by ligation which was facilitated by the use of phosphorylated primers.

2.13 DNA sequencing

Recombinant clones and potential mutants were verified by DNA sequencing. Commercial sequencing services were provided by MWG Biotech/Eurofins. Suitable sequencing primers for standard vectors were provided as part of the service. Samples were sent as plasmid preparations (Section 2.6).

2.14 Bio-Infomatics

Nucleotide and amino acid sequences were analysed using a variety of web-based tools.

- BLAST programs (Altschul *et al.* 1990) at NCBI www.ncbi.nlm.nih.gov were used to identify homologous sequences deposited in GenBank (Benson *et al.* 1996).

- Protein sequences, structure files and tools to calculate protein molecular weights and isoelectric points were obtained from the Swiss-Prot database (Bairoch, Apweiler 1997) at http://us.expasy.org and the Protein Data Bank (PDB) (Berman *et al.* 2000) at http://www.rscb.org/pdb.

- Alignments of DNA and protein sequences were performed using the CustralW (Larkin *et al.* 2007) at http://www.ebi.ac.uk and edited using the GeneDoc program (Nicholas, Nicholas and Deerfield II 1997) available for download at www.psc.edu/biomed/genedoc.

- Plasmid maps were constructed using the BVtech program available for download at http://www.biovisualtech.com/bvplasmid/downloadpl.htm.

- DNA sequences were analysed for restriction enzyme sites using the Webcutter 2.0 tool at http://rna.lundberg.gu.se/cutter2.

- Tertiary protein structure was predicted by the automated SWISS-MODEL server (Schwede *et al.* 2003), http://swissmodel.expasy.org, and subsequently analysed and visualised using the Pymol software (DeLano 2002) available for download at http://pymol.sourceforge.net/

2.15 Protein expression

An LB plate with the appropriate antibiotic was streaked with a glycerol stock of the strain containing the expression plasmid. One colony was selected and used to inoculate 5 ml of LB broth containing the appropriate antibiotic, and grown overnight at 37^{o}C while shaking.

A 250ml conical flask containing 100 ml of LB broth was inoculated with 1 ml of the overnight culture and the appropriate antibiotic added. The culture was incubated at 37^{o}C, shaking at 220 rpm, until an OD_{600} of 0.5-0.6 was reached. IPTG was added to a final concentration of 50 μM to induce expression. The culture was then either;

- Allowed to incubate for a defined period, sampling 2ml routinely at hourly intervals for time course analysis.
- Dropped to 30^{o}C and incubated overnight.

Samples or culture was centrifuged at 5,000 rpm for 5 min (using a Sorvall SLA-1500 rotor) to pellet the cells. The supernatant was discarded and the pellets were stored at -20^{o}C.

2.15.1 Preparation of cleared lysate

Cell pellets from Section 2.15 were washed with 50 ml per 100ml culture pellet with lysis buffer (Section 2.3). The cells were re-suspended in 10 ml lysis buffer and disrupted using a 3 mm micro-tip sonicator (Sonics and Materials Inc). Cells were subjected to 2.5 sec, 40 kHz pulses for 30 sec while on ice. The cell debris was removed by centrifugation at 4,000 rpm for 20 min. The cleared lysate was transferred to a fresh

universal container. For smaller samples/cultures, the volumes were reduced in correlation to the sample/culture size.

2.15.2 Colony blot procedure

Using a multichannel pipette, 400 µl of LB containing the required antibiotic was aseptically aliquoted into an autoclaved deep well 96-well plate. Each well was inoculated with a single colony obtained from a transformation plate. The 96 well plate was sealed with a Breathe-Easy Sealing Membrane (Sigma) and incubated at 37^{o}C, shaking at 800 – 1000 I/min overnight. Using a multichannel pipette 10µl of growth from each well was aseptically aliquoted to a fresh autoclaved deep well 96-well plate. The second 96-well plate was sealed and grown up as described above to ensure that there was sufficient growth to proceed with the experiment. An LBamp agar plate containing 50µM IPTG was overlaid with a nitrocellulose membrane and an equal volume of each overnight culture was spotted from the 96-well plate on the membrane. This was incubated at 37^{o}C for eight hours. A set of dishes for colony lysis was prepared. Each dish contained a sheet of 3MM paper soaked in of the one of the following

- 10% SDS solution (v/w)
- Neutralisation solution (X2)
- Denaturing solution
- 2x SSC

Excess fluid was discarded from the filters so that filter was moist but not wet. The nitrocellulose membrane was removed from the LBamp plate carefully and placed colony side up, on top of each filter, taking care to exclude air bubbles. The filters were incubated at room temp. as outlined in Table 2.4.

Solution	Time
SDS	10 min
Denaturing	5 min
Neutralization	5 min
Neutralization	5 min
2x SSC	15 min

Table 2.4: Incubation of nitrocellulose membrane for in situ lysis
Steps carried out on the nitrocellulose membrane of a 96-well clonoy blot. For denaturing, neutralization and 2X SSC buffers see Section 2.3.

The membrane was then washed twice for 10 min with TBS buffer (Section 2.3) and blocked for 1 hr in TBS containing 2.5% BSA. Unbound BSA was removed by washing twice for 10 min in TBS-Tween/Triton buffer (Section 2.3). The membrane was finally washed for 10 min in TBS buffer before incubation with antibody of 1:10,000 mouse anti histidine HRP-labelled antibody for 1 hr. Unbound antibody was removed by washing twice with TBS for 10 min each. A SIGMAFast 3,3-Diaminobenzidine (DAB) tablet was dissolved in dH_2O as per manufactures instruction and poured on the membrane. This was left to develop for 2-5 min, and stopped by rinsing twice with dH_2O.

2.16 Immobilised metal affinity chromatography (IMAC)

Immobilised Metal Affinity Chromatography (IMAC) is based on coordinate covalent binding between proteins and metal ions. In this work the coordinate binding between the metal ion Ni^{2+} to the amino acid histidine was exploited so that recombinant protein containing 6HIS were purified over the metal ion containing nitrilotriacetic acid (Ni-NTA) resin.

2.16.1 IMAC purification using Ni-NTA resin

A 0.5 ml aliquot of Ni-NTA (Amersham) was gently missed with 10ml of cleared lysate for 15 min at room temperature or overnight at 4^oC. The mixture was poured into a 0.7 x 15 cm column, and the resin was allowed to settle. The column was then washed twice with 10 ml of cell lysis buffer (Section 2.3). Two 10ml washes of lysis buffer supplemented with 30 mM imidazole were then passed through the column. 5 ml lysis buffer supplemented with 350 mM imidazole was used to elute the final protein. Samples taken throughout the procedure were analysed by SDS-PAGE (Section 2.24). The imidazole was removed from the eluted protein by buffer exchange (Section 2.20) over the FPLC. The purified sample was lyophilized (Section 2.21) and stored at -80^0C.

2.16.2 Recharging of Ni-NTA resin

This column was routinely recharged prior to re-use of the Ni-NTA resin. The used resin was poured into a column and washed with 2 column volumes of distilled water, followed by 2 column volumes 50% (v/v) ethanol. The resin was then stripped with 3 column volumns of 100 mM EDTA, pH 8.0. Remaining impurities were removed with

2 column volumes of 200 mM NaCl, followed by 2 column volumes of dH_2O. Hydrophobically bound proteins and lipoproteins were removed by washing with 10 column volumes of 30% isopropanol for 30 minutes, followed by 10 column volumes with water. The resin was the charged by adding 2 column volumes of 100 mM $NiSO_4$. The resin was washed again with 2 column volume of dH_2O before transfer to a sterile universal and stored at 4^oC in 20% ethanol.

2.17 Lectin purification using mannose agarose

PA-IIL has a high sugar binding affinity for L-fucose and D-mannose (see Chapter 1, (Gilboa-Garber 1982)). This binding affinity can be exploited to purify the lectin over agarose beads bearing D-mannose.

A 1ml aliquot of D-mannose agarose (Sigma-aldrich) was poured into a 20 ml column and washed with 2 column volumes of water to remove the storage buffer. The column was then equilibrated by washing with two column volumes of lysis buffer. Cleared cell lysate prepared as outlined in Section 2.15.1 was then added to the column, which was allowed to mix for 15 min at room temperature. Unbound material was collected as flowthrough, followed by a 20 ml wash using lysis buffer. Lectin elution was then facilitated by addition of 5 ml PBS supplemented with 0.3M D-mannose. The mannose was removed from the eluted protein by buffer exchange (Section 2.20) over the FPLC. The purified sample was lyophilized (Section 2.21) and stored at -80^0C.

2.18 StrepTrap™ HP Hiprep™ purification

StrepTrap™ columns are HiPrep™ columns prepacked with Streptactin sepharose™, available from GE healthcare. The StrepTrap column storage buffer (20% ethanol) was washed out with 5 column volumes of water and the resin equilibrated with at least 10 column volumes of lysis buffer (Section 2.3). The protein lysate was filtered through 0.45μM filter prior to loading onto the column using a syringe fitted with a luer adaptor. 5 to 10 column volume washes of PBS (Section 2.3) was then passed over the column. The protein was eluted by 6 column volumes of 2.5mM desthiobiotin. The desthiobiotin was removed from the eluted protein by buffer exchange (Section 2.20) over the FPLC.

2.19 Protein concentration

The two procedures for the quantification of protein were variably employed. The bicinchoninic acid assay combines the reduction of Cu^{2+} to Cu^{1+} by protein in an alkaline medium with the highly sensitive and selective colurimetric detection of Cu^{1+} by bicinchoninic acid. Quantification of the protein concentration by 280nm reading is based on a proteins ability to absorb ultraviolet light in solution with absorbance maxima at 200-280nm.

2.19.1 Quantitative determination of protein by BCA assay

The BCA assay described by Smith, P.K. *et al.* (1985) was utilised to quantify total protein in the range of 20-2,000 µg/ml. All samples were diluted appropriately to achieve a concentration within range of the assay. 150 µl of the protein sample was added in triplicate to 150 µl ml of BCA reagent (Sigma) and incubated at 37^{o}C for two hours. Absorbance was read at 562 nm. Bovine serum albumin (BSA) was used as the reference protein. BSA standards (0-2,000 ug/ml) were prepared in dH_2O and assayed in triplicate to yield a standard curve. Protein concentration of samples was determined from this standard curve.

2.19.2 Quantitative determination of protein by 280nm readings

Prior to reading all samples were spun down at 13,000 rpm for 10min to minimize any debris present in solution that may interfere in the absorbance readings. The UV lamp was adjusted to 280nm and allowed to warm up for approx. 15min before use. The zero absorbance was calibrated with buffer solution. The absorbance of the protein solution was read in triplicate. The concentration of the protein was then calculated using Equation 2.1;

Equation 2.1: Equation for the quantitative determination of protein concentration from absorbance readings at 280nm.

$$\textbf{Protein conc. (mg/ml) = Abs @ 280nm X path length X extinction co-efficient}$$

2.20 Desalting of purified protein using HiPrep 26/10 desalting column

The FPLC maximum post column pressure limit was set to 3.5 bar before attachment of the HiPrep 26/10 Desalting Column (GE Healthcare). The top of the column was attached to the pump outlet no 1 and the bottom of the column was screwed directly into the FPLC. Before the first sample application the column storage buffer (20% ethanol) was removed, and the column equilibrated with sample buffer. This was done by washing with 2 column volumes of water followed by 5 column volumes of sample buffer (either water or PBS see Section 2.3). Washing was done at the flow rate intended for chromatography, typically 20 ml/min. Samples were collected in 0.5 ml fractions in 96-well plates. All buffers and samples were filtered and degassed prior to application to the column.

2.21 Lyophilization

Samples were aliquoted (0.1 – 1ml) in microfuge tubes and stored at -80^{0}C until frozen solid. Microfuge tubes were opened and covered with Parafilm which was pierced prior to lyophilization. Samples were lyophilized at -45^{o}C and at a vacuum of 25 bars. Once lyophilized samples were stored at -80^{0}C.

2.22 Preparation of dialysis tubing

The required amount of tubing was placed in a 1 L glass beaker and rinsed thoroughly with distilled water. The beaker was filled with distilled water and ~ 1 g/L of EDTA was added. The beaker was brought to boil. The dialysis tubing was then boiled for 2 min. The water was allowed to cool and then poured off. The beaker was re-filled with fresh distilled water and brought to the boil. The dialysis tubing was boiled again for 2 min. The water was allowed to cool and poured off. The tubing was thoroughly rinsed with distilled water and stored at 4^{o}C in distilled water.

To perform dialysis the required length of prepared dialysis tubing was washed with dH_2O, one end was sealed, the protein solution transferred to the tubing and the tubing was sealed. The sample was placed in 5L of desired buffer and agitated for 2 hour at 4^{o}C. The buffer was re-freshed and the tubing was agitated overnight at 4^{o}C. The protein solution was removed from the tubing and stored at the appropriate temp.

2.23 Periodic treatment of BSA

50 mM sodium periodate solution made up freshly in dH_2O. The sodium periodate solution was added to 3% BSA solution made in sodium acetate solution pH 4.0. The solution was left in the dark at 4^oC for 2-4 hours constantly rocking. The solution was then buffer exchanged by dialysis against TBS to remove the sodium periodate.

2.24 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Protein samples were analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), based on the method outlined by Laemmli, U.K. (1970).

2.24.1 Preparation of SDS gels

20% or 15% resolving and 4% stacking polyacrylamide gels were prepared as per Table 2.5. Gels were cast using an ATTO vertical mini electrophoresis system. Upon addition of the TEMED to the resolving gel it was immediately poured and overlaid with isopropanol. After polymerisation, the isopropanol was removed and the stacking gel added. A comb was inserted into the top of the gel to form loading wells.

Table 2.5 SDS-PAGE gel recipes

Solution	20% Resolving Gel	15% Resolving Gel	4% Stacking gel
1.5M Tris-HCl, pH 8.8	1.88 ml	1.75 ml	
0.5M Tris-HCl, pH 6.8			0.74 ml
dH_2O	0.55 ml	1.68 ml	1.83 ml
Acrylamdie/Bis-acrylamide 30%/0.8% (w/v)	4.99 ml	3.5 ml	0.4 ml
10% (w/v) Ammonium Persulphate	38 µl	35 µl	15 µl
20% (w/v) SDS	38 µl	35 µl	15 µl
TEMED	8 µl	7 µl	6 µl

2.24.2 Sample preparation

To a microfuge tube, 20 µl of sample was added to 5 µl of 5X gel loading dye (Section 2.3). Samples were boiled for 5 minutes and applied to wells that were flushed of un-polymerised acrylamide. The insoluble fraction should be applied in the same concentration as the soluble fraction, so to the insoluble pellet obtained from a 1ml

sample, 20µl of 5X gel loading dye was added, and boiled for 5 min. The sample was then diluted with dH_2O to the original 1 ml.

2.24.3 Sample application

Routinely 15 µl of the prepared sample was applied to the SDS PAGE well. One lane on each gel was loaded with 10µl of a relative molecular weight protein marker (M_r) solution (SigmaMarker, Sigma, Figure 2.8) consisting of Rabbit Muscle Myosin (205 kDa), *E. coli* β-Galactosidase (116 kDa), Rabbit Muscle Phosphorylase b (97 kDa), Rabbit Muscle Fructose-6-phosphate Kinase (84 kDa), Bovine Serum Albumin (66 kDa), Bovine Liver Glutamic Dehydrogenase (55 kDa), Chicken Egg Ovalbumin (45 kDa), Rabbit Muscle Glyceraldehyde-3-phosphate Dehydrogenase (36 kDa), Bovine Erythrocyte Carbonic Anhydrase (29 kDa), Bovine Pancreas Trypsinogen (24 kDa), Soybean Trypsin Inhibitor (20 kDa), Bovine Milk α-Lactalbumin (14.2 kDa) and Bovine Lung Aprotinin (6.5 kDa). Gels were run at 25 mA for 70 min at room temperature using 1X SDS running buffer (Section 2.3).

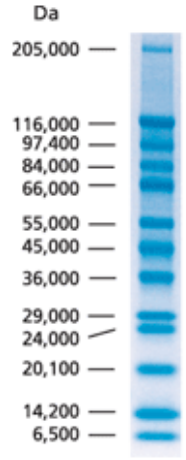


Figure 2.8: Wide range SigmaMarker visualised on 15% SDS-PAGE gel.
A representative image of the wide range protein maker used in this study. Image obtained from www.sigmaaldrich.com.

2.24.4 Gel analysis

Polyacrylamide gels were removed from the electrophoresis chamber and washed with dH_2O. Gels were routinely stained for 60 min in a Coomassie Blue stain solution (Section 2.3). Overnight de-stained followed using Coomassie Blue de-stain solution (Section 2.3). Subsequent soaking in dH_2O enhanced the protein bands further. For gels

which required a greater degree of sensitivity the silver staining method (Blum, Beier *et al.* 1987) was used as outlined in Table 2.6.

Table 2.6: Silver staining of SDS PAGE gels

Step	Duration	Reagent
Fixing	60 min	30% ethanol, 10% acetic acid (to 100ml)
Wash	15 min	20% ethanol
Wash	10 min	dH_2O
Sensitize	1 min	0.1% Sodium thiosulphate (w/v)
Rinse	2 x 20 sec	dH_2O
Sliver stain	30 min	0.1% Silver nitrate, 70 µl of 37% (v/v) formaldehyde stock
Rinse	2 x 20 sec	dH_2O
Develop	10 min (max)	3% Na_2CO_3, 50 µl of 37% (v/v) formaldehyde stock, 0.002% (w/v) $Na_2S_2O_2$
Stop	5 min	50 g/L Trizma Base, 2.5% Acetic Acid (v/v)

2.24.5 Native-PAGE

For the preparation of Native-PAGE gels, the concentrations outlined in Table 2.5 were used with the omission of 20% SDS, and supplemented with the appropriate volume of water. The samples to be analysed were not boiled and the sample buffer contained neither SDS nor 2-Mercaptoethanol. The gels were run in 1X Native-PAGE running buffer (Section 2.3).

2.25 Size exclusion Chromatography

The native molecular mass of the recombinant lectins under native conditions were determined by size exclusion chromatography (Andrews 1964, Whitaker 1963). Two methods were utilised during this study, one manual flow by gravity, the other using Fast Protein Liquid Chromatography (FPLC).

2.25.1 Gravity flow size exclusion chromatography

The molecular mass of recombinant PA-IIL under native conditions was determined by size exclusion chromatography (SEC). A 2.5 x 48 cm Sephadex G-100 gel-filtration column (Sigma-Aldrich) was equilibrated with 5 column volumes of PBS (Section 2.3) at room temp. 250µl of sample was applied to the column and allowed to flow by

gravity through the column. The proteins retention time was monitored by measuring eluent absorbance at an OD of 280nm. A standard curve was produced by running a number of known standards over the column separately. The standard proteins applied were:

Carbonic Anhydrase.; 200 µl of 10 mg/ml stock

BSA; 200 µl of 10 mg/ml stock

Cytochrome C; 200 µl of 5 mg/ml stock

Ovalbumin; 200 µl of 10 mg/ml stock

2.25.2 SEC by FPLC

Before attaching the SEC column to the AKTA Purifier FPLC (Amersham Biosciences) the maximum back pressure limit was set according to the manufactures guidelines. The column storage valve was disconnected from the gel filtration column, and the top connected to pump outlet no. 1. The bottom of the column screwed directly into the FPLC. To remove the column storage buffer (20% ethanol) two column volumes of water were pumped through the column (Pump P-900, Amersham Biosciences) at a flowrate of 1 mL/min (unless otherwise stated). The column was then equilibrated with 5 column volumes of sample buffer at the same flow rate, before sample application. With pump A connected to a stock of degassed sample buffer, 1 mg of sample was applied through the sample injection port in a volume of no more than 100 µl, and the run commenced, with a constant flow rate of 1 ml/min of buffer A. After sample addition, 2.5 column volumes of buffer was passed through the column. The proteins retention time was measured using the online Monitor U-900 (GE Healthcare), which read the eluent absorbance at OD 280nm.

2.26 Peroxidase-Binding Assay

The peroxidase-binding assay is a semi-quantitative assay (Huet, Bernadac 1974) for demonstration of PA-IIL interaction with cells. The assay is based on the ability of the lectin to combine both with the cells and with the mannose-bearing enzyme horseradish peroxidase (Gilboa-Garber, Mizrahi 1973). The assay was carried out in 96 well plate format. A 50µL volume of a 10 µg/ml lectin solution was immobilised in each well of a NUNC MaxiSorp ELISA plates at 4^{o}C overnight. Each sample was assayed in triplicate. The unbound lectin was removed by inverting the plate and the wells were washed with

TBS supplemented with 0.1% Triton X four times. A 50µl aliquot of HRP, supplemented with 10 mM $CaCl_2$ was then added at a concentration of 10µg/ml and left to incubate at room temperature for 1 hour. This was removed by inversion and washed with TBST as before. The wells were developed for 2-10min by the addition of 50µl of TMB substrate (Section 2.3). The reaction was stopped after 5 min by adding 10% H_2SO_4 and the absorbance was read at 415 nm or 450nm.

2.27 Enzyme Linked Lectin Assays on NUNC MaxiSorp plates

A 50µL volume of glycoprotein was immobilised in each well of a NUNC MaxiSorp ELISA plates at 4^{o}C overnight. Each sample was assayed in triplicate, at a concentration of 10 µg/ml. The unbound glycoprotein was removed by inverting the plate and the wells were blocked with 150 µl of 2.5% BSA in TBS for one hour. This solution was then removed by inverting the plate and washing with TBS supplemented with 0.1% Triton X four times. A 50µl aliquot of lectin in TBST-BSA 1%, supplemented with 10 mM $CaCl_2$ was then added at a concentration of 10µg/ml and let incubate at room temperature for 1 hour. This was removed by inversion and washed with TBST as before. This was followed with 50µl of 1:10,000 mouse anti-histidine antibody as appropriate. This was created fresh and diluted with TBST-BSA 1%, and incubated for one hour at room temperature. Unbound antibody was removed by inversion and washed four times with TBST, before the addition of 50µl of TMB substrate (Section 2.3). The reaction was stopped after 5 min by adding 10% H_2SO_4 and the absorbance was read at 415nm or 450 nm.

3.0 Cloning, expression, purification of recombinant PA-IIL

3.1 Overview

Samples enriched by a specific protein are rarely easily obtained from natural host cells. The Gram-negative bacterium *Escherichia coli* is one of the most commonly used hosts of recombinant heterologous protein expression (Baneyx 1999). It facilitates protein expression by its relative simplicity, rapid high density growth rate on inexpensive substrates and well characterized genetics which allow for large scale protein production. Many variable strains are available as well as a broad range of compatible cloning and expression vectors which confer flexible control of protein expression, e.g. location of expressed protein, yields, addition of fusion proteins, etc. The primary disadvantages of *E. coli* as an expression system include the inability to perform many of the post-translation modifications necessary for eukaryotic proteins, in particular glycosylation, and a limited ability to facilitate extensive disulphide bond formation (Makrides 1996). Predominately, available commercial lectins are of plant origin. These eukaryotic proteins require post-translational modification, preventing the use of the *E. coli* expression system for their production. Plant lectins are therefore often expressed from source resulting in poor yields of expressed protein and heterologous glycosylated lectin mixtures. Lectins from prokaryotic sources such as PA-IIL do not require post-translation modification and can therefore exploit the *E. coli* expression system to produce high yields of active recombinant protein to a high level of purity. This chapter describes the cloning of the *PA-IIL* gene from *P. aeruginosa* and development of a recombinant *E. coli* expression system, capable of producing active recombinant PA-IIL at high purity. The optimal conditions for long term storage of PA-IIL were also investigated.

3.2 Cloning and expression of the *P. aeruginosa* lectin PA-IIL in *E. coli*

In order to determine the optimal systems for the expression and purification of PA-IIL in *E. coli*, the *PA-IIL* gene was cloned into a number of *E. coli* compatible expression vectors. A number of the resulting constructs introduced tags to the PA-IIL protein. The addition of a tag enables simple purification via the IMAC methodology (Section 2.16.1). However, the addition of tags can also affect activity, stability, solubility and

folding and so a range of different tagging systems was evaluated. The cloning of the *PA-IIL* gene into a number of different expression vectors is described in the following sections.

3.3 Cloning for expression of untagged recombinant PA-IIL

The *PA-IIL* gene was cloned into the Qiagen pQE expression vector pQE60 (Table 2.3, Figure 2.1). The pQE60 vector features an optimized hybrid promoter-operator element (Bujard, Gentz *et al.* 1987) consisting of the phage T5 transcriptional promoter and a *lac* operator sequence, which increases *lac* protein (LacI) binding, enabling repression of this strong promoter. The β-lactamase gene conferring resistance to ampicillin (amp^R) is present on the vector. The pQE60 vector facilitates the translation of a fusion protein having a C-terminal 6HIS tag (Figure 3.1.).

Figure 3.1: The multiple cloning site (MCS) of the pQE60 expression vector.
Schematic outlining the positions of the relevant restriction sites, 6HIS tag and start and stop codons found within the MCS of the expression vector pQE60.

The *P. aeruginosa* strain PA01 was obtained from Prof. Keith Poole, and the sequence for the organism was obtained from the NCBI data bank, data entry no. NC002516, (Stover, Pham *et al.* 2000). The primers PAIIL-f and QE60.PAIIL-r were designed for the amplification of the *PA-IIL* gene as an *Nco*I-*Bgl*II 350 b.p fragment by PCR. The primed region of the *PA-IIL* gene for which the primers were designed is shown in Figures 3.2 and 3.3.

PAIIL-f: AAAA - (CCA TGG) - CAA CAC AAG GAG TGT TCA CCC
NcoI

QE60PAIIL-r: GTACAT - (AGA TCT) - GTT AGCCGA GCG GCC AGT TGA TCA CC
BglIII

Figure 3.2: Primers used for the amplification of the *PA-IIL* gene from *P. aeruginosa* PA01 for cloning into the pQE60 vector.
A point mutation was used to replace the stop codon TAG with the more frequently occurring stop codon TAA (base highlighted in red). The enzymatic restriction sites added to the *PA-IIL* gene are shown underlined and labelled.

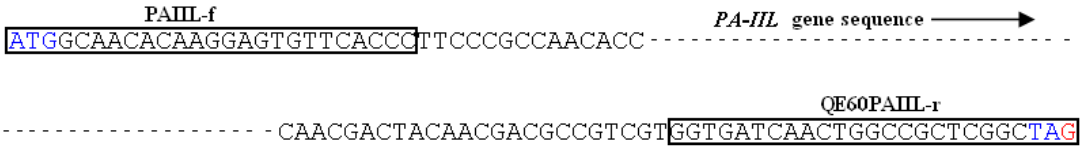


Figure 3.3: Primed region of the *PA-IIL* sequence from *P. aeruginosa* PA01 for suitable amplification for cloning into the pQE60 vector.
The start and stop codons of the *PA-IIL* gene are highlighted in blue. The base mutated by the primer for the introduction of the more frequently occurring TAA stop codon in place of the TAG stop codon is highlighted in red.

Genomic DNA was prepared from *P. aeruginosa* (Section 2.6.3) and used as template DNA for the amplification of the *PA-IIL* gene. The *PA-IIL* gene was amplified with a stop codon preventing the fusion of 6HIS to the C terminus of the translated protein (Figure 3.4). The conditions used for the PCR are shown in Table 3.1 below. The PCR was carried out with the high fidelity *Taq* polymerase, Phusion*Taq* (Invitrogen).

Table 3.1: PCR conditions for the amplification of the *PA-IIL* gene for cloning to the pQE60 vector.
Primers PAIIL-f and pQE60.PAIIL-r used for amplification as described above.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	30 sec
No. of Cycles	50

The PCR product was cleaned using Gel/PCR DNA fragment extraction kit (ISIS). Both the cleaned PCR product and the pQE60 vector (Table 2.3, Figure 2.1) were digested with *Nco*I/*Bgl*II and the restricted products again cleaned (Figure 3.4).

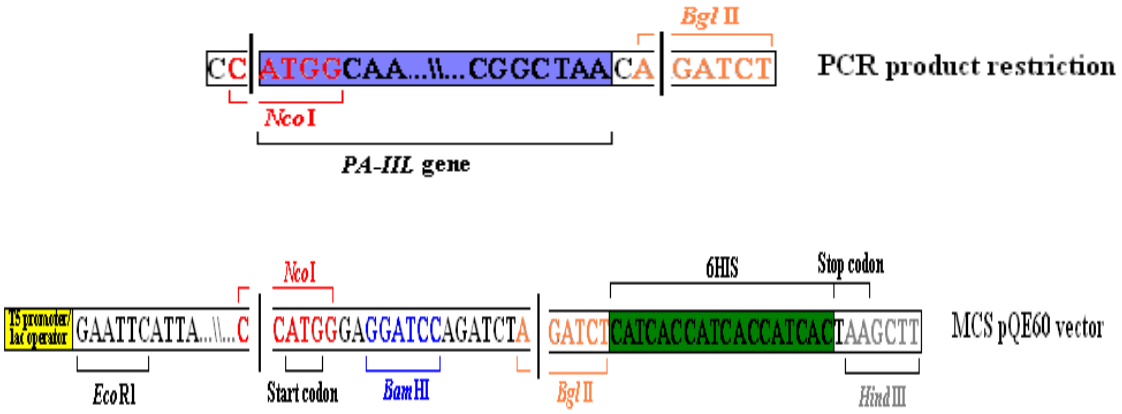


Figure 3.4: Schematic of the *Nco*I/*Bgl*II restricted PCR product amplified by PAIIL-f and QE60PAIIL-r and a schematic of the MCS of the pQE60 vector restricted by *Nco*I/*Bgl*II.
The point of restriction is indicated by a black line and break in the sequence.

The linearised pQE60 plasmid DNA was ligated with the purified *PA-IIL Nco*I/*Bgl*II fragment. The ligation mixture was transformed into *E. coli* JM109 cells, and spread onto LB amp plates (as outlined in Section 2.9.3). Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were screened by performing *Nco*I, *Bgl*II and *Nco*I/*Bgl*II double digests on the plasmid DNA prepared from each clone. The restriction digests were visualised on 0.7% agarose gels. Screening in this manner resulted in the identification of the desired plasmid named pQE60.PAIIL. Figure 3.5 below shows the results of the verifying restriction analysis on the plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pQE60 uncut
3	pQE60.PAIIL uncut
4	pQE60.PAIIL cut *Nco*I
5	pQE60.PAIIL cut *Bgl*II
6	pQE60.PAIIL cut *Nco*I/*Bgl*II
7	1Kb Ladder

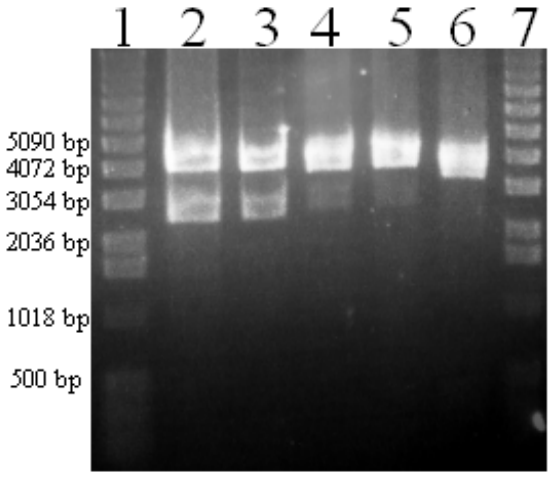


Figure 3.5: Analysis and restriction of plasmid pQE60.PAIIL.
The insert introduced to the pQE60 vector is small and therefore the band shift between pQE60 and pQE60.PAIIL is difficult to see. The restriction digest of the pQE60.PAIIL clone analysed on a 0.7% agarose gel (Section 2.7) showed that single restriction with either *Nco*I and *Bgl*II produced a band corresponding to the expected (~38000 b.p) linearised plasmid. Double restriction with *Nco*I/*Bgl*II excised a band of ~355 b.p resulting in a visible band shift lower of the plasmid which is indicative of the presence of the *PA-IIL* fragment within the clone indicating that the cloning of *PA-IIL* to the pQE60 vector was successful.

The presence of the correct insert within the pQE60.PAIIL vector (Figure 3.6) was further confirmed by DNA sequencing. Translation of protein from this expression vector results in untagged PA-IIL.

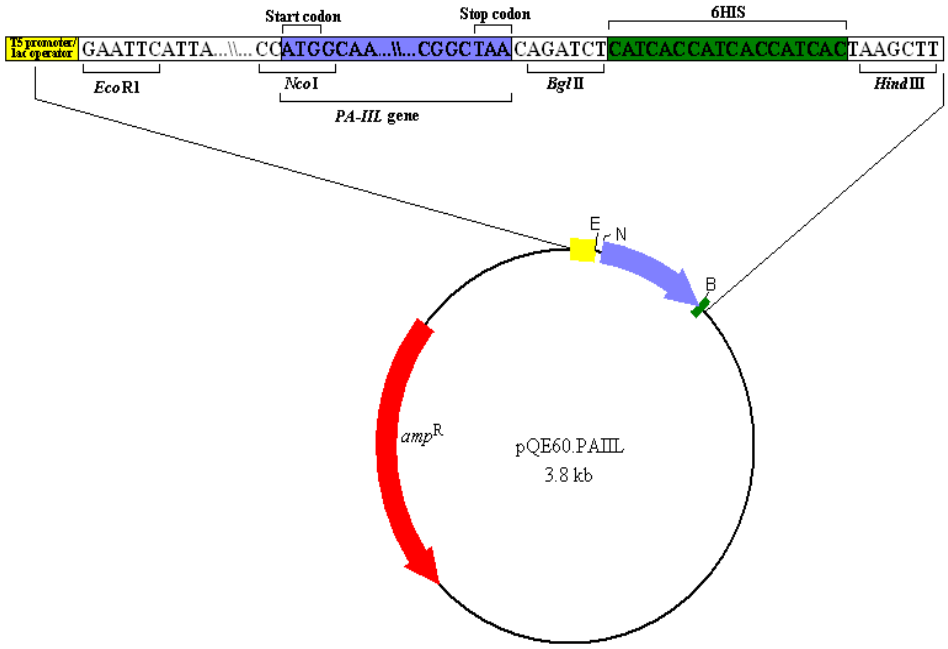


Figure 3.6: Schematic of pQE60.PAIIL vector.
Emphasised region of the pQE60.PAIIL vector outlines the relative positions of the *PA-IIL* gene, the relevant restriction sites, start and stop codon and 6HIS. Image generated using BVtech.

3.4 Cloning for expression of 6HIS tagged recombinant PA-IIL

3.4.1 Cloning for the expression of C-terminally 6HIS tagged PA-IIL

The pPC6 vector has been established as an expression vector within the laboratory. This expression vector is based on the pQE60 vector with the T5promoter/lac operator replaced by the *tac* promoter/operator (P*tac*). This hybrid promoter-operator element consists of the trp promoter and the lac UV5 promoter/operator, which increases *lac* protein (LacI) binding and enables repression of this strong promoter. The β-lactamase gene conferring resistance to ampicillin (amp^R) is present on the vector. The pPC6 vector facilitates the translation of a fusion protein having a C-terminal 6HIS tag (Figure 3.7).

Figure 3.7: The MCS of the pPC6 expression vector.
Schematic outlining the positions of the relevant restriction sites, 6HIS tag (green) and start and stop codons found within the MCS of the expression vector pPC6.

The primers PAIIL-f and PAIIL-r2 were designed for the amplification of the *PA-IIL* gene as an *Nco*I-*Bgl*II 353 b.p fragment by PCR. The primed region of the *PA-IIL* gene for which the primers were designed is shown in Figures 3.8 and 3.9.

PAIIL-f: AAAA - (CCA TGG) - CAA CAC AAG GAG TGT TCA CCC
*Nco*I

PAIIL-r2: AAAA - (AGA TCT) - GCC GAG CGG CCA GTT GAT CAC C
*Bgl*II

Figure 3.8: Primers used for the amplification of the *PA-IIL* gene from *P. aeruginosa* PA01 for cloning into the pPC6 vector.
The enzymatic restriction sites added to the *PA-IIL* gene are shown underlined and labelled.

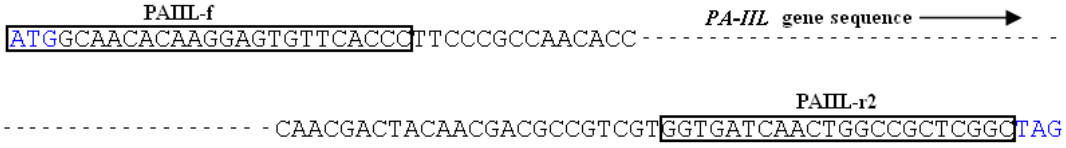


Figure 3.9: Primed region of the *PA-IIL* sequence from *P. aeruginosa* PA01 for suitable amplification for cloning into the pPC6 vector.
The start and stop codons of the *PA-IIL* gene are highlighted in blue. The stop codon of the *PA-IIL* gene lies outside the primed region to allow for the fusion of 6HIS to the C-terminus of the translated protein.

Genomic DNA was prepared from *P. aeruginosa* and used as template DNA for the amplification of the *PA-IIL* gene. The stop codon of the *PA-IIL* gene lies outside the binding region of the primers to facilitate the fusion of 6HIS to the C terminus of the translated protein. The conditions used for the PCR are shown in Table 3.2 below.

Table 3.2: PCR conditions for the amplification of the *PA-IIL* gene for cloning to the pPC6 vector.
Primers PAIIL-f and PAIIL-r2 used for amplification as described above.

PCR Condition	
Annealing Temp.	68°C
Annealing Time	30 sec
Extension Time	30 sec
No. of Cycles	30

The PCR was carried out with the *Taq* polymerase Red*Taq* (Invitrogen). Red*Taq* polymerase has a nontemplate-dependent terminal transferase activity that adds a single deoxyriboadenosine (A) to the 3' terminus of PCR products (Figure 3.10).

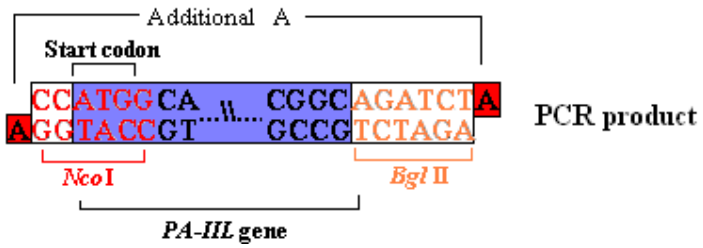


Figure 3.10: *PA-IIL Nco*I-*Bgl*II fragment after PCR amplification by Red*Taq* polymerase.
Addition of A overhangs to 3' terminus of the *PA-IIL* PCR product by Red*Taq* are shown in red. The relevant restriction sites are indicated.

The commercial pCR2.1 vector (Table 2.3, Fig. 2.2) has been linearised within the *lacZα* gene and the cleavage site contains deoxyribothymidine (T) overhangs. The T overhangs are complementary to the A overhangs of PCR products amplified using Red*Taq* polymerase, and can thus be ligated easily. The PCR *PA-IIL Nco*I-*Bgl*II product was cleaned and ligated with the linearised TA cloning vector pCR2.1 as outlined in Figure 3.11.

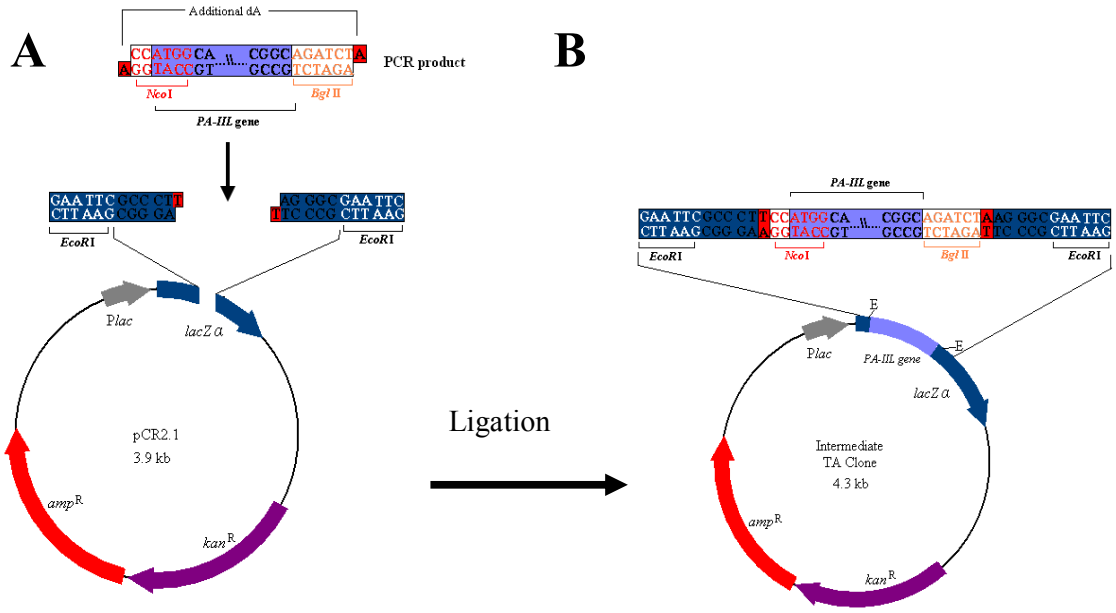


Figure 3.11: Schematic of the intermediate clone formed by TA cloning of the *PA-IIL Nco*I-*Bgl*II fragment to the pCR2.1 vector
[A] Schematic of the PCR *PA-IIL* fragment amplified by Red*Taq* polymerase and the linearised pCR2.1 prepared for ligation. [B] Schematic of the intermediate cloned formed by TA cloning. The complementary T overhangs of the commercial linerized pCR2.1 vector and the A overhangs of the Red*Taq* polymerase amplified *PA-IIL Nco*I-*Bgl*II are highlighted in red. The relevant restriction sites are indicated. Image generated using BVtech.

The PCR product was ligated with the TA pCR2.1 vector according to the TA protocol (Section 2.10). The resulting ligation was transformed into *E. coli* JM109 cells, and spread onto LB amp plates containing IPTG and X-Gal and the transformants were screened using the blue/white screening method as described in Section 2.10.1. White colonies were isolated and screened by performing *Eco*RI digests on plasmid DNA isolated from each clone. In this way the desired intermediate TA clone was obtained. Figure 3.12 shows the results of the verifying restriction analysis on the intermediate TA plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pCR2.1 uncut
3	pCR2.1 cut *EcoR*I
4	Inter. TA clone uncut
5	Inter. TA clone *EcoR*I

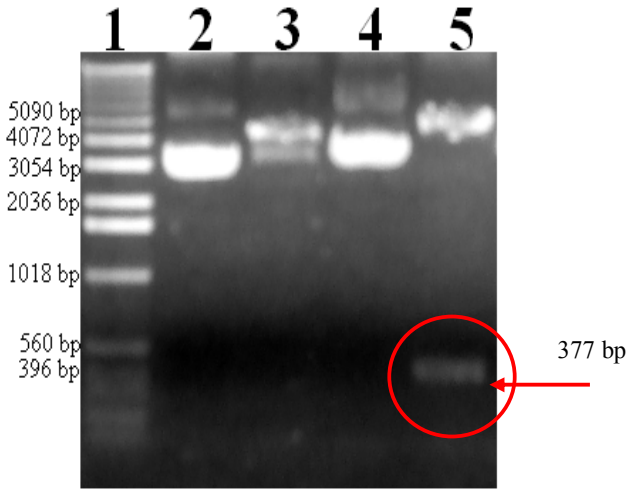


Figure 3.12: Analysis and restriction of the intermediate TA vector.
The restriction digest of the intermediate TA clone analysed on a 0.7% agarose gel (Section 2.7). *EcoR*I sites are situated on either side of the insert point of pCR2.1 (Figure 3.11). A restriction digest of the empty pCR2.1 TA clone with *EcoR*I resulted in linearised plasmid with a minor band shift. A restriction digest of the intermediate clone with *EcoR*I excised a band of approx 380 b.p, circled in red, indicating that *PA-IIL* fragment is present and therefore the cloning of *PA-IIL* to the pCR2.1 vector was successful.

Plasmid DNA was prepared from the intermediate TA vector and digested with *Nco*I/*Bgl*II. The digested DNA was separated on a 1% agarose gel and the band corresponding to the *PA-IIL* fragment was purified from the gel Gel/PCR DNA fragment extraction kit (Section 2.8). pPC6 was linearised with *Nco*I*Bgl*II and ligated with the purified *PA-IIL Nco*I-*Bgl*II fragment (Figure 3.13).

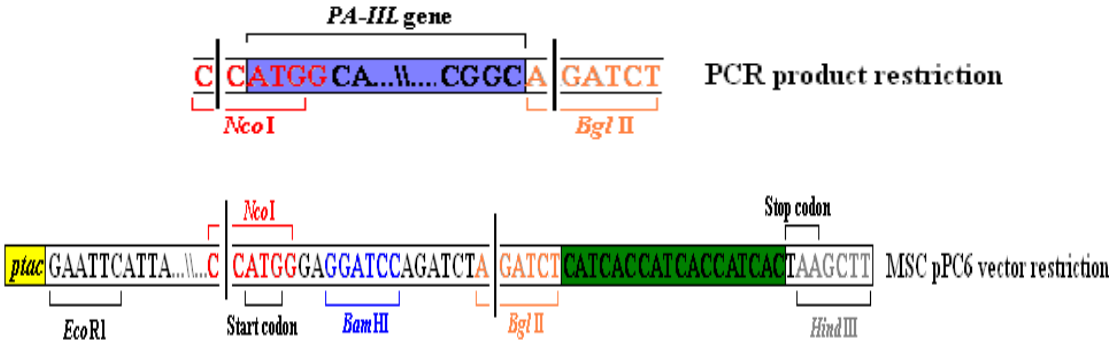


Figure 3.13: Schematic of the *Nco*I/*Bgl*II restricted *PA-IIL* fragment excised from the intermediate TA clone and the schematic of the MCS of the pPC6 vector restricted by *Nco*I/*Bgl*II.
The point of restriction is indicated by a black line and break in the sequence.

The ligated DNA was transformed to *E. coli* JM109 and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were

screened by performing *Nco*I, *Bgl*II and *Nco*I/*Bgl*II double digests on plasmid DNA prepared from each clone. The restriction digests were visualised on a 0.7% agarose gel. Screening in this manner resulted in the identification of the desired plasmid named pPAIIL3. Figure 3.14 below shows the results of the verifying restriction analysis on the plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pPAIIL3 uncut
3	pPAIIL3 cut *Nco*I
4	pPAIIL3 cut *Bgl*II
5	pPAIIL3 cut *Nco*I/*Bgl*II
6	*PAIIL3* PCR product

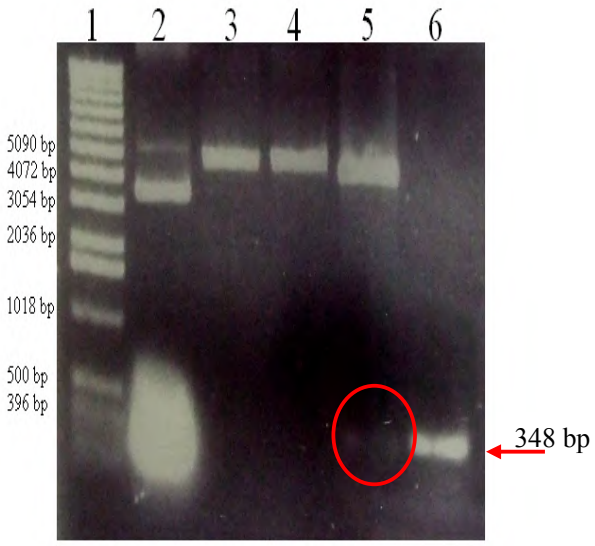


Figure 3.14: Analysis and restriction of the pPAIIL3 plasmid.
The restriction digest of the pPAIIL3 clone analysed on a 0.7% agarose gel (Section 2.7) showed that a single restriction with *Nco*I and *Bgl*II produced a band corresponding to the expected (~5000 b.p) linearised plasmid. Double restriction with *Nco*I/*Bgl*II excised a band of ~350 b.p, corresponding to the *PA-IIL* gene. The size of the excised band was confirmed by comparison to the PCR product produced by primers PAIIL-f and PAIIL-r2. The insert is indicative of the presence of the *PA-IIL* fragment within the clone indicating that the cloning of *PA-IIL* into the pPC6 vector was successful.

The presence of the correct insert within the pPAIIL3 plasmid (Figure 3.15) was further confirmed by DNA sequencing. Translation of protein from this expression vector results in C-terminal 6HIS tagged PA-IIL.

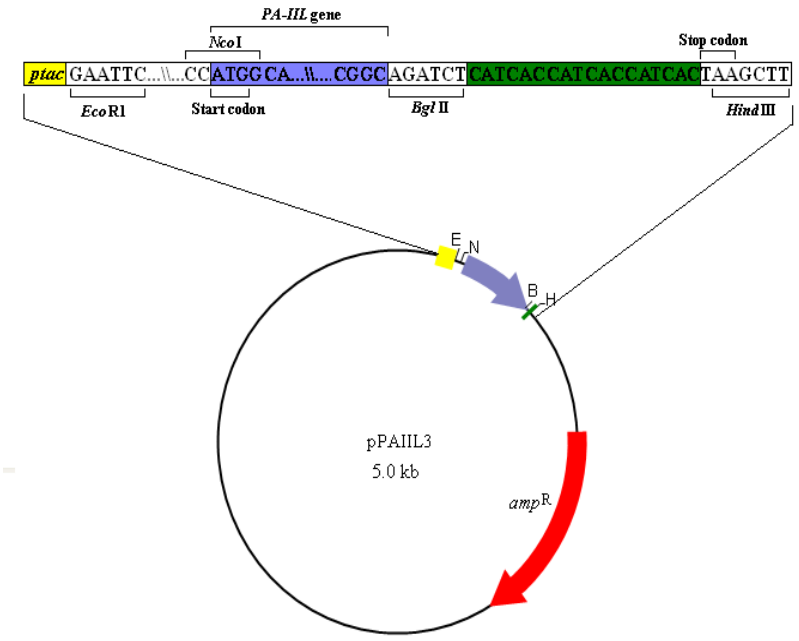


Figure 3.15: Schematic of pPAIIL3 construct.
Emphasised region of the pPAIIL3 vector outlines the relative positions of the*PA-IIL* gene, the relevant restriction sites, start and stop codon and 6HIS. Image generated using BVtech.

3.4.3 Cloning for the expression of N-terminally 6HIS tagged PA-IIL

The *PA-IIL* gene was cloned into the Qiagen pQE expression vector pQE30 (Table 2.3, Figure 2.4). The vector pQE-30 features the same promoter operator system and β-lactamase gene conferring resistance to ampicillin (amp^R) as the pQE60 vector. The pQE30 vector facilitates the translation of a fusion protein having an N-terminal 6HIS tag (Figure 3.16).

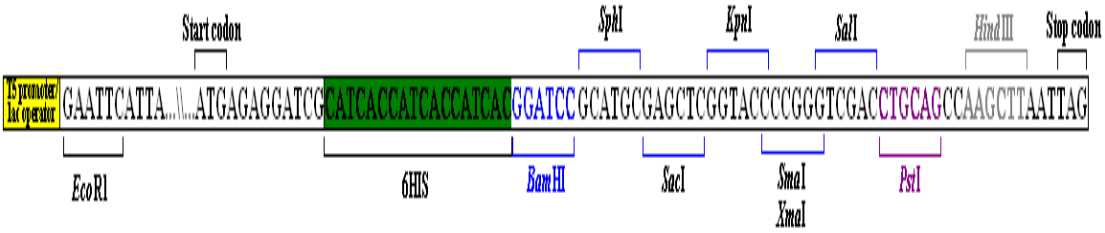


Figure 3.16: The MCS of the pQE30 vector.
Schematic outlining the positions of the relevant restriction sites, 6HIS tag and start and stop codons found within the MCS of the expression vector pQE30.

The primers QEPAIIL-f and QEPAIIL-r were designed for the amplification of the *PA-IIL* gene as a *BamH*I-*Hind*III 360 b.p fragment by PCR. The primed region of the *PA-IIL* gene for which the primers were designed is shown in Figures 3.17 and 3.18.

QEPAIIL-f: (GGA TCC) - ATG GCA ACA CAA GGA GTG TTC
BamH I

QEPAIIL-r: (AAG CTT) - TTA GCC GAG CGG CCA GTT G
Hind III

Figure 3.17: Primers used for the amplification of the *PA-IIL* gene from *P. aeruginosa* PA01 for cloning into the pQE30 vector.
The enzymatic restriction sites added to the *PA-IIL* gene are shown underlined and labelled

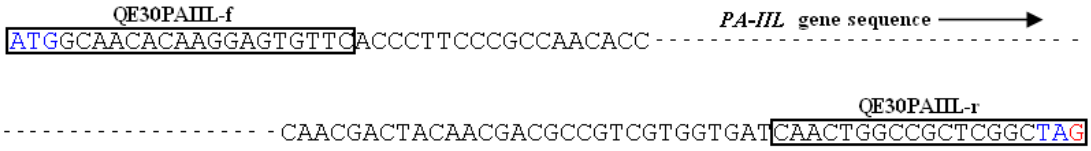


Figure 3.18: Primed region of the *PA-IIL* sequence from *P. aeruginosa* PA01 for suitable amplification for cloning into the pQE30 vector.
The start and stop codon of the *PA-IIL* gene highlighted in blue. The base mutated by the primer for the introduction of the TAA stop codon is highlighted in red.

Genomic DNA was prepared from *P. aeruginosa* and used as template DNA for the amplification of the *PA-IIL* gene. The conditions used for the PCR are shown in Table 3.3 below. The PCR was carried out with the high fidelity *Taq* polymerase, Phusion*Taq*.

Table 3.3: PCR conditions for the amplification of the *PA-IIL* gene for cloning to the pQE30 vector.
Primers QEPAIIL-f and QEPAIIL-r used for amplification as described above.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	30 sec
No. of Cycles	50

The PCR product was cleaned using Gel/PCR DNA fragment extraction kit. Both the cleaned PCR product and the pQE30 vector were digested with *BamH*I/*Hind*III and the restricted products cleaned (Figure 3.19).

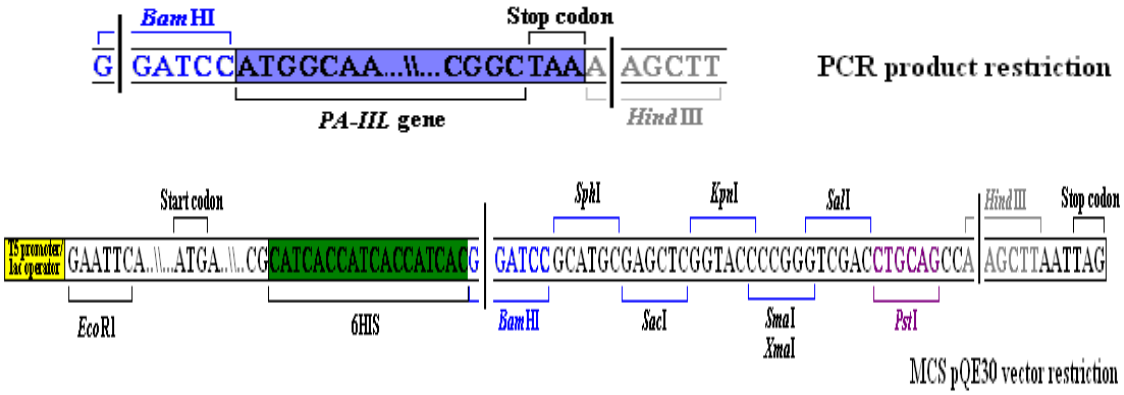


Figure 3.19: Schematic of *Bam*HI/*Hind*III restricted PCR product amplified by QE30.PAIIL-f and QE30.PAIIL-r and schematic of the MCS of the pQE30 vector restricted by *Bam*HI/*Hind*III.
The point of restriction is indicated by a black line and break in the sequence.

The linearised pQE30 plasmid DNA was ligated with the purified *PA-IIL BamH*I-*Hind*III fragment. The ligated DNA was transformed to *E. coli* XL10-Gold and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were screened by performing *BamH*I, *Hind*III and *BamH*I/*Hind*III double digests on plasmid DNA prepared from each clone. The restriction digests were visualised on 0.7% agarose gels. Screening in this manner resulted in the identification of the desired plasmid named pQE30.PAIIL. Figure 3.20 below shows the results of the verifying restriction analysis on the plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pQE30 uncut
3	pQE30.PAIIL uncut
4	pQE30.PAIIL cut *Bam*HI
5	pQE30.PAIIL cut *Hind*III
6	pQE30.PAIIL cut *Bam*HI/*Hind*III

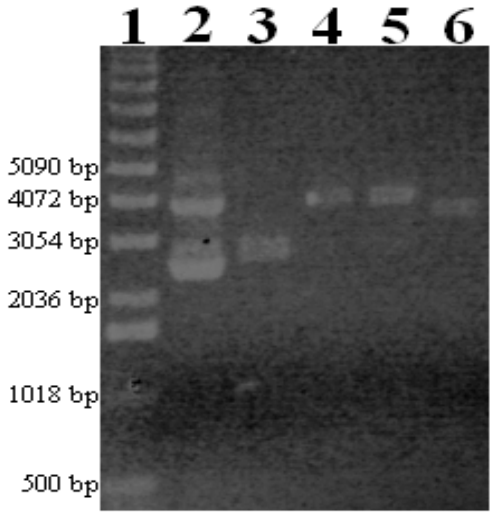


Figure 3.20: Analysis and restriction of the pQE30.PAIIL plasmid.
The restriction digest of the pQE30.PAIIL clone analysed on a 0.7% agarose gel (Section 2.7) showed single restriction with *BamH*I and *Hind*III produced a band corresponding to the expected (~3800 b.p) linearised plasmid. Double restriction with *BamH*I/*Hind*III excised a band of ~350 b.p resulting in a visible band shift lower of the plasmid which is indicative of the presence of the *PA-IIL* fragment within the clone indicating that the cloning of *PA-IIL* into the pQE30 vector was successful.

The presence of the correct insert within the pQE30.PAIIL plasmid (Figure 3.21) was further confirmed by DNA sequencing. Translation of protein from this expression vector results in N-terminal 6HIS tagged PA-IIL.

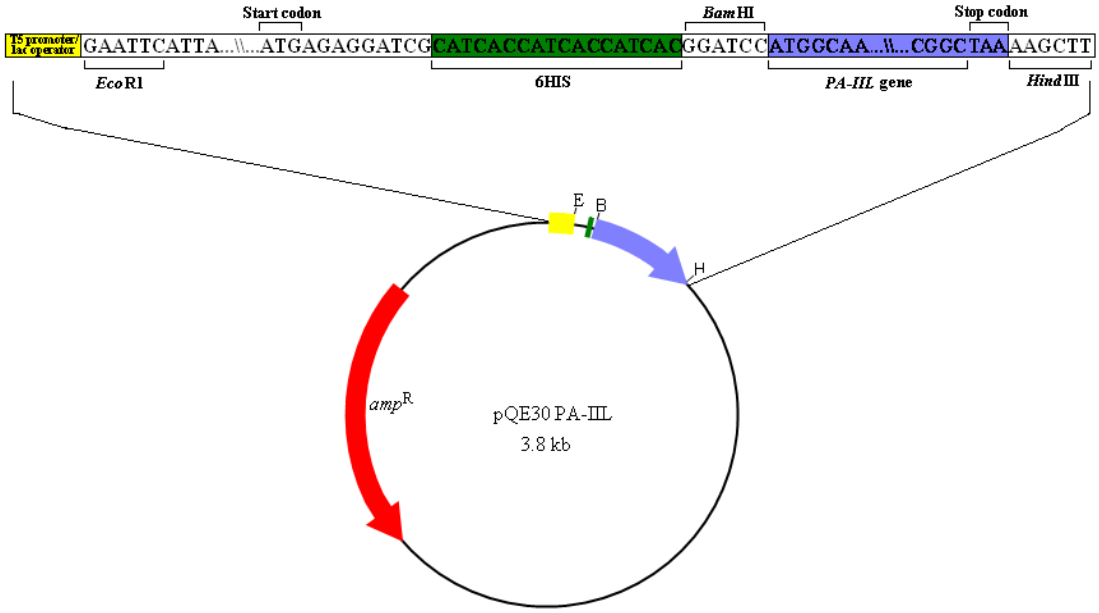


Figure 3.21: Schematic of pQE30.PAIIL construct.
Emphasised region of the pQE30.PAIIL vector outlines the relative positions of the *PAIIL* gene, the relevant restriction sites, start and stop codon and 6HIS. Image generated using BVtech.

3.4.4 Cloning for the expression of PA-IIL with a cleavable N-terminal 6HIS tag

The *PA-IIL* gene was cloned into the Qiagen pQE expression vector pQE30Xa (Table 2.3, Figure 2.5). The vector pQE30Xa features the same promoter operator system and β-lactamase gene conferring resistance to ampicillin (amp^R) as the pQE60 vector. The pQE30Xa vector enables the translation of PA-IIL protein with an N-terminal 6HIS affinity tag that can be removed after purification via specific digestion by FactorXa protease (Figure 3.22).

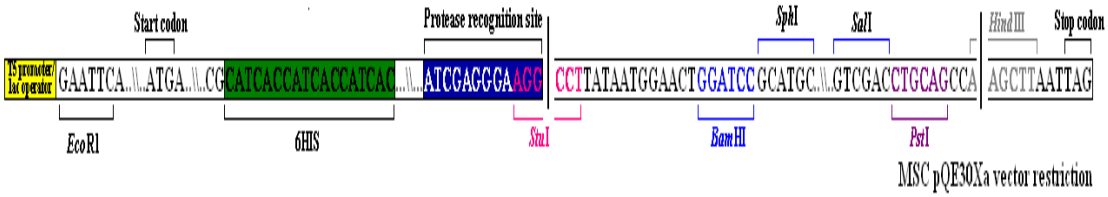


Figure 3.22: The MCS of the pQE30Xa vector
Schematic outlining the positions of the relevant restriction sites, protease recognition site, 6HIS tag and start and stop codons found within the MCS of the expression vector pQE30Xa.

The primers PAIIL-f2 and QEPAIIL-r were designed for the amplification of the *PA-IIL* gene as a 354 b.p *Hind*III fragment by PCR. The primed region of the *PA-IIL* gene for which the primers were designed is shown in Figures 3.23 and 3.24.

PAIIL-f2: ATG GCA ACA CAA GGA GTG TTC ACC C

QEPAIIL-r: (AAG CTT) - TTA GCC GAG CGG CCA GTT G

Hind III

Figure 3.23: Primers used for the amplification of the *PA-IIL* gene from *P. aeruginosa* PA01 for cloning into the pQE30Xa vector.
The enzymatic restriction sites added to the *PA-IIL* gene are shown underlined and labelled

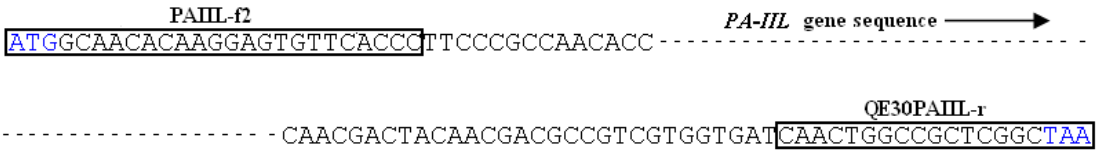


Figure 3.24: Primed region of the *PA-IIL* sequence from pQE30.PAIIL for suitable amplification for cloning into the pQE30Xa vector.
The start and stop codon of the *PA-IIL* gene highlighted in blue.

Plasmid DNA was prepared from the pQE30.PAIIL plasmid and used as template DNA for the amplification of the *PA-IIL* gene. The conditions used for the PCR are shown in Table 3.4 below. The PCR was carried out with the high fidelity *Taq* polymerase, Phusion*Taq*.

Table 3.4: PCR conditions for the amplification of the *PA-IIL* gene for cloning to the pQE30Xa vector.
Primers PAIIL-f2 and QEPAIIL-r used for amplification as described above.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	15 sec
Extension Time	30 sec
No. of Cycles	50

The PCR product was cleaned using Gel/PCR DNA fragment extraction kit. The cleaned PCR product was restricted with *Hind*III and the pQE30Xa vector was restricted with *Stu*I/*Hind*III and the restricted products were cleaned (Figure 3.25). Restriction with *Stu*I creates a blunt end to which the blunt end of the PCR product can ligate resulting in the loss of the *Stu*I restriction site.

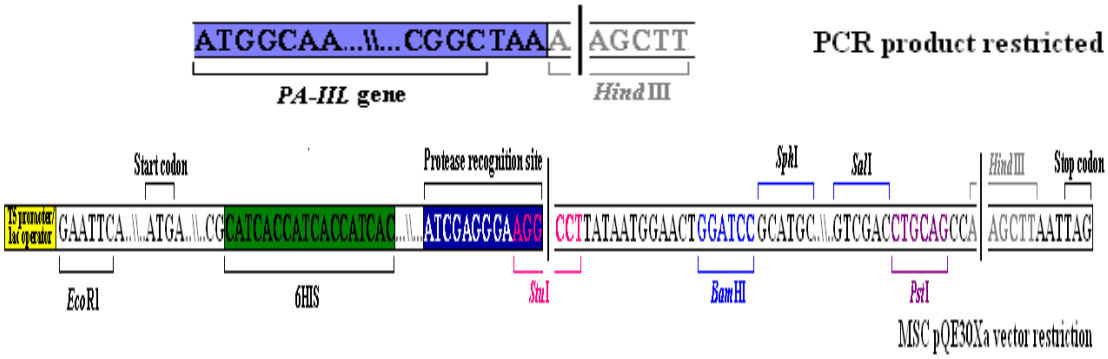


Figure 3.2515: Schematic of *Hind*III restricted PCR product amplified by PAIIL2-f and QE30.PAIIL-r and schematic of MCS of pQE30Xa vector restricted by *Stu*I/*Hind*III.
The point of restriction is indicated by a black line and break in the sequence.

The linearised pQE30Xa plasmid DNA was ligated with the purified *PA-IIL Hind*III fragment. The ligated DNA was transformed to *E. coli* XL10-Gold and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were screened by performing *Stu*I, *Eco*RI, *Hind*III and *Eco*RI/*Hind*III double digests on plasmid DNA prepared from each clone. The restriction digests were visualised on 0.7% agarose gels. Screening in this manner resulted in the identification of the desired plasmid named pQE30Xa.PAIIL. Figure 3.26 below shows the results of the verifying restriction analysis on the plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pQE30Xa.PAIIL uncut
3	pQE30Xa.PAIIL cut *Stu*I
4	pQE30Xa.PAIIL cut *Eco*RI
5	pQE30Xa.PAIIL cut *Hind*III
6	pQE30Xa.PAIIL cut *Eco*RI/*Hind*III
7	QE30Xa.*PAIIL* PCR product

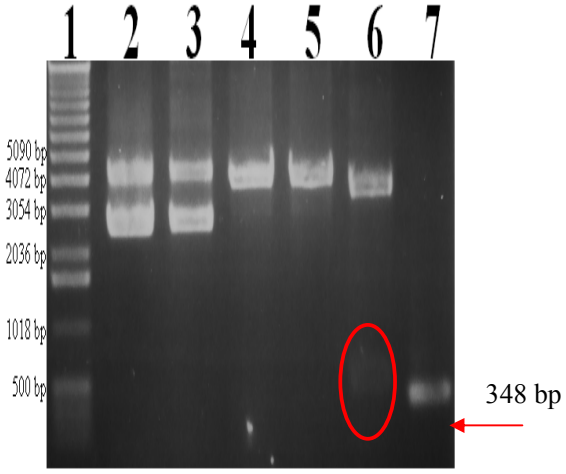


Figure 3.26: Analysis and restriction of the pQE30Xa.PAIIL vector.
The restriction digest of the pQE30Xa.PAIIL clone analysed on a 0.7% agarose gel (Section 2.7) found *Stu*I did not cut the plasmid which is indicative of the loss of the restriction site. Single restriction with *Eco*RI and *Hind*III produced a band corresponding to the expected (~3800 b.p) linearised plasmid. Double restriction with *Eco*RI/*Hind*II excised a band of ~350 b.p corresponding to the *PA-IIL* gene. The size of the excised band was confirmed by comparison to the PCR product produced by primers QE30Xa.PAIIL-f and QE30XaPAIIL-r. The insert is indicative of the loss of the *PA-IIL* fragment within the clone indicating that the cloning of *PA-IIL* to the pQE30Xa vector was successful.

The presence of the correct insert within the pQE30Xa.PAIIL vector (Figure 3.27) was further confirmed by DNA sequencing. Translation of protein from this expression vector results in recombinant PA-IIL with an N terminal 6HIS tag that can be removed after purification via specific digestion by FactorXa protease.

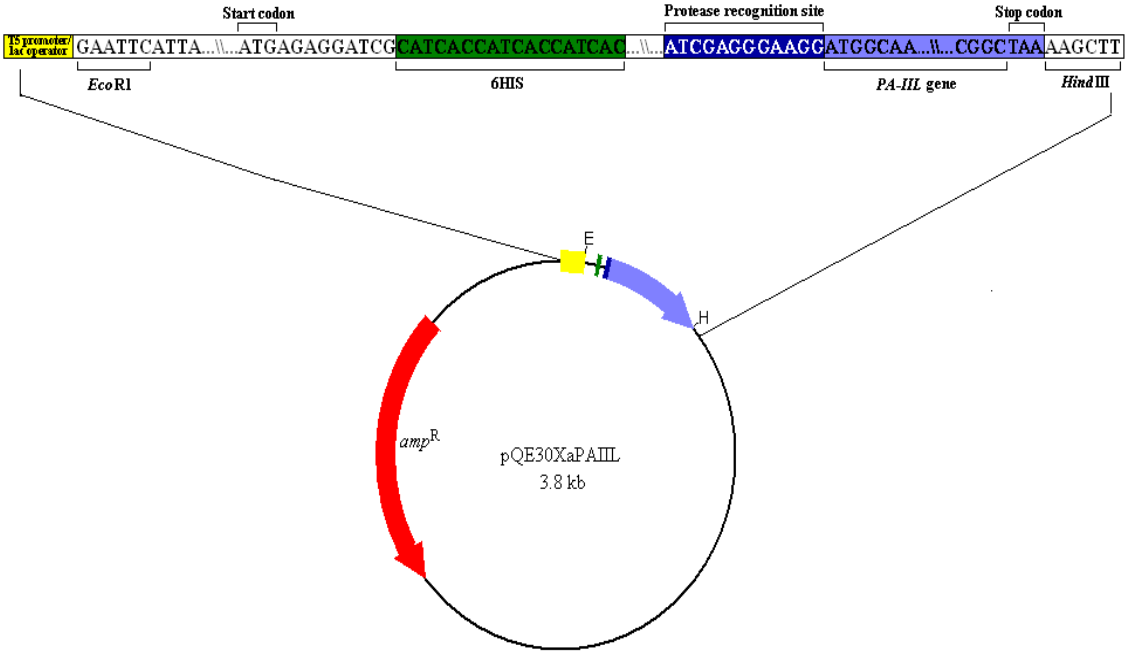


Figure 3.27: Schematic of pQE30Xa.PAIIL construct
Emphasised region of the pQE30Xa.PAIIL vector outlines the relative positions of the *PA-IIL* gene, relevant restriction sites, protease recognition site, start and stop codon and 6HIS. Image generated using BVtech.

3.4.5 Cloning for the expression of StrepII tagged *PA-IIL*

A vector that enables the translation of fusion protein having an N-terminal StrepII-tag was created in house based on the pQE30 vector. The 6HIS of pQE30.PAIIL were replaced by the StrepII tag sequence by site specific mutation. The primers PAIIL.StrepII-f and PAIIL.StrepII-r were designed for the incorporation of the StrepII sequence in place of 6HIS and the RGS-His epitope by mutagenesis. The presence of the RGS-HIS epitope enables the detection of 6HIS tagged protein by the RGS-His antibody eliminating the need for a protein-specific antibody. This feature is not required for this work and thus the additional bases of the RGS-His epitope were removed during mutagenesis. The primed region of the pQE30.PAIIL vector for which the primers were designed is shown in Figures 3.28 and 3.29.

PAIIL.StrepII-f (5' Phosphorylation) CAG TTC GAA AAG GGA TCC ATG GCA ACA CAA GGA GTG TTC ACC C

StrepII tag 3' terminus

PAIIL.StrepII-r (5' Phosphorylation) CGG GTG GCT CCA CAT AGT TAA TTT CTC CTC TTT AAT GAA TTC

StrepII tag 5' terminus

Figure 3.28: Primers used for the mutation of 6HIS of the pQE30.PAIIL vector to the StrepII tag sequence.
The StrepII sequence was divided between the forward primer and reverse primers. Phosphorylation of the 5' terminus of the primers facilitates ligation of the PCR product.

Figure 3.29: Primed region of the pQE30.PAIIL plasmid for the mutagenesis of the 6HIS tag to a StrepII tag to form the pQE30StrepII.PAIIL plasmid.
The start codon of the *PA-IIL* gene is highlighted in blue. The bases forming the RGS-His epitope site (highlighted in red) were eliminated by the primer sequences. The 6HIS to be mutated is highlighted in green.

Plasmid DNA was prepared from the pQE30.PAIIL vector and used as template DNA for the site specific mutagenesis. The conditions used for the PCR are shown in Table 3.5 below. The use of the high fidelity *Taq* polymerase, Phusion*Taq*, enabled full vector amplification.

Table 3.5: PCR conditions for the incorporation of StrepII tag in place of 6HIS in the pQE30.PAIIL vector.
Primers PAIIL.StrepII-f and PAIIL.StrepII-r used for amplification as described above. Elongated extension time enables full vector amplification.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	2min 30 sec
No. of Cycles	50

The PCR product (Figure 3.30) was cleaned using Gel/PCR DNA fragment extraction kit. The restriction enzyme *Dpn*I is effective against dam methylated DNA. Treatment of the PCR product with *Dpn*I leads to the cutting of the template pQE30.PAIIL while the un-methylated PCR product remains intact (see Section 2.12.1).

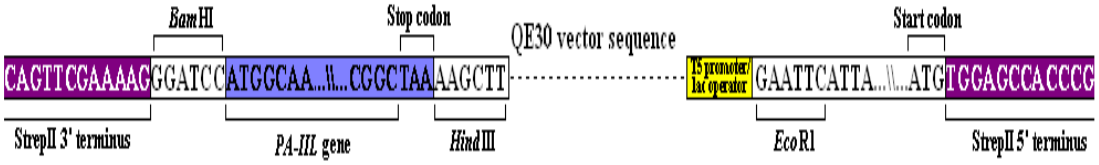


Figure 3.30: PCR product amplified by primers PAIIL.StrepII-f and PAIIL.StrepII-r, using the pQE30.PAIIL vector DNA as a template.
Schematic outlines the positions of the *PA-IIL* gene, the relevant restriction sites, StrepII tag and start and stop codons found within the MCS of the expression vector pQE30.Strep.PAIIL.

Phosphorylation of the primers facilitated ligation of the PCR product to form a plasmid which was then called pQE30.StrepII.PAIIL. The ligated DNA was transformed to *E. coli* XL10-Gold and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Restriction analysis cannot differentiate between re-circularised template DNA and pQE30.StrepII.PAIIL as all restriction sites are identical. The presence of the correct insert within the pQE30.StrepII.PAIIL vector (Figure 3.31) was confirmed by DNA sequencing. Translation of protein from this expression vector results in N-terminal StrepII tagged PA-IIL.

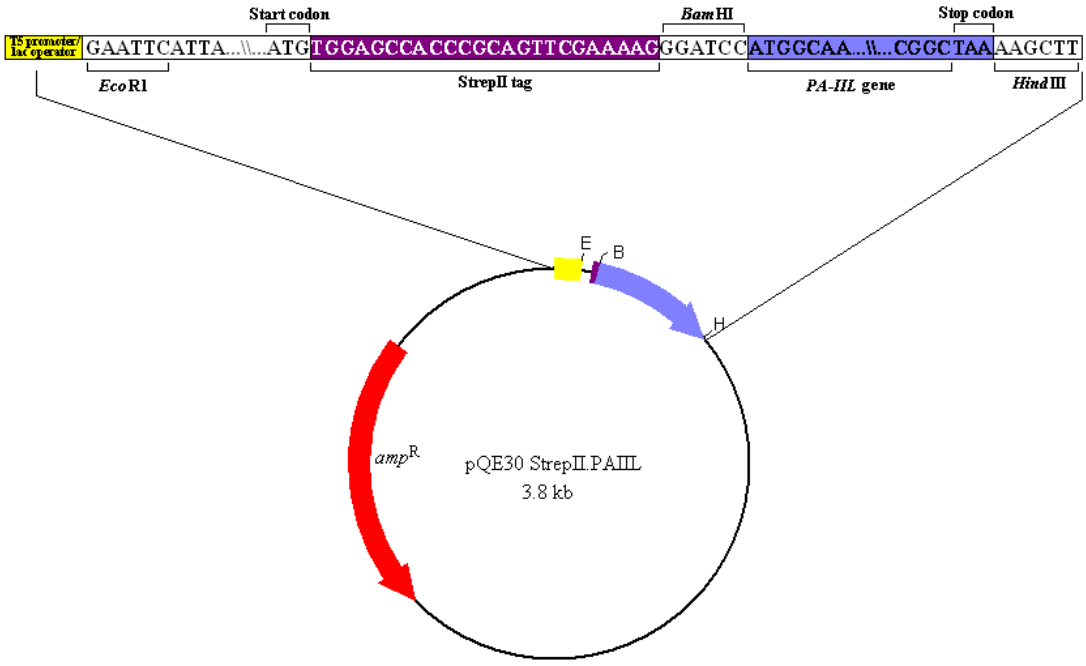


Figure 3.31: Schematic of the pQE30.StrepII.PAIIL construct
Emphasised region of the pQE30.StrepII.PAIIL vector outlines the relative positions of the *PA-IIL* gene, the relevant restriction sites, StrepII tag, start and stop codon. Image generated using BVtech.

3.4.5 Mutagenesis of the sugar binding loop of PA-IIL

A number of lectins of high similarity to PA-IIL have been identified and characterised, in particular the mannose binding lectin RS-IIL from *R. solanacearum.* Alignment of the amino acid sequence of PA-IIL and RS-IIL shows a 69% identity and 85% similarity between the two proteins (Figure 3.32).

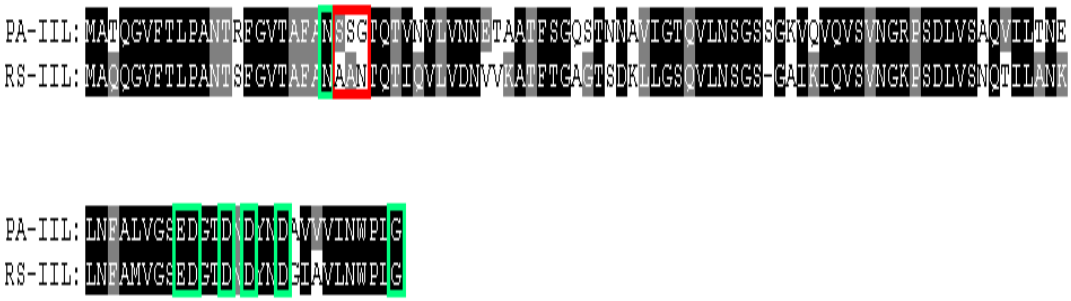


Figure 3.32: Amino acid alignment of protein sequences of *P. aeruginosa* lectin PA-IIL and *R. solanacearum* lectin RS-IIL.
Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Identical residues highlighted in black and similar in grey. Image generated using ClustalW align and Genedoc.

Significantly, the three amino acids, Ser22-Ser23-Gly24, identified as directly involved in the sugar binding activity of PA-IIL show little similarity to the three amino acids, Ala-22-Ala-23-Asn-24, identified as directly involved in the sugar binding activity of RS-IIL (Figure 3.32). The differing sugar specificities (PA-IIL preference – L-fucose, RS-IIL preference – D-mannose) of PA-IIL and RS-IIL are attributed to the variation across these three amino acids. Previous studies, as discussed in Section 1.6 (Adam *et al.* 2007), have point mutated each of the sugar binding amino acids of PA-IIL to those found in RS-IIL to investigate the effect on sugar specificity. While the sugar binding specificity was found to alter from a preference for fucose to mannose a binding affinity directly comparable to that found for RS-IIL was not obtained. In order in investigate the effect of the collective mutation of all three amino acids the primers PAIIL.MU-f and PAIIL.MU-r were designed to site specifically mutate the *PA-IIL* gene. A diagnostic restriction site, *Hpa*I, was also introduced by the primers into the gene by a silent point mutation to facilitate restriction analysis of colonies. The primed region of the pQE30Xa.PAIIL vector for which the primers were designed is shown in Figures 3.33 and 3.34.

PAIIL.MU-f: AAC GTG CTG GTC AAC AAC GAG ACG GCC GCG ACC
{5' Phosphorylated}

PAIIL.MU-r: AAC CGT CTG GGT GTT GGC CGC GTT GGC GAA GGC GGT CAC G
{5' Phosphorylated}

Figure 3.33: Primers used for the site specific mutation of the *PA-IIL* gene of pQE30Xa.PAIIL.
Mutated nucleotides highlighted in red. A silent point mutation for the incorporation of a *Hpa*I diagnostic restriction site highlighted in blue.

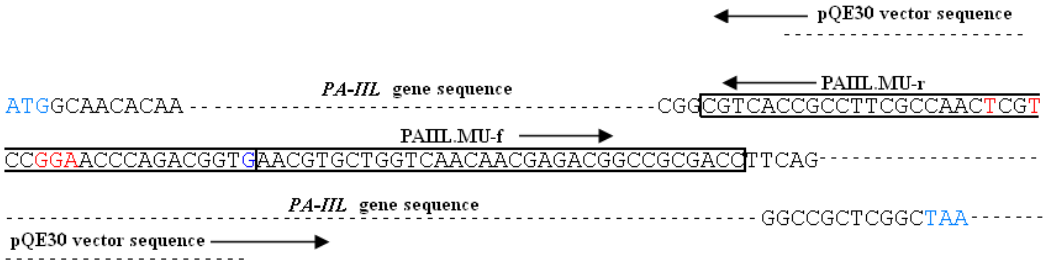


Figure 3.34: Primed region of the pQE30Xa.PAIIL plasmid for the specific mutagenesis of the *PA-IIL* gene.
The start and stop codons of the *PA-IIL* gene are highlighted in light blue. The nucleotides to be mutated are highlighted in red. The silent point mutation for the incorporation of the diagnostic restriction site *Hpa*I is highlighted in dark blue.

Plasmid DNA was prepared from the pQE30Xa.PAIIL vector and used as template DNA for the site specific mutagenesis. The conditions used for the PCR are shown in Table 3.6 below. The use of the high fidelity *Taq* polymerase, Phusion, enabled full vector amplification.

Table 3.6: PCR conditions for the mutation of the *PA-IIL* gene within the pQE30Xa.PAIIL vector.
Primers PAIIL.MU-f and PAIIL.MU-r used for the site specific mutagenesis as described above. Elongated extension time enables full vector amplification.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	2 min 30 sec
No. of Cycles	50

The PCR product (Figure 3.35) was cleaned and restricted with *Dpn*I which leads to the cutting of the template pQE30Xa.PAIIL while the un-methylated PCR product remains intact.

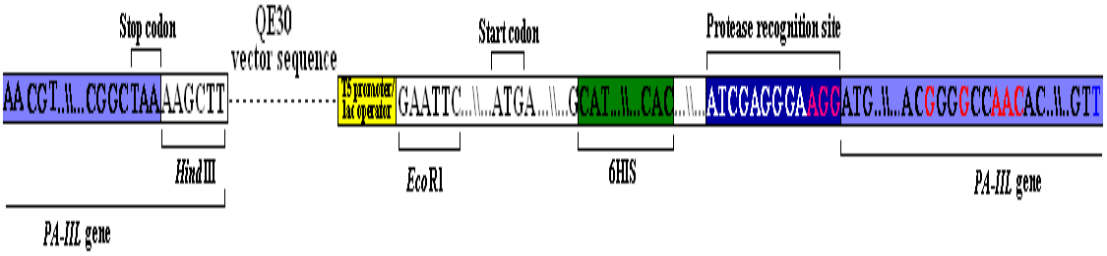


Figure 3.35: PCR product amplified by primers PAIIL.MU-f and PAIIL.MU-r, using the pQE30XA.PAIIL vector DNA as a template.
Schematic outlines the positions of the *PA-IIL* gene, the relevant restriction sites, 6HIS tag, protease site, mutated nucleotide bases and start and stop codons found within the MCS of the expression vector pQE30Xa.PAIILMU plasmid.

The restricted PCR product was again cleaned and the phosphorylation of the primers facilitated the ligation of the PCR product to form a plasmid which was then called pQE30Xa.PAIILMU. The ligated DNA was transformed to *E. coli* XL10-Gold and the transformation was plated on LBamp media (as outlined in Section 2.9.3). The introduction of the unique restriction site, *Hpa*I, provides a means of differentiating between pQE30Xa.PAIILMU and re-circularised pQE30Xa.PAIIL. Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were screened by performing *Hpa*I digests on plasmid DNA prepared from each clone. The restriction digests were visualised on a 0.7% agarose gel. Screening in this manner resulted in the identification of the desired plasmid named pQE30Xa.PAIILMU. Figure 3.36 below shows the results of the verifying restriction analysis on the plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pQE30Xa.PAIILMU uncut
3	pQE30Xa.PAIILMU cut *Hpa*I

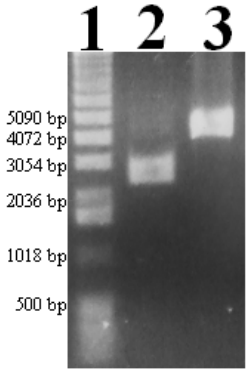


Figure 3.36: Analysis and restriction of the pQE30Xa.PAIILMU vector
The restriction digest of the pQE30Xa.PAIILMU clone analysed on a 0.7% agarose gel (Section 2.7) found that restriction with *Hpa*I resulted in linearised plasmid. The confirmation of the presence of the *Hpa*I restriction site is indicative of the successful mutation of the *PA-IIL* gene.

The presence of the correct mutations within the pQE30Xa.PAIIL.MU vector was confirmed by DNA sequencing (Figure 3.37).

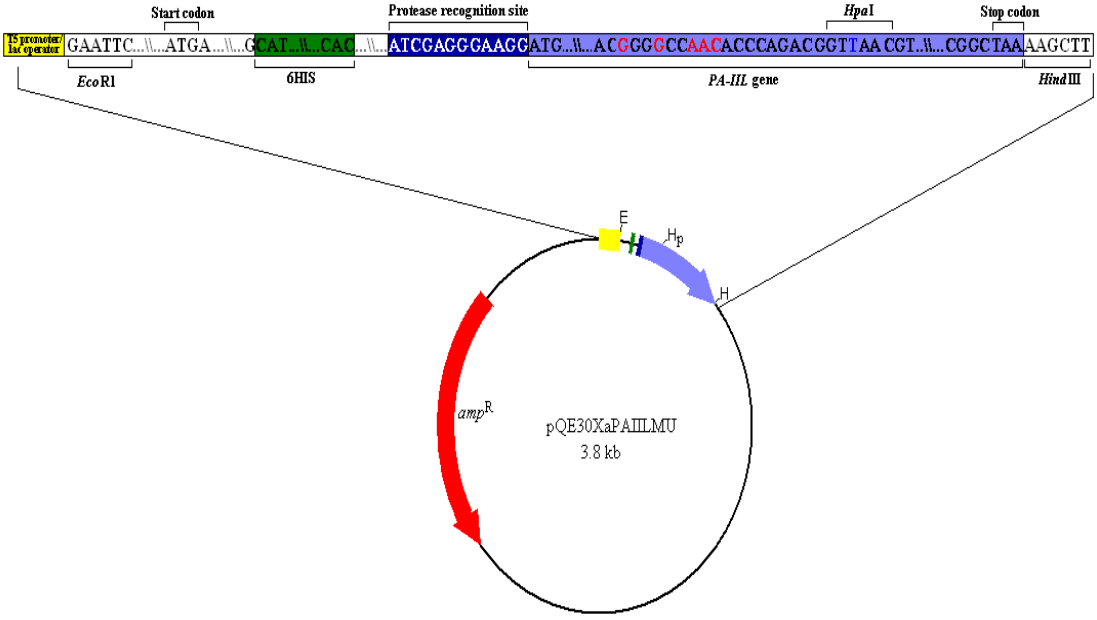


Figure 3.37: Schematic of pQE30Xa.PAIILMU construct
Emphasised region of the pQE30Xa.PAIILMU vector outlines the relative positions of the relevant restriction sites, protease recognition site, start and stop codon and 6HIS. The mutated nucleotide bases are highlighted in red. Image generated using BVtech.

Translation of protein from this expression vector results in N-terminal 6HIS tagged PA-IIL with a mutated sugar binding loop as shown in Figure 3.38.

```
PAIIL   : AFANSSGTQTV
PAIILMU : AFANAANTQTV
```

Figure 3.38: The mutated sugar binding loop found for protein expressed from the pQE30Xa.PAIILMU vector.
Upper read; original PA-IIL sequence as expressed from pQE30Xa.PAIIL. Lower read; PA-IIL with 3 point mutation within the sugar binding loop as expressed from pQE30Xa.PAIILMU. Image created with GeneDoc (Section 2.11).

3.5 Preliminary characterization of recombinant PA-IIL

The pPAIIL3 and QE30PAIIL plasmids were transformed into *E. coli* XL-10 Gold and grown up in LBamp media (as outlined in Section 2.9.3). The culturing conditions for both plasmids were as described in Section 2.15. Samples were taken 4 hours after induction with 50 µM isopropyl-β-thiogalactopyranoside (IPTG). Samples were lysed following the procedure described in methods and materials. Figure 3.39 shows SDS-PAGE analysis of the soluble protein fractions. Prominent bands corresponding to an over expressed protein running at the expected MW of 6HIS C-tag PA-IIL of 12,929 Da

and of 6HIS N-tag PA-IIL of 14,062 Da as deduced from their amino acid sequences can be seen (Figure 3.39).

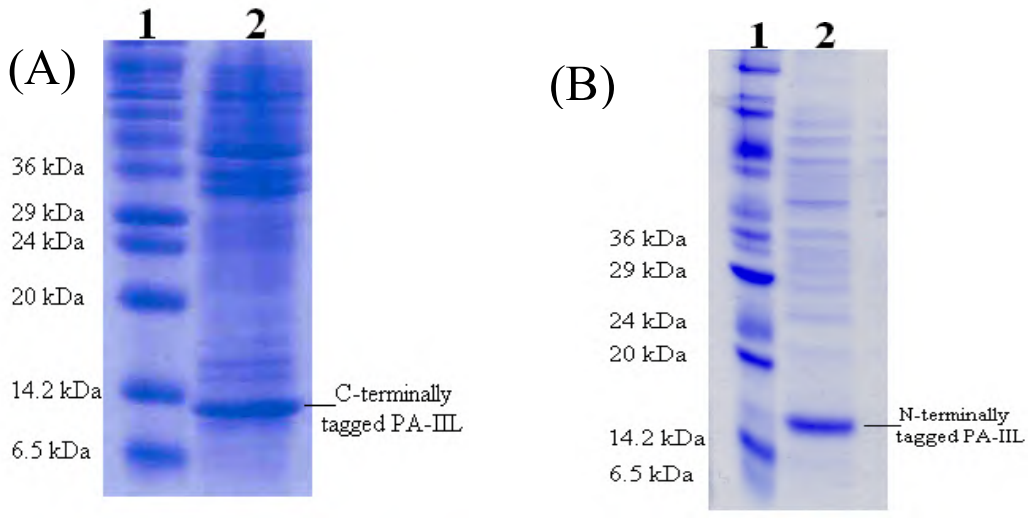


Figure 3.39: Expression of C-terminally and N terminally tagged PA-IIL in *E .coli*.
Analysis of cleared lysate samples by 20% SDS-PAGE (Section 2.24). (**A**): Lane 1, M_r (sizes as in Figure 2.8); Lane 2, C-terminally tagged PA-IIL (expressed from pPA-IIL3 in *E. coli* XL10-Gold as outlined in Section 2.12). (**B**): Lane 1, M_r (sizes as in Figure 2.8); Lane 2, N-terminally tagged PA-IIL (expressed from pQE30Xa.PAIIL in *E. coli* XL10-Gold as outlined in Section 2.12).

3.5.1 IMAC purification of recombinant PA-IIL

The addition of a 6HIS tag allows PA-IIL to be purified by single-step immobilised metal affinity chromatography, IMAC (Nilsson *et al.* 1997, Porath *et al.* 1975). Purification of 6HIS tagged lectins through IMAC involves exploitation of the strong interaction between 6HIS residues in the introduced tag and the two ligand binding sites of the Ni^{2+} ion, which has been immobilised on a solid matrix, in this case nitrilotriacetic acid linked to sepharose/sephadex (Ni-NTA). Many cellular proteins display histidine residues on their surface which may loosely bind to Ni-NTA resin. Washing with low concentrations of imidazole, the active side chain of histidine, displaces loosely bound proteins. High concentrations of imidazole are required to displace 6HIS tagged recombinant protein.

Protein was expressed from pPAIIL3 and pQE30.PAIIL, *E. coli* XL10-Gold, in 100ml LBamp cultures. Samples were lysed following the procedure described in the methods and materials. The elution profiles of the cleared lysate samples of 6HIS C-tag and 6HIS N-tag PA-IIL were established within the standard lysis buffer (Section 2.3), by carrying out a series of washes supplemented with increasing concentrations of

imidazole. Figure 3.40 shows SDS-PAGE analysis of the highly purified 6HIS N and C tagged PA-IIL produced by IMAC purification.

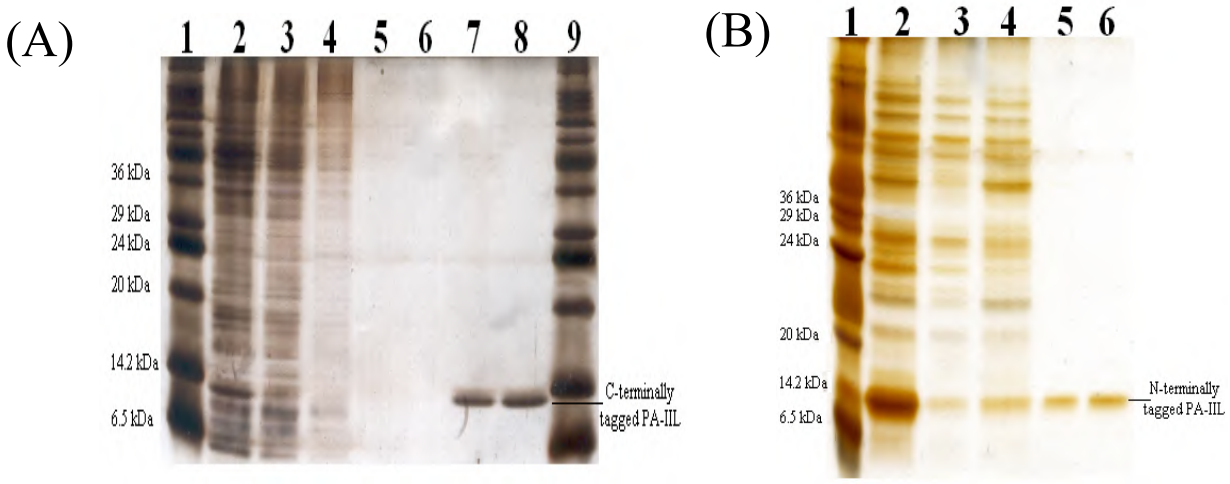


Figure 3.40: Analysis by 20% SDS page of PA-IIL purification on Nickel-NTA resin
Analysis by 20% SDS-PAGE of fractions collected from purification on a nickel-NTA resin of a 100ml *E. coli* lysate containing PA-IIL.(A) Analysis by 20% SDS-PAGE Silver stained (Section 2.24) of recombinant 6HIS C-tagged PA-IIL. Lane 1, Mr (sizes given in Figure 2.8); Lane 2, Clear lysate; Lane 3, flow-through; Lane 4, wash 1 (80mM imidazole); Lane 5 wash 2 (80mM imidazole); Lane 6, wash 3 (120mM imidazole) Lane 7, 350mM elution; Lane 8, 350mM elution, Lane 9, Mr. (B) Analysis by 20% SDS-PAGE Silver stained (Section 2.24) of recombinant 6HIS N-tagged PA-IIL purification Lane 1, Mr (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, wash 1 (30mM imidazole); Lane 5, 350mM elution; Lane 6, 350mM imidazole elution.

3.5.2 Biological activity of PA-IIL

PA-IIL has a high affinity for L-fucose and D-mannose as discussed, see Chapter 1 (Garber *et al.* 1987). Horse radish peroxidase (HRP) is a mannose bearing compound, therefore the ability of PA-IIL to bind to HRP is indicative of sugar specific lectin activity. The impact of the addition of 6HIS on the activity of PA-IIL was investigated using the semi-quantitative HRP binding assay (see Section 2.26). This assay was used to compare the activity of 6HIS N-tagged and C-tagged PA-IIL. Figure 3.41 shows a histogram of the PA-IIL activity.

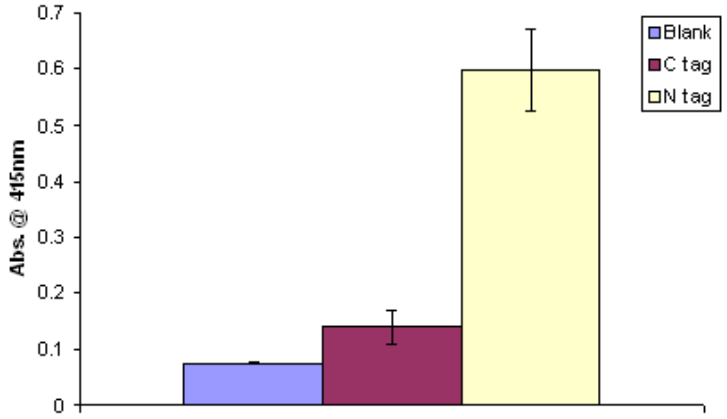


Figure 3.41: Histogram of PA-IIL activity assessed by Horse radish peroxidase (HRP) assay.
HRP assay carried out as per Section 2.26. Histogram ; PBS used in place of PA-IIL as a negative control (blue). 10 µg/ml of 6HIS N-tagged (yellow) and 6HIS C-tagged (purple) PA-IIL per well. (Error bars represent mean ± SD, n = 3).

There was a 1.84 fold increase in activity of 6HIS C-tag PA-IIL over the signal for the control compared to a 7.9 fold increase for 6HIS N-tag PA-IIL. Therefore 6HIS N-tag PA-IIL has a 4.3 fold greater activity then 6HIS C-tag PA-IIL. The addition of the 6HIS residues to the C-terminus of PA-IIL can therefore be said to disrupt binding activity. Addition of 6HIS residues to the N-terminus of PA-IIL appeared not to disrupt binding activity.

3.5.3 Relative molecular mass

To investigate the quaternary structure of the recombinant lectins size exclusion chromatography (SEC) was employed using the G-100 Sephadex gel filtration matrix (Whitaker 1963). Size exclusion chromatography separates proteins based on molecules hydrodynamic radius. Proteins of lower relative molecular mass are retarded within the porous resin beads and are thus separated from proteins of larger relative molecular mass which are excluded from the small pores. The larger proteins will pass through the packed column more rapidly than the smaller proteins and their exact size can be determined by comparison with a standard curve created using a set of known standards (Figure 3.42). The standards include BSA, Carbonic Anhydrase, Cytochrome C and Ovalbumin. Equation 3.1 was used to create the standard curve.

Equation 3.1: Formula for the calculation of the K_{av}

$$K_{av} = \frac{V_e - V_o}{V_t - V_o}$$

V_e - the elution volume of the protein standard
V_o - the void volume
V_t - the total volume

The K_{av} was then plotted against the log of the MW of each of the standard proteins to create a standard curve, see Figure 3.42. The elution profile of each protein on the column can be seen in Figure 3.43.

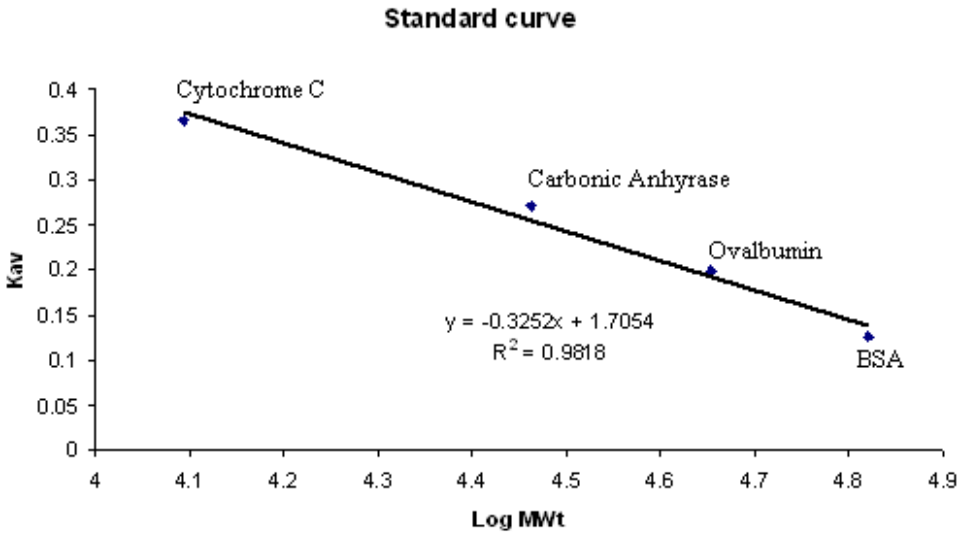


Figure 3.42: Development of size exclusion chromatography standard curve
The protein molecular weight standard curve was created using size exclusion chromatography with a G-100 Sephadex column (Section 2.25.1). Void volume (V_o) of 26.4ml was determined from the elution volume of blue dextran, and total volume (V_t) was 132ml, determined by elution of bromophenol blue. The elution volumes of the protein standards were as follows; BSA (39.6 ml), Carbonic Anhydrase (55 ml), Ovalbumin (47.3 ml) and Cytochrome C (64.9 ml) as determined by readings at OD_{280}. These values were used to construct a plot of K_{av} versus log MW. This could be in turn used to calculate the relative MW of all lectin samples that were eluted in the MW range of the standards.

Table 3.7: Construction of a protein molecular weigh standard curve for the G-100 Sephadex column at pH 7.8 in PBS.

Standard	Elution Vol.	Expected MW	logMW	Kav	Actual MW	% Error
BSA	39.6 ml	66000	4.82	0.13	72407	9.71
Ovalbumin	47.3 ml	45000	4.65	0.20	43207	-3.98
Carbonic anhydrase	55 ml	29000	4.46	0.27	25783	-11.09
Cytochrome C	64.9 ml	12400	4.09	0.36	13276	7.06

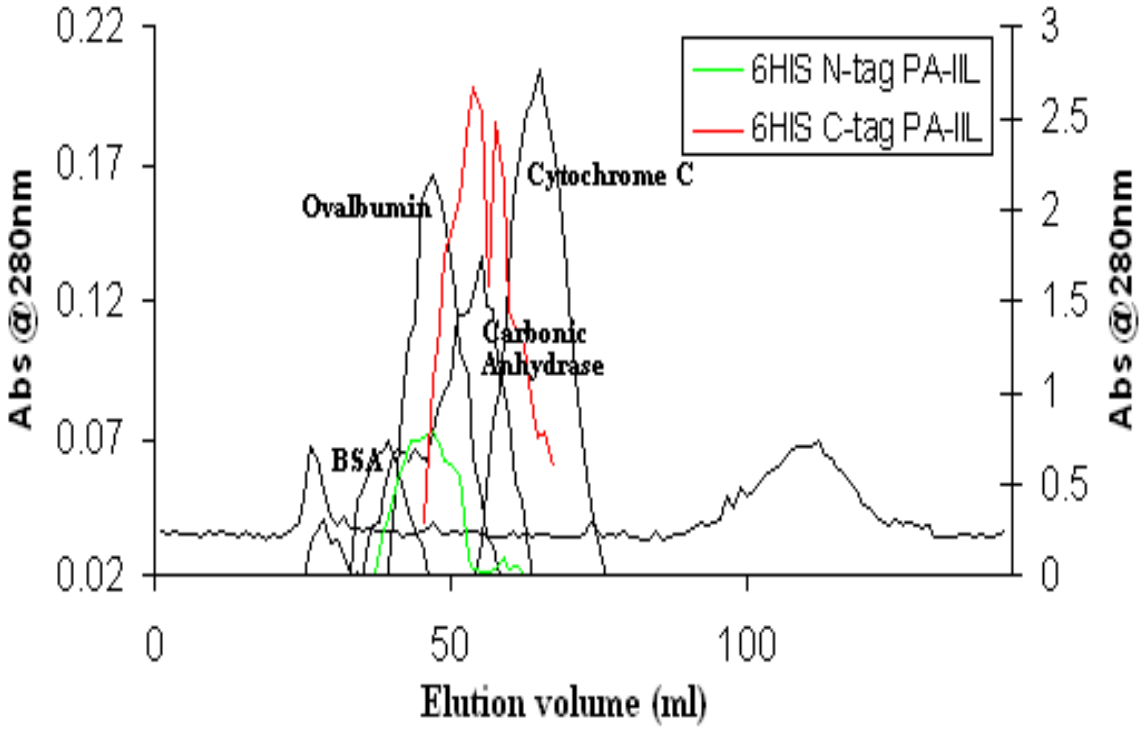


Figure 3.43: Elution of protein standards on a G-100 Superdex column.
The elution profiles of standard proteins using size exclusion chromatography with a G-100 Sephadex column. The optical density at OD_{280nm} of eluted protein is plotted against the elution volume. The elution peak volumes were used in the construction of the standard curve (Figure 3.42). The elution profile of 6HIS N-tag PA-IIL is highlighted in green. The elution profile of 6HIS C-tag PA-IIL is highlighted in red.

As outlined in Section 2.25.1, purified 6HIS C-tagged and 6HIS N-tagged PA-IIL, was analysed by size exclusion chromatography. In each instance the lectin was suspended in PBS, pH7.8, in order to ascertain its multimeric structure under native conditions. Samples were applied to the column in 250µl volumes at a concentration of approx. 5 mg/ml, and the peak heights recorded were converted into relative molecular weights using the stand curve created (see Figure 3.42)

Table 3.8: Calculated native MW of recombinant PA-IIL from a G-100 Sephadex column.
The theoretical monomer size of the subunits of C-tag and N-tag differ due to addition of a protease cleavage site and a RGS sequence before the 6HIS N-tag, while no such additions are present on the 6HIS C-tag.

Lectin	Elution Vol.	Theoretical Monomer Size (Da)	Calculated Molecular Weight
6HIS N-tag PA-IIL	46.8 ml	14,062	44,680
6HIS C-tag PA-IIL	54 ml	12,929	27,571

The calculated molecular weights for each recombinant lectin are provided in Table 3.8. Each lectin was run on the column separately and their elution profiles are plotted in Figure 3.43. The 6HIS C-tag PA-IIL molecule was found to form a dimer (expected MW of dimer 26kDa). The calculated molecular weight of the 6HIS N-tag PA-IIL molecule was found to be between a trimer and tetramer (expected MW of a trimer 42kDa, expected MW of a tetramer 56.5kDa). PA-IIL is a small compact protein that would be likely to diffuse through the column at a slower rate then a globular protein of the same molecular weight (similar to the globular standards used for the creation of the standard curve) and therefore would elute at a greater volume then predicted (Dubin, Principi 1989). Therefore 6HIS N-tagged PA-IIL is most likely to be a tetramer.

3.6 Optimisation of expression conditions for recombinant expression of PA-IIL in *E. coli*

Having elucidated in the previous section that the addition of 6HIS to the C-terminus of PA-IIL disrupts the sugar binding activity of the lectin and disrupts the native tetrameric structure, addition of 6HIS to the C-terminus was deemed unsuitable. The expression conditions for 6HIS N-tagged PA-IIL were further optimised.

3.6.1 Selection of an *E. coli* strain for recombinant expression of PA-IIL

E. coli is the most commonly used host for recombinant protein expression. It facilitates protein expression by its relative simplicity, rapid high density growth rate on inexpensive substrates and well characterized genetics which allow for large scale protein production. As the recombinant lectin being investigated is prokaryotic in source, the question of eukaryotic PTM requirement in not a consideration in the choice of an expression host. The merits of different *E. coli* genotypes for the expression of PA-IIL were investigated and compared for optimal expression of the PA-IIL protein. The expression plasmid pQE30.PA-IIL, which expressed 6HIS N-tagged PA-IIL (see Section 3.4.2) was transformed into various *E. coli* strains, namely XL10-Gold, BL21 (DE3), JM109 and XL1-Blue (Table 2.1). XL1-Blue and JM109 were used routinely in this work for the cloning and propagation of plasmid constructs. BL21 (DE3) is a strain deficient in the OmpT protease, purportedly allowing higher recovery of heterologous recombinant proteins. XL10-Gold contains F factors, which carry the *lac*I^q allele. The

lacI^{q} allele is a promoter mutation that expresses the LacI repressor protein at high levels, resulting in strong repression of the T5 promoter/*lac* operon in the absence of IPTG. The effect of individual clonal selection was investigated by the methodology outlined in Section 2.15.2 (Figure 3.44).

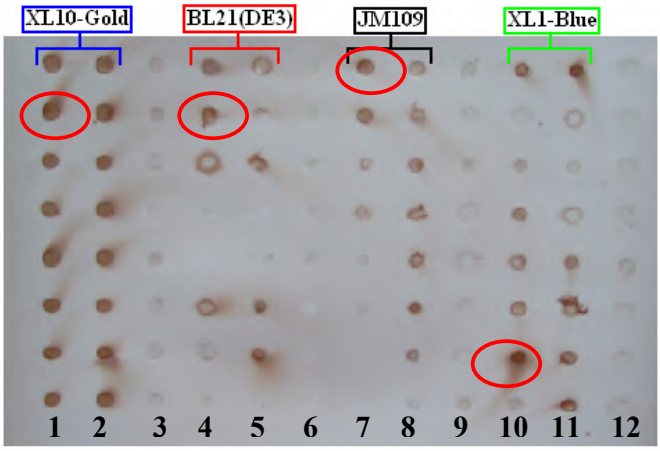


Figure 3.44: Effect of colony selection on lectin expression in *E. coli*
Colony blot of pQE30.PAIIL plasmid expressed in *E. coli* by the methodology as described in Section 2.15.2. Lane 1 and 2, pQE30PAIIL *E. coli* XL10-Gold; Lane 3, pQE30 *E. coli* XL10-Gold; Lane 4 and 5, pQE30PAIIL *E. coli* BL21 (DE3): Lane 6, pQE30 *E. coli* BL21 (DE3): Lane 7 and 8, pQE30PAIIL *E. coli* JM109; Lane 9, pQE30 *E. coli* JM109; Lane 10 and 11, pQE30PAIIL *E. coli* XL1-Blue; Lane 12, pQE30 *E. coli* XL1-Blue. Highlighted in red are individual clones that appeared to express higher levels of protein then their counter parts and were picked for expression in 10ml of LBamp

A high expressing candidate clone from each strain of *E. coli* was picked. 10ml expression cultures of each clone chosen were grown in LBamp broth containing 50µM IPTG at 37^{o}C overnight. Figure 3.45 shows the SDS-PAGE analysis of the soluble and insoluble fractions of PA-IIL expression from each *E. coli* strain. Clear bands corresponding to an over expressed protein could be seen corresponding to the expected MW of 13,261 Da in all four strains of *E. coli*.

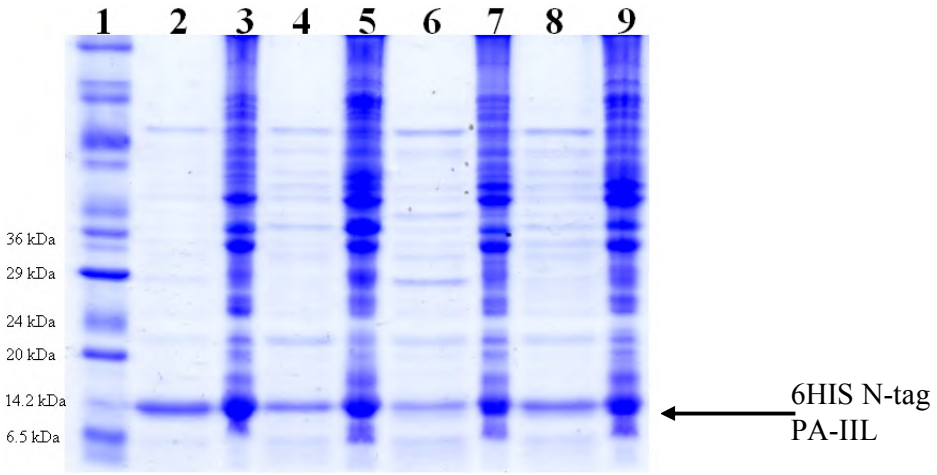


Figure 3.45: Expression of N- terminal 6HIS tagged PA-IIL from *E. coli* strains
Expression carried out from 10ml LBamp cultures as described in Section 2.15.Lane 1, Mr (sizes as in Figure 2.8); Lane 2, pQE30PAIIL *E. coli* XL10-Gold soluble fraction; Lane 3, pQE30PAIIL *E. coli* XL10-Gold insoluble fraction; Lane 4, pQE30PAIIL *E. coli* BL21 (DE3) soluble fraction: Lane 5, pQE30PAIIL *E. coli* BL21 (DE3) insoluble fraction: Lane 6, pQE30PAIIL *E. coli* JM109 soluble fraction; Lane 8, pQE30PAIIL *E. coli* JM109 insoluble fraction; Lane 9, pQE30PAIIL *E. coli* XL1-Blue soluble fraction; Lane 10, pQE30PAIIL *E. coli* XL1-Blue insoluble fraction.

3.6.2 Effect of temperature and time on recombinant expression of PA-IIL

Protein was seen to be over expressed in all four *E. coli* strains (Figure 3.45). *E. coli* XL10-Gold and BL21 (DE3) are common expression strains. Further optimisation of PA-IIL expression was therefore carried out using *E .coli* XL10-Gold and BL21 (DE3). The optimal temperature for expression and the optimal harvest point of PA-IIL were investigated by examining the soluble cell extract at set points post-induction when cultures were grown at 37^{o}C, 30^{o}C and grown to the point of induction at 37^{o}C while protein expression was then carried out at 30^{o}C. Expression of PA-IIL was carried out by the methodology outlined in Section 2.15. Cultures were induced with 50µM IPTG when an OD_{600} of 0.5 -0.6 was reached. An initial sample (T_0) was taken prior to induction and samples were subsequently taken every hour after induction for 6 hours. Cultures were then grown overnight after which a final sample was taken. Samples were lysed following the procedure described in Section 2.15.1 and the soluble protein fractions were analysed by SDS-PAGE as shown in Figures 3.46 and 3.47.

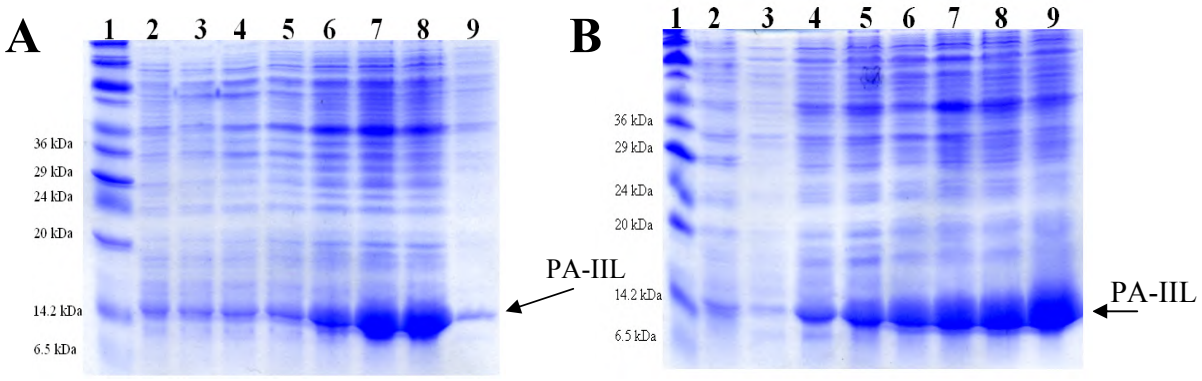


Figure 3.46: Expression of N-terminal 6HIS tagged PA-IIL in *E. coli* BL21 (DE3)
Analysis of the expression of the protein PA-IIL from the plasmid pQE30.PAIIL in *E. coli* BL21 (DE3) by 20% SDS-PAGE. (A) Expression at 37°C; Lane 1, Protein marker; Lane 2, cell extract at point of induction; Lane 3, 1 hr; Lane 4, 2 hrs; Lane 5, 3hrs; Lane 6, 4 hrs; Lane 7, 5 hrs; Lane 8, 6 hrs; Lane 9, overnight culture. (B) Growth at 37°C and expression at 30°C; Lane 1,Protein marker; Lane 2, cell extract point of induction; Lane 3, 1 hr; Lane 4, 2 hrs; Lane 5, 3hrs; Lane 6, 4 hrs; Lane 7, 5 hrs; Lane 8, 6 hrs; Lane 9, overnight culture.

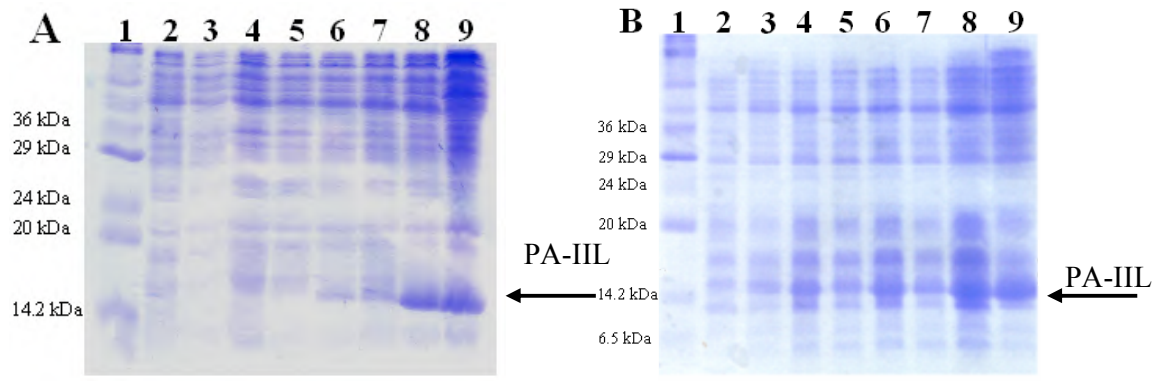


Figure 3.47: Expression of N-terminal 6HIS tagged PA-IIL in *E. coli* XL10-Gold
Analysis of the expression of the protein PA-IIL from pQE30PA-IIL in *E. coli* XL10-Gold by 20% SDS-PAGE. (A) Growth and induction at A_{600} of 0.5 – 0.6, expression at 37°C; (B) growth at 37°C and induction at A_{600} of 0.5, expression at 30°C. Lane 1, M_r (sizes given in Figure 2.7); Lane 2, point of induction; Lane 3, 1 hr induction; Lane 4, 2 hrs induction, Lane 5, 3 hrs induction; Lane 6, 4 hrs induction; Lane 7, 5 hrs induction, Lane 8, 6 hrs induction, Lane 9, 24 hrs induction. Location of PA-IIL is indicated by arrows.

A decrease in protein yield was noted when PA-IIL was expressed from *E. coli* BL21 (DE3) at 37°C overnight. Little or no protein was also found to be expressed by cultures grown at 30°C (gels not shown). The growth rate of the cultures at 37°C, 30°C and the 37°C and 30°C combination, during the exponential phase, were measured by OD_{600} as shown in Figure 3.48. There was no significant variation in growth rates between those cultures found to express soluble protein and those that did not, namely those expressed at 30°C. This indicated that the lack of protein expression was not due to cell death. The

failure to over express soluble PA-IIL at 30°C was found regardless of *E. coli* strain utilised for expression.

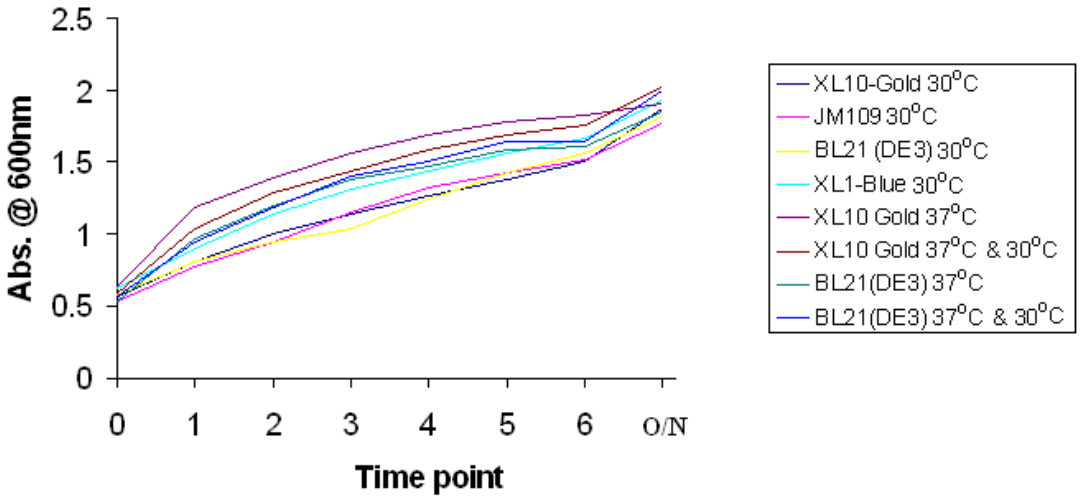


Figure 3.48: Growth curve for expression of N-terminal 6HIS tagged PA-IIL from pQE30.PAIIL in *E. coli*
Absorbance readings at OD_{600} taken during expression of PA-IIL from pQE30.PAIIL in *E. coli* BL21, XL10-Gold, JM109 and XL1-Blue with induction using 50uM IPTG. Temperature conditions varied between growth and expression at 37°C or 30°C and growth at 37°C to induction point and expression at 30°C.

The failure to express soluble PA-IIL at 30°C and the decreased yield found for soluble PA-IIL expressed from BL21 (DE3) overnight was attributed to the aggregation of PA-IIL when expressed at high levels. Studies such as Berrow, N. *et al.* (2006), have found that lower temperature tends to be more effective for the expression of soluble protein. Expression of PA-IIL at 30°C was likely to have led to greater PA-IIL yields resulting in aggregation. BL21 (DE3) was shown to be a leaky expressor, as PA-IIL expression is visible in Figure 3.46A and B at time point 0 (Lane 2) prior to induction. Leaky expressing of PA-IIL may also have lead to accumulation of a high yield of protein resulting in aggregation. *E. coli* XL10-Gold contains F factors carrying the *lac*I^q allele conferring greater control of protein expression. Strong control of protein expression by XL10-Gold led to consistent protein expression, as shown in Figure 3.44 by expression of protein for every XL10-Gold colony investigated by the colony blot method. Expression from *E. coli* XL10-Gold produced suitable yields of PA-IIL for future work while preventing excess protein production leading to aggregation and loss of soluble protein. Therefore XL10-Gold was used for all subsequent expression of PA-IIL.

3.6.3 Recovery of PA-IIL from the insoluble fraction

E. coli XL10-Gold was elucidated in the previous section to be the optimal strain for the production of PA-IIL. PA-IIL was found to be present in both the soluble and insoluble fractions as shown in Figure 3.45. A study by Tsumoto, K. *et al.* (2003) has shown that insoluble proteins that are at least in part in their native conformation may be solubilized under non-denaturing conditions to yield active protein by the addition of L-arginine. Therefore, an attempt was made to solubilize PA-IIL from the insoluble fraction using L-arginine. An expression culture was prepared and lysed by the methodology outlined in Section 2.15. The cell pellet (i.e. the insoluble protein fraction) formed by centrifugation following lysis was retained and resuspended in lysis buffer supplemented with L-arginine hydrochloride at various concentrations and left mixing overnight at 4^{o}C. The mixture was centrifuged at 13,000 rpm for 10 min and the supernatant was removed. The SDS-PAGE analysis of the supernatant and the remaining protein within the insoluble fraction is shown in Figure 3.49.

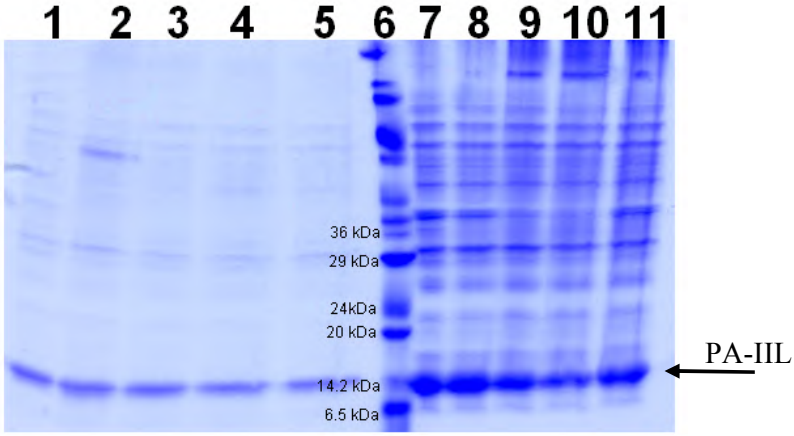


Figure 3.49: SDS PAGE analysis of the effect of L-arginine on the insoluble fraction of PA-IIL
Analysis of L-arginine treatment of the insoluble PA-IIL fraction. Lane 1, supernatant from insoluble fraction treated with no L-arginine; Lane 2, supernatant from insoluble fraction treated with 0.5M L-arginine; Lane 3, supernatant from insoluble fraction treated with 1M L-arginine; Lane 4, supernatant from insoluble fraction treated with 1.5M L-arginine; Lane 5, supernatant from insoluble fraction treated with 2M L-arginine; Lane 6, M_r (sizes given in Figure 2.8); Lane 7, insoluble fraction treated with no L-arginine; Lane 8, insoluble fraction treated with 0.5M L-arginine; Lane 9, insoluble fraction treated with 1M L-arginine; Lane 10, insoluble fraction treated with 1.5M L-arginine; Lane 11, insoluble fraction treated with 2M L-arginine

L-arginine had no greater effect on the levels of protein recovered from the insoluble fraction than lysis buffer (Figure 3.49) therefore this is not a suitable method to recover PA-IIL from the insoluble fraction.

3.6.4 Effect of IPTG concentration on recombinant expression of PA-IIL

Lectin expressed from plasmids derived from the pQE range of expression plasmids are under the control of the T5 promoter/lac operator element system, one of the most widely used expression systems. This promoter/operator system consisting of the coliphage T5 promoter (recognized by the *E. coli* RNA polymerase) and the *lac* repressor operator sequence which binds the *lac* repressor and ensures efficient repression of the T5 promoter. Expression of T5 promoter/lac operator is repressed by the LacI protein. The *lacI*^{q} allele, present in the *E. coli* XL10-Gold strain, is a promoter mutation that increases the intracellular concentration of LacI repressor, resulting in the strong repression. Addition of IPTG, the inducer molecule, inactivates the LacI repressor. Thus, the amount of expression from the T5 promoter/lac operator is proportional to the concentration of IPTG added. Therefore low concentrations of IPTG result in relatively low expression from the T5 promoter/lac operator element and high concentrations of IPTG result in high expression from the T5 promoter/lac operator element. The IPTG concentration was varied in PA-IIL expression and the amount of soluble protein product expressed from T5 promoter/lac operator element was examined by SDS-PAGE (Figure 3.50).

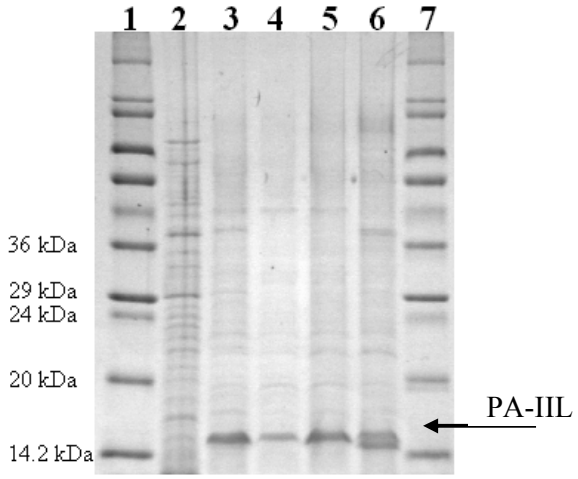


Figure 3.50: Effect of varying IPTG concentration on the expression of PA-IIL in *E. coli* XL10-Gold
Analysis of increasing IPTG concentration on the expression of recombinant PA-IIL by 20% SDS-PAGE. Lane 1, Protein marker, Lane 2, induction with 0 µM IPTG. Lane 3, induction with 50 µM IPTG. Lane 4, induction with 100µM IPTG. Lane5, induction with 250 µM IPTG. Lane 6, induction with 500 µM IPTG. Lane 7, Protein marker.

It was deduced from the SDS-PAGE analysis that there was no significant difference between the yield of soluble protein produced upon induction with 50 µM IPTG and

any higher concentration of the compound. Cultures could be induced by higher concentrations of IPTG, however to save on reagent 50 µM of IPTG was concluded to be the optimal induction concentration.

3.7 Purification of recombinant PA-IIL

Purification of histidine tagged lectins through IMAC involves the exploitation of the strong interaction between the six histidine residues of the introduced tag and the two ligand binding sites of the Ni^{2+} ion of the Ni-NTA resin. Many cellular proteins display histidine residues on their surface, several wash steps are therefore employed using imidazole, the active side-chain of histidine, to displace loosely bound proteins. A high concentration is then employed to displace the 6HIS tagged recombinant protein.

3.7.1 Optimisation of IMAC purification

The elution profile of 6HIS N-tagged PA-IIL was determined using lysis buffer (Section 2.3) supplemented with increasing concentrations of imidazole up to 350 mM imidazole. The resulting profile is shown in Figure 3.51.

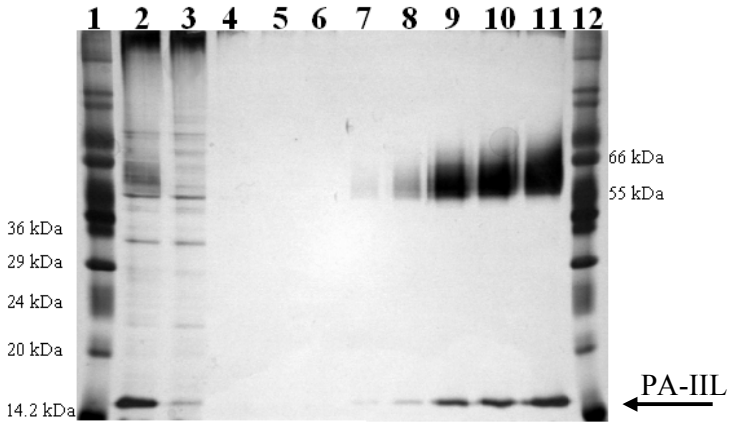


Figure 3.51: SDS gel analysis of wash optimisation of PA-IIL silver stain.
Analysis by 15% SDS-PAGE Silver stained (Section 2.24) of fractions collected from the purification on Amersham nickel resin, of cell lysate from a 100ml culture of *E. coli* expression PA-IIL protein. Lane 1, Protein ladder; Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole; Lane 5, 40mM imidazole; Lane 6, 50mM imidazole; Lane 7, 80mM imidazole; Lane 8, 100mM imidazole; Lane 9, 150mM imidazole; Lane 10, 350mM imidazole elution 1; Lane 11, 350mM imidazole elution 2; Lane 12, Mr.

From the SDS-PAGE analysis PA-IIL was found to remain bound to the column in the presence of imidazole washes up to 50 mM. Washing with a 30 mM imidazole was found to sufficiently remove all loosely bound proteins. The standardised procedure to

purify PA-IIL was established as; two 10 ml 30 mM imidazole wash steps followed by elution with 5 ml 350 mM imidazole buffer (Figure 3.52).

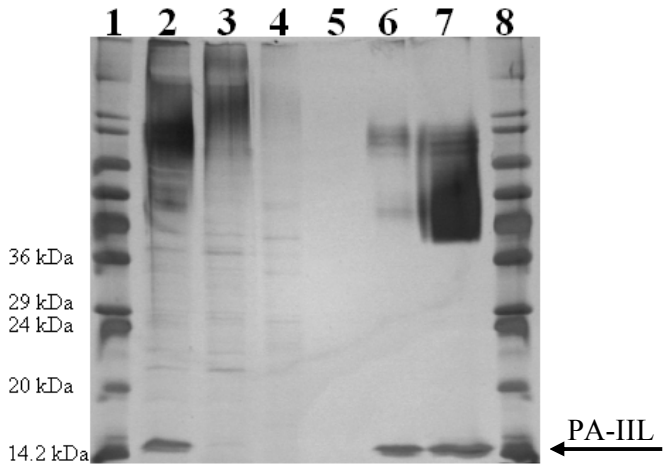


Figure 3.52: SDS gel analysis of purified PA-IIL
Analysis by 20% SDS-PAGE Silver stained (Section 2.24) of recombinant N-terminally tagged PA-IIL purification using IMAC (Section 2.13.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash 1; Lane 5, 30mM imidazole wash 2; Lane 6, 350mM imidazole elution 1; Lane 7, 350mM imidazole elution 2; Lane 8, Mr Ladder.

SDS analysis reveals PA-IIL forming multimers, appearing as large bands and smears, when PA-IIL is highly expressed and purified (see lanes 6 and 7 of Figure 3.52). Elongation of boiling time when purified protein was prepared for SDS-PAGE analysis (Section 2.24) to 20 minutes resulted in the loss of these bands confirming that they are due to multirmer formation and not contamination.

The average yield of protein from a 100ml culture was found to be approximately 1 -2 mg/ml. Expression of PA-IIL was upscaled to a 500ml culture, with the purification protocol upscaled in accordance. Ni-NTA resin bed increased to 1ml and 30 mM wash volume increased to 30 ml (see Figure 3.53). The average yield of protein from a 500ml culture was found to be approx. 4 -5 mg/ml.

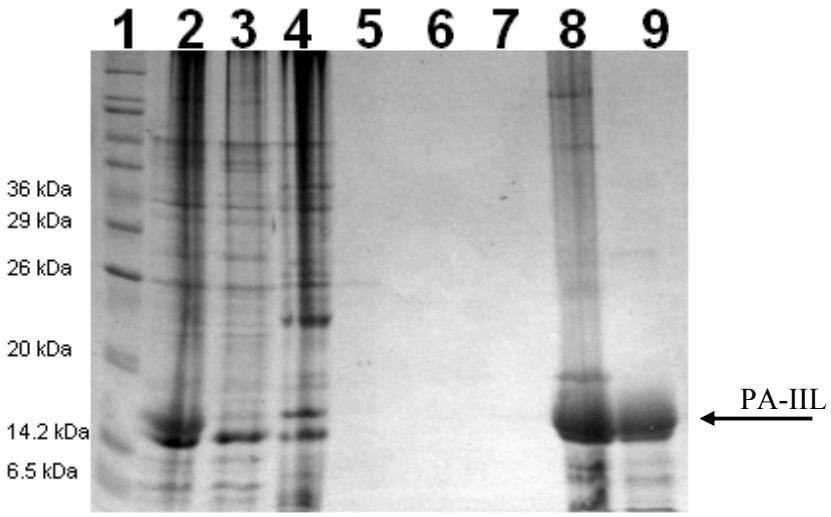


Figure 3.53: SDS gel analysis of large scale PA-IIL purification
Analysis by 15% SDS-PAGE (Section 2.24) of recombinant N-terminally 6HIS tagged PA-IIL purification using IMAC (Section 2.16.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash 2; Lane 5, 30mM imidazole wash 10; Lane 6, 30mM imidazole wash 20; Lane 7, 30mM imidazole wash 30; Lane 8, 350mM imidazole elution 1; Lane 9, 350mM imidazole elution 2.

3.8 Recombinant expression of PA-IIL in *E. coli* from other constructs

The optimised procedures established for the expression of PA-IIL from pQE30.PAIIL were used to express PA-IIL from a number of constructs based on the same pQE vectors and utilising the T5 promoter/lac operator system.

3.8.1 Expression of recombinant PA-IIL in *E. coli* from pQE60 construct

The effect of the *E. coli* strain and the effect of individual clonal selection on the expression of untagged PA-IIL (from pQE60.PAIIL, see Section 3.3) was investigated by the colony blot method (Section 2.15.2). A high expressing candidate clone from each strain of *E. coli* was picked (see Figure 3.54A). 10ml expression cultures of each clone chosen were grown in LBamp broth containing 50µM IPTG at 37^{o}C overnight. Figure 3.54B shows the SDS-PAGE analysis of the soluble and insoluble fractions of PA-IIL expression from each *E .coli* strain. Clear bands corresponding to an over expressed protein could be seen corresponding to the expected MW of untagged PA-IIL at approximately 12 kDa for *E .coli* XL10-Gold and JM109.

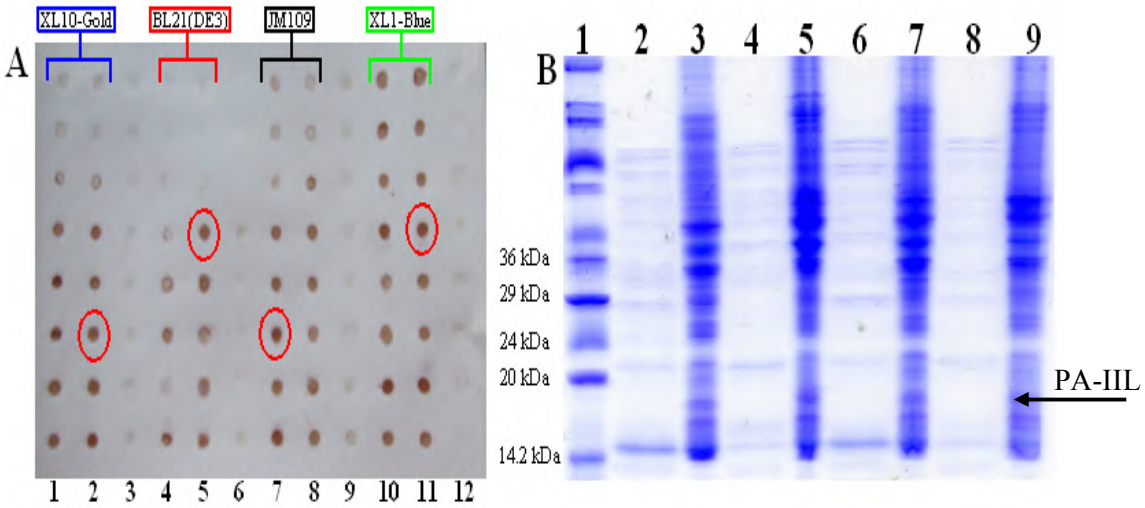


Figure 3.54: Colony selection of untagged PA-IIL expressed from pQE60.PAIIL from *E. coli*
[A] Lane 1 and 2, pQE60PA-IIL *E. coli* XL10-Gold; Lane 3, pQE60 *E. coli* XL10-Gold; Lane 4 and 5, pQE60PA-IIL *E. coli* BL21 (DE3): Lane 6, pQE60 *E. coli* BL21 (DE3): Lane 7 and 8, pQE60PA-IIL *E. coli* JM109; Lane 9, pQE60 *E. coli* JM109; Lane 10 and 11, pQE60PA-IIL *E. coli* XL1-Blue; Lane 12, pQE60 *E. coli* XL1-Blue. Highlighted in red are clones that appeared to express higher levels of protein and were picked for expression in 10ml LBamp[B] Lane 1, Mr (sizes as in Figure 2.8); Lane 2, pQE60PA-IIL *E. coli* XL10-Gold soluble fraction; Lane 3, pQE60PA-IIL *E. coli* XL10-Gold insoluble fraction; Lane 4, pQE60PA-IIL *E. coli* BL21 (DE3) soluble fraction: Lane 5, pQE60PA-IIL *E. coli* BL21 (DE3) insoluble fraction: Lane 6, pQE60PA-IIL *E. coli* JM109 soluble fraction; Lane 8, pQE60PA-IIL *E. coli* JM109 insoluble fraction; Lane 9, pQE60PA-IIL *E. coli* XL1-Blue soluble fraction; Lane 10, pQE60PA-IIL *E. coli* XL1-Blue insoluble fraction.

The high expressing untagged PA-IIL candidate clone from *E. coli* XL10 Gold was then expressed at 100ml using the optimised expression conditions established for QE30.PAIIL expression (Section 3.6.2).

3.8.1.1 Purification of recombinant untagged PA-IIL

The lectin PA-IIL was originally purified using mannose agarose (Gilboa-Garber 1982). This method exploited the proteins natural affinity for the carbohydrate D-mannose to separate it from other proteins present in the cytoplasm of *P. aeruginosa*. This technique (see Section 2.17) was carried out on the cell extract of a 100 ml culture of *E. coli* expressing the lectin from the pQE60.PAIIL plasmid (Section 3.3). Figure 3.55 shows the SDS-PAGE analysis of the successful purification of untagged PA-IIL over mannose agarose. 0.3M mannose was used to elute the bound protein.

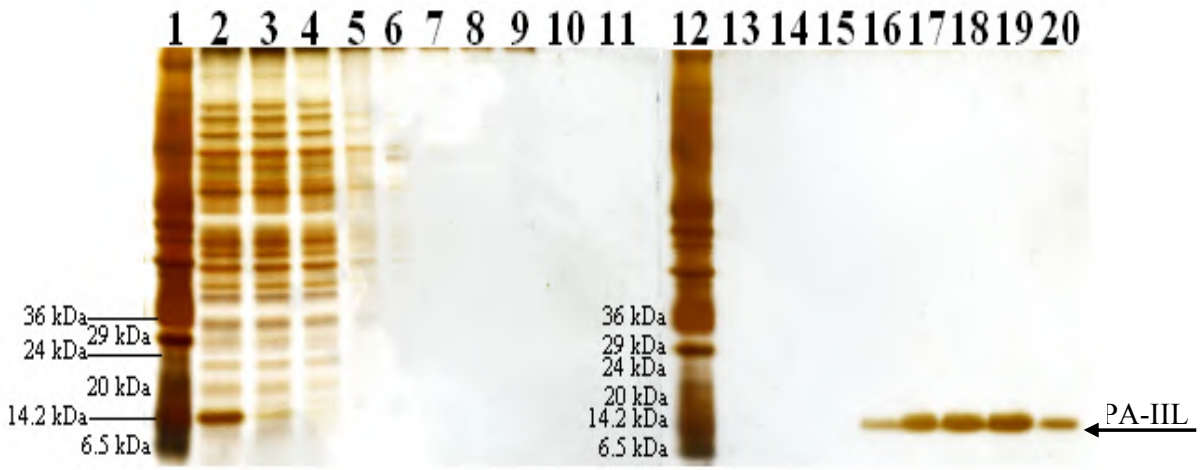


Figure 3.55: Purification of untagged Pa-IIL by affinity chromatography using mannose agarose.

Analysis by 15% SDS-PAGE, silver stained, of recombinant untagged PA-IIL purification by affinity chromatography using mannose agarose (Section 2.17). Lane 1, Molecular weight marker; Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4-11 potassium phosphate wash 1 - 6; Lane 12, Mr Ladder ; Lane 13-15, potassium phosphate wash 7-10; Lane 16-20, 0.3M elution fraction 1-5.

3.8.2 Expression of recombinant PA-IIL in *E. coli* from pQEStrep construct

The effect of the *E. coli* strain and the effect of individual clonal selection on the expression of StrepII tagged PA-IIL was investigated by the colony blot method (Section 2.15.2). A high expressing candidate clone from each strain of *E. coli* was picked (see Figure 3.56A). 10ml expression cultures of each clone chosen were grown in LBamp broth containing 50µM IPTG at 37^{o}C overnight. Figure 3.56B shows the SDS-PAGE analysis of the soluble and insoluble fractions of PA-IIL expression from each *E .coli* strain. Clear bands corresponding to an over expressed protein could be seen corresponding to the expected MW of approximately 13.2 kDa from all four *E. coli* strains.

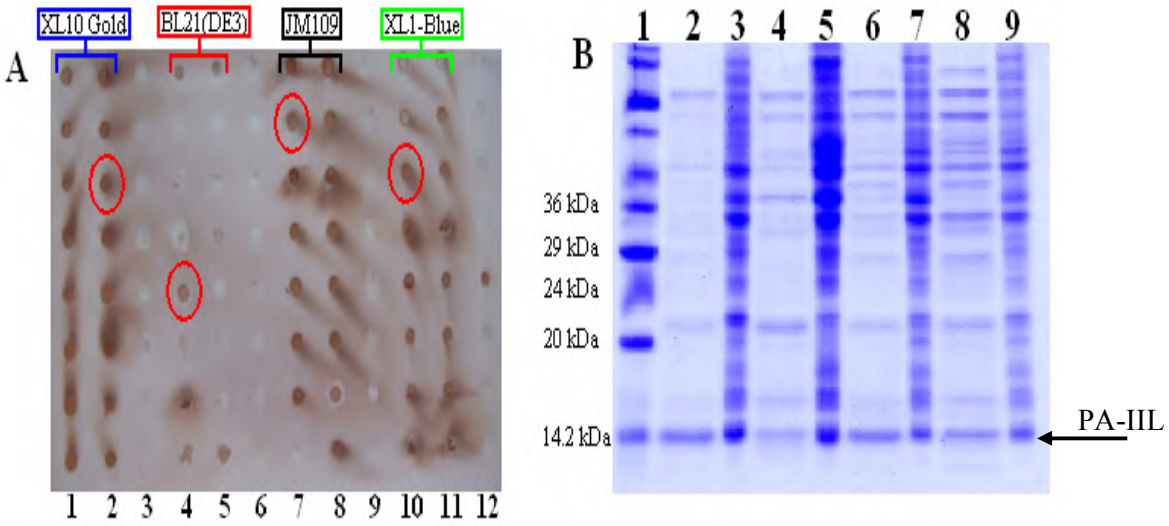


Figure 3.56: Colony selection of StrepII tagged PA-IIL from pQEStrepII.PAIIL from *E. coli*
[A] Lane 1 and 2, pQE30.StrepPA-IIL *E. coli* XL10-Gold; Lane 3, pQE30.Strep *E. coli* XL10-Gold; Lane 4 and 5, pQE30.Strep PA-IIL *E. coli* BL21 (DE3): Lane 6, pQE30.Strep *E. coli* BL21 (DE3): Lane 7 and 8, pQE30.StrepPA-IIL *E. coli* JM109; Lane 9, pQE30.Strep *E. coli* JM109; Lane 10 and 11, pQE30.StrepPA-IIL *E. coli* XL1-Blue; Lane 12, pQE30.Strep *E. coli* XL1-Blue. [B] Lane 1, Mr (sizes as in Figure 2.8); Lane 2, pQE30.StrepPA-IIL *E. coli* XL10-Gold soluble fraction; Lane 3, pQE30.StrepPA-IIL *E. coli* XL10-Gold insoluble fraction; Lane 4, pQE30.StrepPA-IIL *E. coli* BL21 (DE3) soluble fraction: Lane 5, pQE30.StrepPA-IIL *E. coli* BL21 (DE3) insoluble fraction: Lane 6, pQE30.StrepPA-IIL *E. coli* JM109 soluble fraction; Lane 8, pQE30.StrepPA-IIL *E. coli* JM109 insoluble fraction; Lane 9, pQE30.StrepPA-IIL *E. coli* XL1-Blue soluble fraction; Lane 10, pQE30.StrepPA-IIL *E. coli* XL1-Blue insoluble fraction.

The high expressing StrepII tagged PA-IIL candidate clone from *E. coli* XL10 Gold was then expressed at 100 ml using the optimised expression conditions established for QE30.PAIIL expression (Section 3.6.2).

3.8.2.1 Purification of StrepII recombinantly tagged PA-IIL

StrepTrap HP columns (GE Healthcare) are pre-packed StrepTactin sepahrsoe columns that facilitate one-step-purification of StrepII tagged proteins. StrepII tagged PA-IIL was purified over a StrepTrap HP column as outlined in Section 2.18. Figure 3.57 shows the SDS-PAGE analysis of the successful purification of StepII tagged PA-IIL.

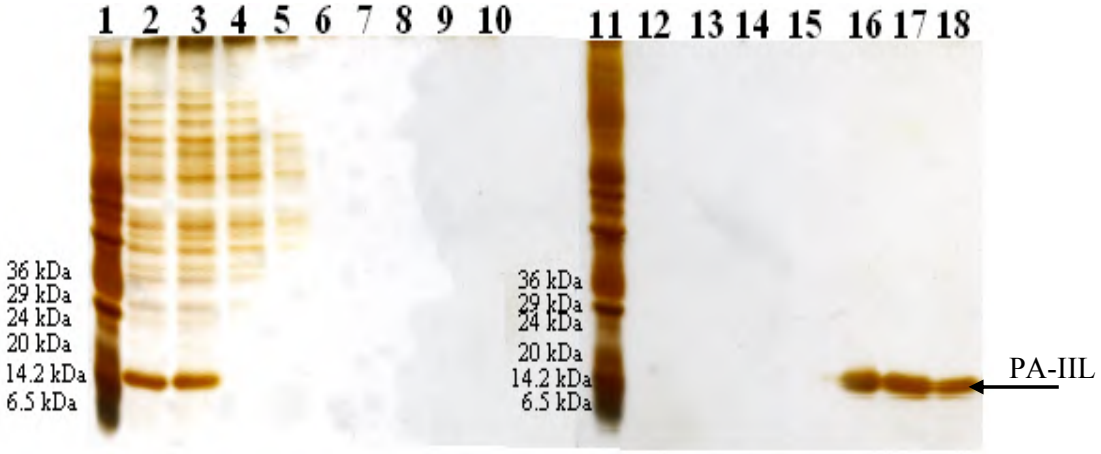


Figure 3.57: Purification of StrepII tagged PA-IIL by affinity chromatography
Analysis by 15% SDS-PAGE Silver stained (Section 2.24) of recombinant StrepII tagged PA-IIL, purification over StrepTrap™ HP column (Section 2.18). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, filtered clear lysate; Lane 4, flow-through; Lane 5-10 potassium phosphate wash 1 - 6; Lane 11, Mr Ladder ; Lane 12-15, wash 7-10; Lane 16-18, 2.5mM desthiobiotin elution fraction 1-3.

3.8.3 Expression and purification of pQE30PAIILMu1

The expression plasmid pQE30.PAIIL was site-specifically mutated (Section 3.4.5) to create the pQE30.PAIIL.MU plasmid. Figure 3.58 shows the expression and purification of the mutated PA-IIL using the expression and purification procedures established for the expression of PA-IIL from pQE30.PAIIL (Section 3.6).

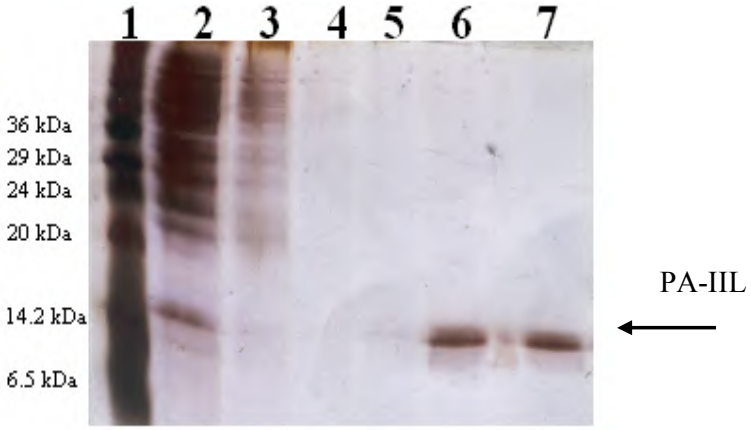


Figure 3.58: Purification of PA-IIL mutant from pQE30PAIILMU by affinity chromatography using IMAC
Analysis by 20% SDS-PAGE, silver stained (Section 2.24), of recombinant 6HIS N-tagged PA-IILMU purification using IMAC (Section 2.16.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash 1; Lane 5, 30mM imidazole wash 2; Lane 6 and 7, 350mM imidazole elution 1 and 2.

3.8.4 Activity of recombinant PA-IIL from pQE constructs.

The impact of the addition of each tag to PA-IIL was investigated using the semi-quantitative HRP assay (Section 2.26). The assay is based on the ability of PA-IIL to sugar specifically bind to mannose bearing HRP. To ensure any signal obtained was due to sugar specific activity 0.3 M fucose was added to competitively inhibit PA-IIL binding, therefore quenching any signal produced by specific protein-carbohydrate activity. Figure 3.59 shows a histogram of the PA-IIL activity.

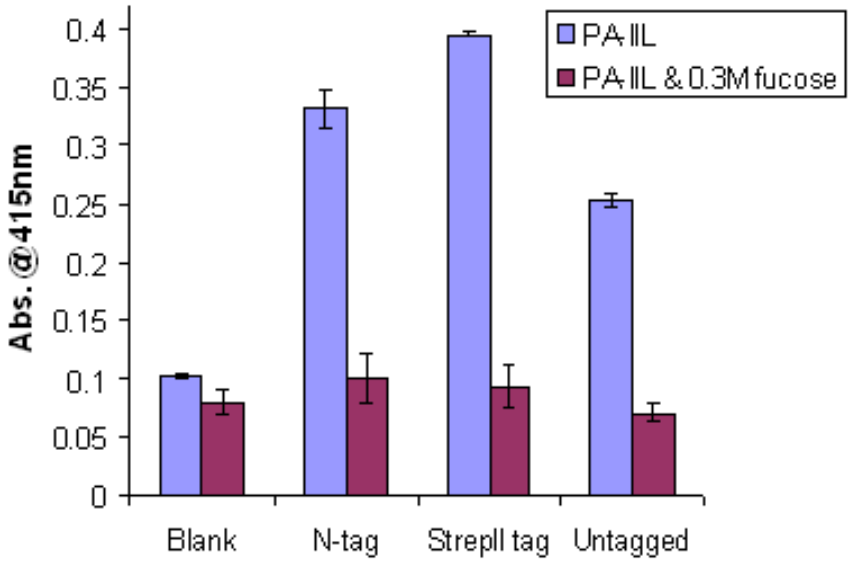


Figure 3.59: Activity of recombinantly tagged PA-IIL
HRP assay carried out as per Section 2.26. PBS used in place of PA-IIL as a negative control (blue). A concentration of 10 μg/ml of PA-IIL used per reaction. The activity was measured at OD_{415}. (Error bars represent mean ± SD, n = 3).

The variation of the activity between the alternatively tagged PA-IIL may be due to inaccuracies in calculating protein concentration due to the precipitation of the protein (as discussed in Section 3.9). The HRP assay indicates that all expressed recombinant PA-IIL was active and the activity was due to sugar specific binding.

3.9 Protein stability of recombinant PA-IIL

PA-IIL was found to precipitate in solution after a short period of time when highly purified and stored at both 4°C and -20°C. A number of additives were investigated to improve the stability of PA-IIL in solution.

3.9.1 Effect of the addition of a 6HIS tag to PA-IIL on protein stability

The addition of a 6HIS recombinant tag may be contributing to PA-IIL precipitation. The addition of a tag with a hydrophilic nature such as 6HIS may mediate protein-protein interaction leading to precipitation. To ascertain if this is what was occurring untagged PA-IIL was expressed and purified (Section 3.8.2, Figure 3.57). Untagged PA-IIL was found to precipitate when purified and stored at 4^{o}C and -20^{o}C. Precipitation of PA-IIL was therefore attributed to be the inherent nature of protein.

3.9.2 Use of stabilizing additives

Additives are commonly used to stabilize protein during storage, e.g. BSA, glycerol, Triton X etc. A preliminary stabilisation study was carried out to investigate the effect of a variety of additives on the stability of PA-IIL in solution. 6HIS N-tagged PA-IIL was expressed and purified (Figure 3.52). Protein was aliquoted 500 µl per epindorf at low concentration (50 – 100 µg/ml) and stored at 4^{o}C. Samples were spun down daily at 13,000 rpm for 10 min and the presence of visible precipitate pellet noted. The additives investigated included, Triton X, KCl, Polyethylene glycol (PEG), BSA, glycerol and ethanol at various concentrations. Two additives were found to stabilize PA-IIL for 14 days and prevent visible precipitation; 40 % glycerol and 1 % ethanol. After 14 days, protein storage at 4^{o}C is not recommended. The confirmation of proteins have been found to be stabilized and preferentially hydrated in a glycerol solvent system (Gekkot, Timasheff 1981).The addition of a relatively high concentration of glycerol to PA-IIL leads to a high viscosity solution. A highly viscose solution may be problematic in future lectin characterisation work e.g. use in ELLA would require increased development times etc. Ethanol like glycerol is used as a cryoprotectant for proteins. It is preferentially excluded from the surface of proteins while serving to prevent stress from denaturing the protein. However, addition of ethanol can create a harsh environment and lead to denaturation of protein. PA-IIL has been shown to be a robust protein, tolerating temperatures up to 80^{o}C (Gilboa-Garber 1982) as well as extreme pH values of 4.5 -11.5 (Tielker 2006). It is unknown wheather homologs of PA-IIL will retain its robust nature therefore 1% ethanol was deemed an inappropriate stability additive for future work.

3.10 Storage conditions

Both additives identified in Section 3.9, 40% glycerol and 1% ethanol, as the most effective stabilization agents for PA-IIL have inherent problems with use. Lyophilization, whereby a frozen liquid is vacuum dried to a powder, was identified as a possible method of long term storage of PA-IIL. Water is removed from protein by freeze drying, therefore removing the proteins ability to precipitate.

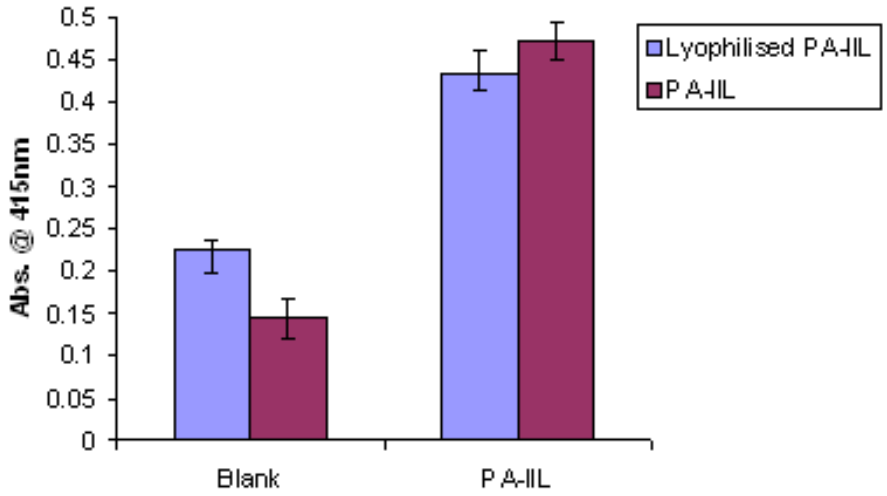


Figure 3.60: Histogram of the effect of lyophilization on activity of PA-IIL.
HRP assay carried out as per Section 2.26. PBS used in place of PA-IIL as a negative control (blue). 10 µg/ml of 6HIS N-tagged PA-IIL used per well. Lyophilised PA-IIL – blue, untreated PA-IIL – purple. (Error bars represent mean ± SD, n = 3).

PA-IIL was expressed and purified as outlined in Section 2.15. The purified protein was split and half of the sample was aliquoted in 250µL to sterile microcentrifuge tubes and lyophilized as outlined in Section 2.21. The effect of lyophilization on the activity of PA-IIL was assessed by HRP assay (Section 2.26). There was no significant variation in the activity of PA-IIL which was lyophilised and that which was not (Figure 3.60). Lyophilization was used subsequently as a storage method for purified PA-IIL.

3.11 Discussion

To ensure a high quantity of purified PA-IIL could be produced, as would be required for analytical methods such as mass spectrometry and chromatography, PA-IIL expression was extensively optimised. Despite some variation on the amino acid level (Figure 1.13) it was hoped that the expression conditions in the *E .coli* expression system for PA-IIL would be suitable for the expression of the lectin region of the

Photopexins of *P. luminescens* (see Chapter 4). PA-IIL has previously been expressed in *E. coli* (Loris *et al.* 2003). This was repeated in Section 3.4, with the inclusion of 6HIS to the N and C-terminus of the lectin by the cloning of the PA-IIL gene into the pPAIIL3 and pQE30.PAIIL expression plasmids. The affect of 6HIS fusion to both the N and C terminus of the *PA-IIL* gene on lectin activity and structure was investigated (Section 3.5). C-terminally tagged PA-IIL was found, by HRP assay, to have poor activity (Figure 3.41). The loss of activity correlates with data from PA-IIL crystallography studies (Mitchell *et al.* 2002). It was found that the C-terminal amino acid (Gly114) plays a role in the binding pocket of the neighbouring subunit in the tetrameric PA-IIL structure (Figure 1.2). Addition of histidine peptides to the C-terminal glycine disrupts the activity within the binding pocket, impairing the lectins ability to bind. The addition of 6HIS to the C-terminus was also found to disrupt the tetrameric structure of the lectin. The molecular weight of C-tagged PA-IIL, as determined by size exclusion chromatography (Section 3.5.3) was found to be 27kDa (see Table 3.8), i.e. a dimer. There are three possible dimer formations (see Figure 3.61).

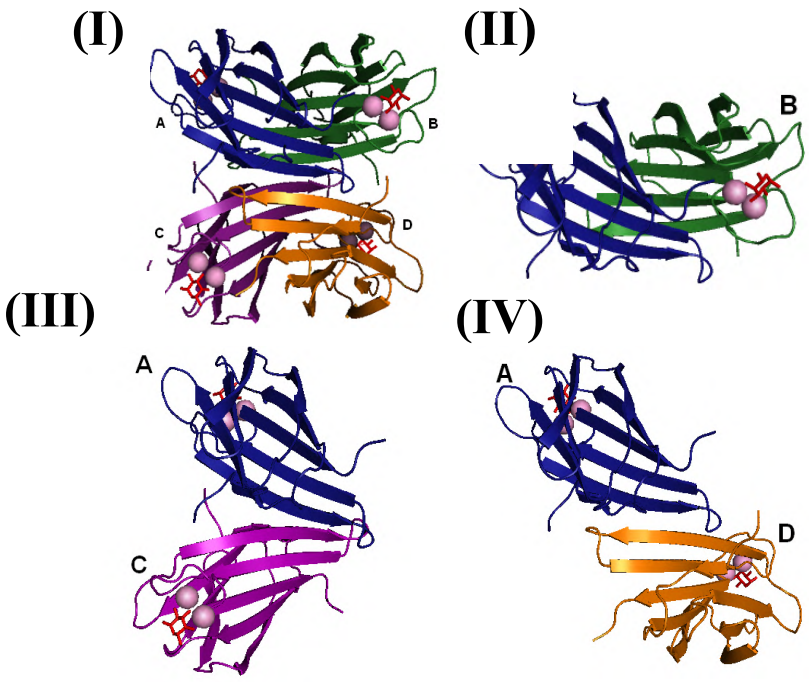


Figure 3.61: Three possible dimer formations of C-terminally 6HIS tagged PA-IIL.
(I) Tetrameric structure of the PA-IIL-fucose complex (Mitchell *et al.* 2002), each monomer shown in a different colour. Stick representation of fucose (red) and two calcium ions shown as pink spheres. Each monomer is labelled A, B, C and D according to the nomenclature used in the text. (II) Formation of A-B PA-IIL dimer; (II) Formation of A-C dimer; (IV) Formation of A-D dimer. Images generated using Pymol and PDB file 1UZV.

Loris, R. *et al.* (2003) discerned the interfaces within the PA-IIL tetramer. Interactions between A-B were found to be hydrophobic, A-D interaction hydrophilic in nature while the A-C interface has limited contact between hydrophilic side chains. This implies that dimer formation of A-B or A-D is more likely due to stronger bond formation then those found for A-C. Size exclusion chromatography cannot determine the exact dimer formation. Further work, such as mass spectrometry or density gradient ultracentrifugation (Cole, Hansen 1999) would be required to determine this (not done in this study). N-terminally 6HIS tagged PA-IIL was found to be active**,** i.e. addition of 6HIS to the N-terminus did not affect binding activity (Figure 3.41). The relative molecular mass (by SEC) of 6HIS PA-IIL was found to be below the predicted MW of a tetramer. However, it is not unusual for a small compact protein, such as PA-IIL, to diffuse at a slower rate through a size exclusion chromatography column then the globular standards used to establish the linear curve (Dubin, Principi 1989). The N-tagged PA-IIL was therefore assumed to be a tetramer, i.e. addition of 6HIS to the N-terminus of PA-IIL did not disrupt the quaternary structure of the protein.

Having successful active expression of PA-IIL in *E. coli* from pQE30.PAIIL (Section 3.5), the *PA-IIL* gene was cloned into a number of commercial pQE vectors (Qiagen). These are high yield expression plasmids with strong promoter systems. The optimal *E .coli* expression strain for the recombinant expression of PA-IIL was determined from four commonly used *E. coli* strains, BL21 (DE3), XL10-Gold, JM109 and XL1-Blue. Freshly transformed colonies can posses highly varying expression rates of soluble protein. The colony blot method (see Section 2.15.2) was utilised to distinguish between freshly transformed clones of pQE30.PAIIL in each of the *E. coli* strains that have high protein expression rates. Figure 3.44 shows the dot blot of the four *E. coli* strains and variation of expression between the clones. XL10-Gold was found to have consistent levels of expression, while expression varied greatly for the remaining strains. This experiment has its limitations as it does not distinguish between soluble and insoluble proteins.

Further optimisation was carried out on *E. coli* XL10-Gold and BL21 (DE3) as these strains are the more commonly used expression strains. The optimal expression temperature and harvest time for fermentation from both of these strains was investigated. Expression of PA-IIL overnight from BL21 (DE3) at 37^{o}C found a decrease in protein yield. Protein also failed to be expressed when cultures were grown

at 30°C, regardless of *E. coli* expression strain. It is possible that expression at high levels of PA-IIL leads to aggregation of the protein within the cytoplasm. PA-IIL was found to inherently precipitate when purified (see Section 3.9) and therefore it is not unlikely that protein may also aggregate during expression. The BL21 (DE3) strain is known to be a leaky expressor as it does not contain the *lac*I^q mutation, which results in the overproduction of the Lac reporter gene. Consequently expression from BL21 (DE3) can occur without the addition of the repressor binding molecule IPTG (see Figure 3.46 A and B Lane 2). The leaky expression from BL21 (DE3) may contribute to prolong expression of PA-IIL leading to high levels of protein and ultimately aggregation and reduced soluble protein yield, as found for overnight expression from BL21 (DE3) at 37°C. The optimal *E. coli* expression strain for PA-IIL was therefore determined to be XL10-Gold as XL10-Gold contains F factors carrying the *lac*I^q allele conferring greater control of protein expression. Strong control of protein expression by XL10-Gold led to consistent protein expression (Figure 3.44) at yields suitable for future work while preventing excess protein production leading to aggregation and loss of protein.

SDS-PAGE analysis of the high expression clones identified found the insoluble fractions of these cultures showed a significant quantity of PA-IIL present (Figure 3.45). This may be due to the degradation of the protein, improper folding, or protein aggregation. Attempts to recover PA-IIL from the insoluble fraction by L-arginine had no greater effect on the levels of protein recovered from the insoluble fraction than lysis buffer (Figure 3.49) therefore this is not a suitable method to recover PA-IIL from the insoluble fraction. A sufficient quantity of PA-IIL was expressed to the soluble fraction for protein characterisation; therefore no further attempts were made to solublise PA-IIL from the insoluble fraction.

Induction by IPTG was routinely done at 50μM concentration once the OD_{600nm} reached 0.5-0.6. The effect of induction with varying levels of IPTG was investigated by SDS-PAGE analysis. It was concluded that increasing the IPTG concentration did not result in increase yields of recombinant protein being expressed in the soluble fraction (Figure 3.50).

Once the optimal conditions of recombinant protein expression were evaluated, a purification strategy was investigated to maximise the recovery of recombinant protein from the soluble cell extract while maintaining a high level of purity. IMAC purification strategies were employed for the purification of 6HIS tagged lectins. PA-IIL remained bound to the column in washes up to 50mM imidazole and the final

protein was eluted with 350mM imidazole. The high level of purity obtained by using IMAC can be seen by the highly sensitive silver stain SDS-PAGE gel, where no impurities were detected (see Figure 3.52). The sugar binding ability of PA-IIL was exploited for the purification of untagged PA-IIL from *E. coli* cell extract by affinity chromatography. PA-IIL, as expected, bound to mannose agarose and PA-IIL of high purity was then eluted with mannose (see Figure 3.55). PA-IIL was also purified by the affinity tag StrepII. StrepII tagged PA-IIL was bound to Streptacin sepharose and was then eluted by desthiotiotin. The high level of purity obtained by StrepII tag purification can be seen in Figure 3.57. The semi-quantitative HRP assay found all purified PA-IIL to be active.

PA-IIL was found to aggregate in a purified state (also noted by Gilbo-Garver, N. (1982)). Aggregation was found for wild type untagged PA-IIL, indicating that the inclination of PA-IIL to aggregate is an inherent property of the protein and is not induced or promoted by the addition of a recombinant tag. A number of combined factors may lead to precipitation such as slight changes in temperature, pH, ionic strength, protein concentration and hydrophobic and hydrophilic residues on the protein surface (Wang 2005). It is generally recommended to work at low ionic strength in order to prevent 'salting' out of protein and at low concentration of protein as the higher the concentration the faster the precipitation. There are a number of known additives and methods for improving protein stability and storage although there are tradeoffs associated with each method. Examples of these would include; Cryoprotectants such as glycerol used to a final concentration of 25-50% are often found to aid stabilization of proteins (Gekkot, Timasheff 1981) but may lead to highly viscose solutions. Dilute protein solutions (< 1 mg/ml) are more prone to inactivation and loss of acitivity as a result of low-level binding to the storage vessel. Therefore, it is common practice to add 'carrier' or 'filler' protein, such as purified bovine serum albumin (BSA), to supplement protein solutions to protect against such loss. Polyethylene glycol (PEG) is a chemical compound composed of repeating ethylene glycol units. It is commonly available commercially as mixtures of different oligomere sizes in broadly or narrowly defined molecular weight ranges. PEG is hydrophilic in nature and attachment of PEG to proteins decreases aggregation and increases solubility. Detergents (such as Triton X-100 or Tween20) may also prevent aggregation but often bind strongly to hydrophobic proteins. Suitable storage conditions to enhance the stability of PA-IIL in solution were investigated. The pH of each solution was kept at pH 7.8 as it is the required pH for

future steps with the protein. The detergents Triton X and Tween 20 were also used as they too are compatible with future characterization experiments, e.g. ELLAs. A crude precipitation study carried out at room temperature with a wide variety of additives commonly used for protein storage found that 40% glycerol and 1% ethanol prevented the visible appearance of precipitation for the longest period of time. Although both glycerol and ethanol stabilize the protein neither are compatible for further characterization steps. It is important to note that the addition of a relatively high concentration of glycerol will lead to a high viscosity solution. High viscosity may lead to problems in use for ELLA assay, e.g. require extra mixing or extended reaction times etc., and may be problematic in future surface immobilization strategies. The addition of ethanol is considered unsuitable as a generic additive as although the stability of PA-IIL is enhanced by the addition of ethanol, PA-IIL has been shown to be a robust protein, tolerating temperatures up to 80^{o}C (Gilboa-Garber 1982) as well as extreme pH values of 4.5 -11.5 (Tielker 2006). It is unknown wheather homologs of PA-IIL can tolerate the presence of ethanol, therefore it was considered an unsuitable additive.

Lyophilization allows for long-term storage of protein with little threat of degradation but the protein must be reconstituted before use and may be damaged by the lyophilization process. Activity of lyophilized PA-IIL was not greatly affected by lyophilization and reconstituted as determined by the HRP assay (Figure 3.60). This method was subsequently used for all protein storage.

4.0 Cloning, expression, purification of recombinant Photopexins A and B

4.1 Overview

Two homologs to PA-IIL have been identified in the insect pathogenic bacterium *P. luminescens* (Crennell 2000) as discussed in Section 1.6. The first homolog identified, Photopexin A (ppxA), is composed of two 220 amino acid homologous domains and a 100 amino acid C-terminal domain that has a 58% similarity to PA-IIL (Table 1.3). The second homolog identified, Photopexin B (ppxB), is homologous to two domains of ppxA, i.e consists of one repeating domain and a 100 amino acid C-terminal domain that has a 58% similarity to PA-IIL (Table 1.3). The calcium binding loop is highly conserved between Photopexin A and B and PA-IIL. The amino acids involved in monosaccharide binding differ across the three proteins (Figure 4.1 and 4.9). From the protein alignments it was hypothesised that the homologs found within *P. luminescens* may have differing sugar specificities to both PA-IIL and each other. This hypothesis was investigated by the cloning and expression of the C-terminal lectin-like domain of Photopexin A (ppxA/L) and the C-terminal lectin-like domain of Photopexin B (ppxB/L).

4.2 Cloning of the *P. luminescens* lectin region of Photopexin A

The optimal expression and purification protcols of PA-IIL in *E. coli* were established as discussed in Chapter 3. Protein sequence alignment identified the C-terminal domain of ppxA as having significant similarity to PA-IIL (see Section 1.6, Figure 4.1). While the amino acids involved in calcium binding are conserved across the proteins, those involved in sugar binding vary (Figure 4.1). The impact of this variation in the sugar binding loop was investigated by cloning of the lectin region of Photopexin A (ppxA/L). The significant sequence similarity between PA-IIL and ppxA/L suggested that ppxA/L could be cloned, expressed and purified under similar conditions as PA-IIL. The cloning and expression of ppxA/L under the optimised conditions for PA-IIL is discussed in the following section.

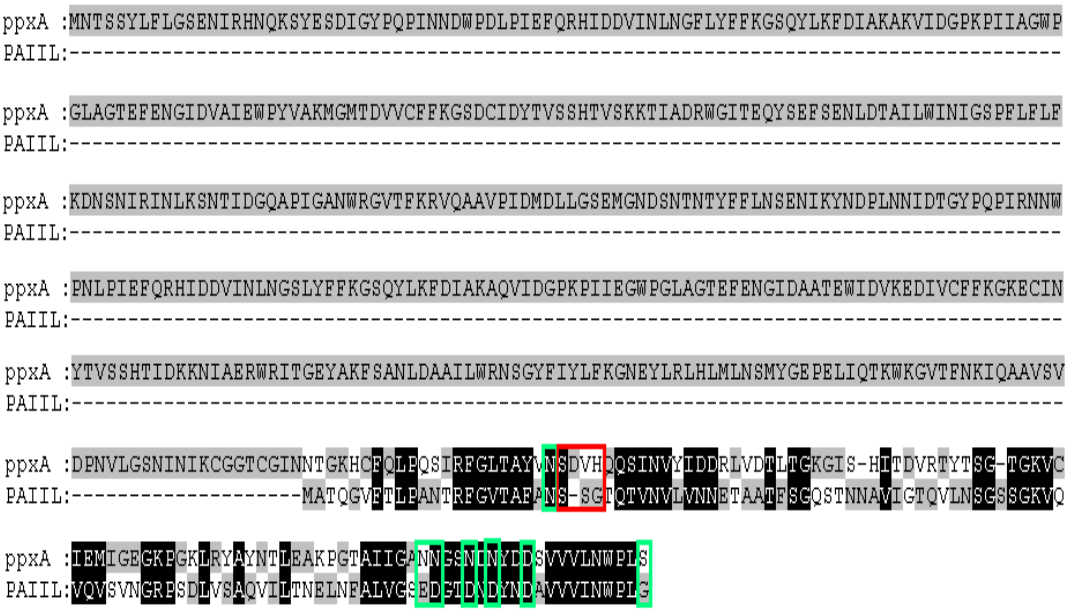


Figure 4.1: Amino acid alignment of PA-IIL and Photopexin A (ppxA) from *P. luminescens*.
Amino acids involved in monosaccharide binding highlighted in red and amino acids involved in calcium binding highlighted in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

Optimal cloning and expression of PA-IIL was carried out using the Qiagen pQE expression vector pQE30. The pQE30 vector was therefore utilised for the cloning of ppxA/L. The pQE-30 vector (Table 2.3, Figure 2.4) features the T5 promoter/*lac* operator element and the β-lactamases gene conferring resistance to ampicillin (*amp*^{R}). The pQE30 vector facilitates the expression of a fusion protein having an N-terminal 6HIS tag. The Photopexin A protein can be viewed as a 450 amino acid N-terminal protein fused to a smaller 114 amino acid C-terminal lectin. Since the C-terminal lectin can tolerate this large N-terminal fusion the addition of a 6HIS tag to the N-terminus of the lectin domain should not greatly affect the lectin activity.

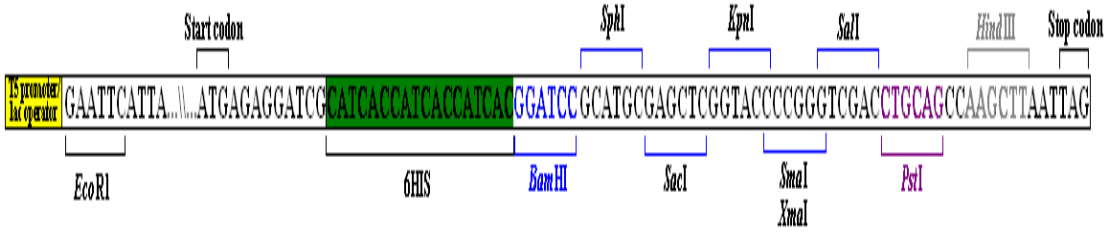


Figure 4.2: The multiple cloning site (MCS) of the pQE30 vector.
Schematic outlining the positions of the relevant restriction sites, 6HIS tag and start and stop codons found within the MCS of the expression vector pQE30.

The *P. luminescens* TTO1 strain was provided by Dr. David Clarke, UCC, and the sequence for the organism was obtained from the NCBI data bank, data entry no. NC_005126, (Duchaud *et al.* 2003). The primers ppxA/L-f and ppxA-r were designed for the amplification of the lectin region of the *ppxA* gene as a *BamH*I-*Hind*III 348 b.p fragment by PCR. The primed region of the *ppxA* gene for which the primers were designed is shown in Figures 4.3 and 4.4.

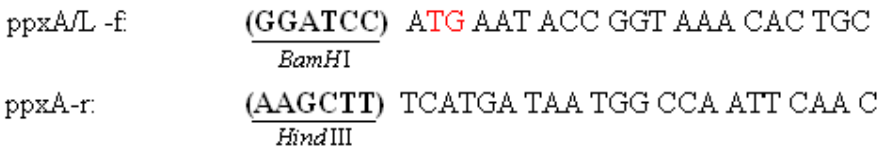


Figure 4.3 Primers used for the amplification of the lectin region of the *Photopexin A* (*ppxA/L*) from *P. luminescens* TTO1 for cloning into the pQE30 vector.
The enzymatic restriction sites added to the *ppxA/L* gene are shown underlined and labelled. Mutated nucleotides for the addition of a start codon are highlighted in red.

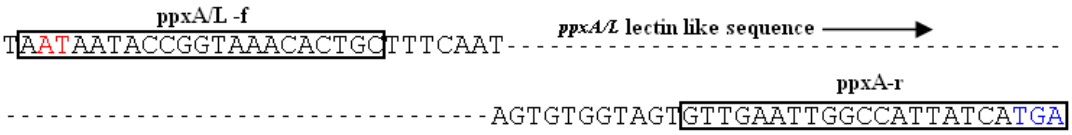


Figure 4.4: Primed region of *ppxA/L* sequence from *P. luminescens* TTO1 for suitable amplification for cloning into the pQE30 vector. The stop codon of *ppxA/L* is highlighted in blue. The bases mutated by the primer for the introduction of a start codon is highlighted in red.

Genomic DNA was prepared from *P. luminescens* and used as template DNA for the amplification of the *ppxA/L* gene. The conditions used for the PCR are shown in Table 4.1 below. The PCR was carried out with the high fidelity *Taq* polymerase, Phusion*Taq*.

Table 4.1: PCR conditions for the amplification of the *ppxA/L* gene for cloning to the pQE30 vector.
Primers ppxA/L-f and ppxA-r used for amplification as described above.

PCR Condition	
Annealing Temp.	50°C
Annealing Time	20 sec
Extension Time	30 sec
No. of Cycles	50

The PCR product was cleaned using Gel/PCR DNA fragment extraction kit. Both the cleaned PCR product and the pQE30 vector were digested with *BamH*I/*Hind*III and the restricted products cleaned (Figure 4.5).

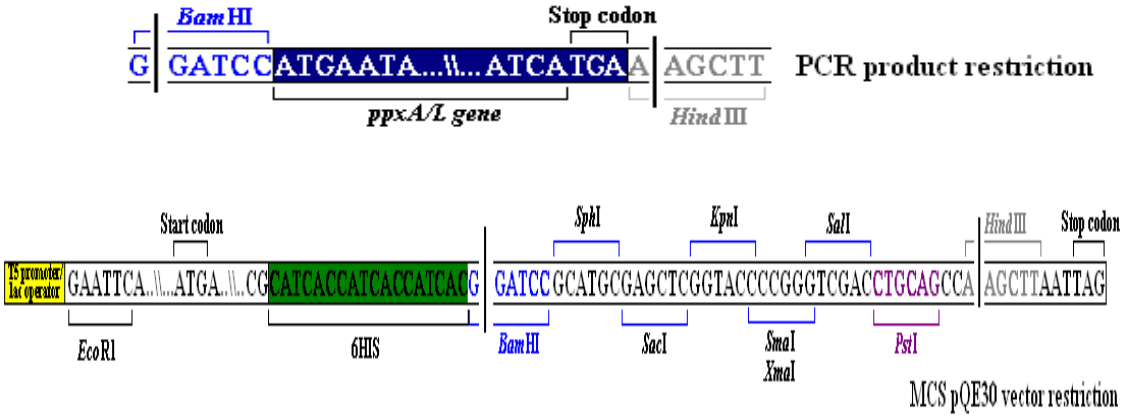


Figure 4.5: Schematic of *Bam*HI/*Hind*III restricted PCR product amplified by ppxA/L-f and ppxA-r and schematic of the MCS of the pQE30 vector restricted by *Bam*HI/*Hind*III.
The point of restriction is indicated by a black line and break in the sequence.

The linearised pQE30 plasmid DNA was ligated with the purified ppxA/L *BamH*I-*Hind*III fragment. The ligated DNA was transformed to *E. coli* JM109 and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were screened by performing *BamH*I, *Hind*III and *BamH*I/*Hind*III double digests on plasmid DNA prepared from each clone. The restriction digests were visualised on 0.7% agarose gels. Screening in this manner resulted in the identification of the desired plasmid named pQE30.ppxA/L. Figure 4.6 below shows the results of the verifying restriction analysis on the plasmid.

Lane	Sample
1	1Kb Ladder
2	pQE30.ppxA/L uncut
3	pQE30. ppxA/L cut *Bam*HI
4	pQE30. ppxA/L cut *Hind*III
5	pQE30.ppxA/L cut *Bam*HI/*Hind*III

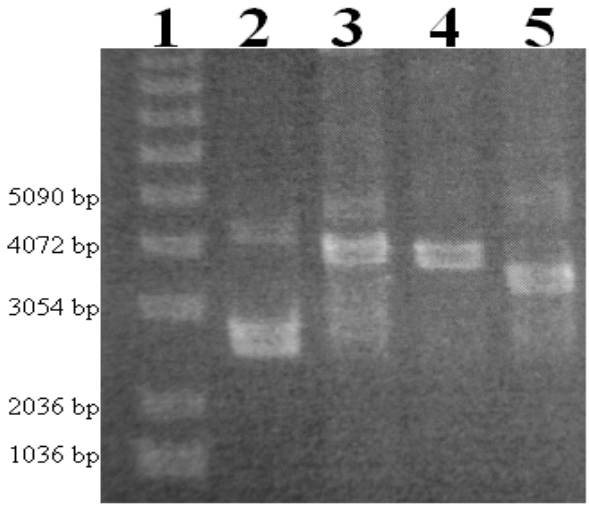


Figure 4.6: Analysis and restriction of pQE30.ppxA/L.
The restriction digest of the pQE30.ppxA/L clone analysed on a 0.7% agarose gel (Section 2.7) showed that single restriction with either *BamH*I and *Hind*III produced a band corresponding to the expected (~3800 b.p) linearised plasmid. Double restriction with *BamH*I/*Hind*III excised a band of ~350 b.p resulting in a visible band shift lower of the plasmid which is indicative of the presence of the *ppxA/L* fragment.

The presence of the correct insert within the pQE30.ppxA/L vector (Figure 4.7) was further confirmed by DNA sequencing. Translation of protein from this expression vector results in N-terminal 6HIS tagged ppxA/L.

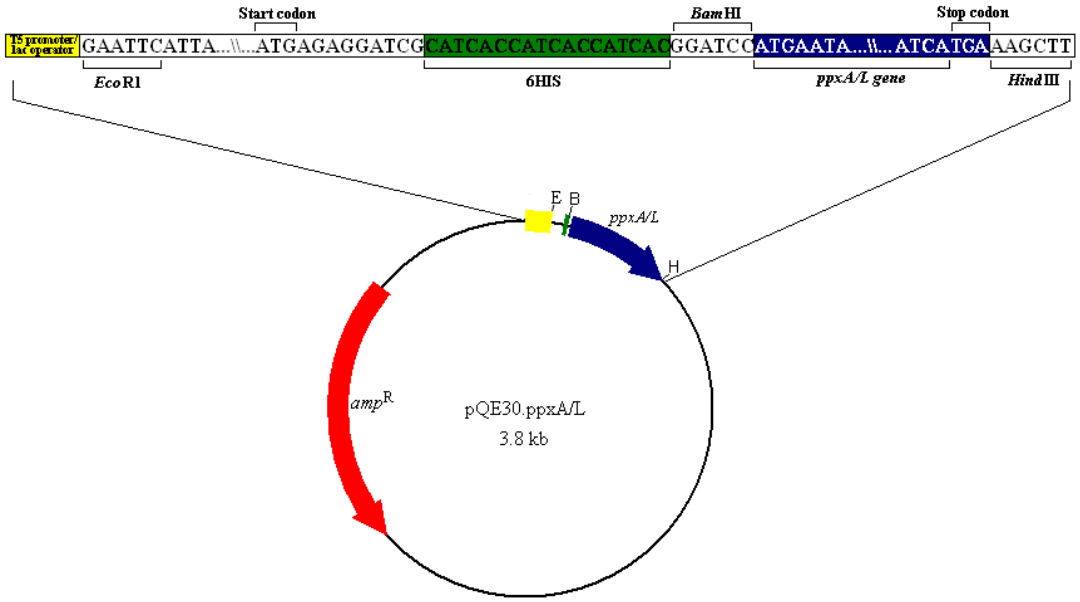


Figure 4.7: Schematic of pQE30.ppxA/L plasmid.
The emphasised region of the pQE30.ppxA/L vector outlines the relative positions of the *ppxA/L* gene, the relevant restriction sites, start and stop codon and 6HIS. Image generated using BVtech.

Three nucleotides were found to differ between the sequenced pQE30.ppxA/L clone, (confirmed in both forward and reverse reads) and the sequence obtained for ppxA from Genebank (Figure 4.8).

```
Blast  :                                                AATACCGGTAAACACTGCTTTCAATTACCTCAAAGTATTCGATTTG
ppxA/L : ATGACAGCATCGCATCACCATCACCATCACGGATCCATGAATACCGGTAAACACTGCTTTCAATTACCTCAAAGTATTCGATTTG

Blast  : GCCTGACTGCCTACGTAAATAGTGATGTTCATCAACAATCAATCAATGTCTATATCGATGACCGATTGGTCGATACCTTAACCGG
ppxA/L : GCCTGACTGCCTACGTAAATAGTGATGTTCATCAACAACAATCAATGTCTATATCGATGACCGATTGGTCGATACCTTAACCGG

Blast  : TAAAGGAATAAGTCACATAACGATGTTAGGACCTACACATCAGGTACCGGCAAAGTCTGCATCGAGATGATAGGAGAGGGTAAA
ppxA/L : TAAAGGAATAAGTCACATAACGGATGTTAGGACCTACACATCAGGTACCGGCAAAGTCTGCATCGAGATGATAGGAGATGGTAAA

Blast  : CCCGGCAAACTCCGTTATGCCTATAACACTCTCGAAGCAAAACCCGGCACAGCCATTATCGGCGCTAACAATGGTTCCAATGACA
ppxA/L : CCCGGCAAACTCCGTTATGCCTATAACACTCTCGAAGCAAAACCCGGCACAGCCATTATCGGCGCTAACAATGGTTCCAATGACA

Blast  : ATTATGATGATAGTGTGGTAGTGTTGAATTGGCCATTATCATGA
ppxA/L : ATTATGATGATAGTGTGGTAGTGTTGAATTGGCCATTATCATGA
```

Figure 4.8: Alignment of nucleotide sequence of pQe30.ppxA/L and ppxA sequence obtained from GeneBank.
Start and stop codons highlighted in orange, 6HIS codon highlighted in green. Variation in nucleotide bases highlighted in red.

The variation of nucleotides resulted in a silent mutation of a serine (S) to threonine (T), a threonine (T) codon for an alternative threonine (T) codon and a glutamic acid (E) to an aspartic acid (D) (Figure 4.9). The variation between the sequence obtained from Genebank and the sequence from pQE30.ppxA/L may be due to single-nucleotide polymorphism (SNP). A SNP denotes a single nucleotide change in a genome that differs between members of a species. SNPs have been more widely associated with human disorders (Hirschhorn *et al.* 2002). SNPs do not always impact the activity of a protein although a variety of SNPS have been discovered that enhance the pathogenicity of bacteria highlighting the need to study SNP of bacteria (Weissman *et al.* 2003). The SNPs of ppxA/L are located outside the key locations, sugar binding loop and calcium binding loop, and due to the similar nature of the nucleotides involved it was theorized that the variation of the nucleotides would not affect protein activity.

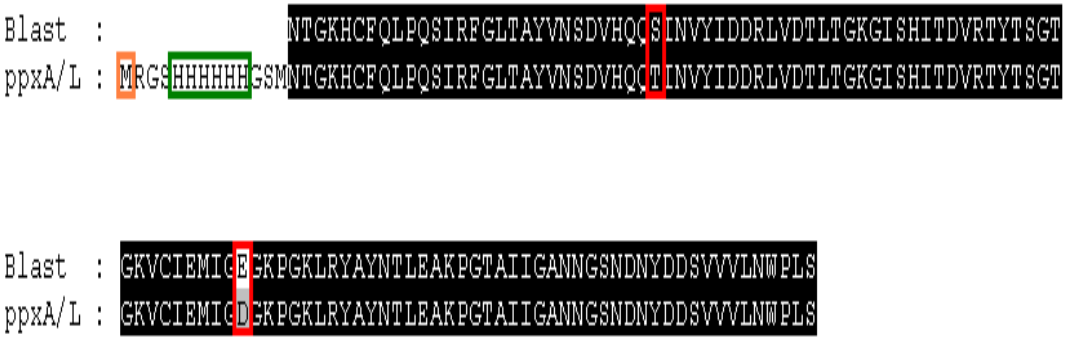


Figure 4.9: Amino acid alignment of protein sequences of the expected *ppxA* sequence obtained from Genebank and the DNA sequencing reads obtained fro pQE30.ppxA/L.
Differing amino acid highlighted in red, start codon highlighted in orange, 6HIS tag highlighted in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

4.3 Cloning of the *P. luminescens* lectin region of Photopexin B

Protein sequence alignment identified the C-terminal domain of Photopexin B (ppxB) as having significant similarity to PA-IIL (Section 1.6). The amino acids involved in calcium binding are conserved across these proteins while those involved in sugar binding vary (Figure 4.10). The impact of this variation in the sugar binding loop was investigated by cloning of the lectin region of *ppxB*. The significant sequence similarity between PA-IIL and ppxB/L suggest that ppxB/L could be cloned, expressed and

purified under the same conditions as PA-IIL. The cloning and expression of ppxB/L under the optimised conditions for PA-IIL is discussed in the following sections.

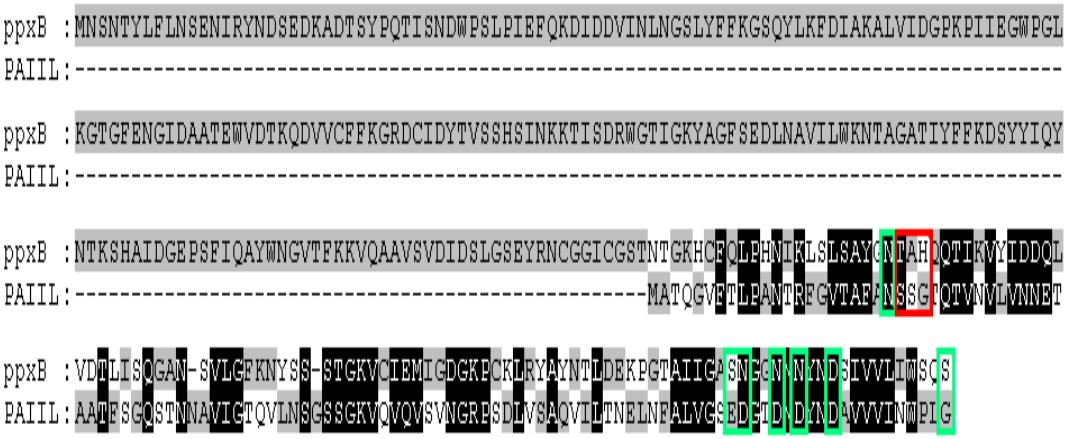


Figure 4.10: Amino acid alignment of PA-IIL and Photopexin B (ppxB) from *P. luminescens*
Amino acids involved in monosaccharide binding highlighted in red and amino acids involved in calcium binding highlighted in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

4.3.1 Cloning of the lectin region of Photopexin B

The expression vector pQE-30 was shown to be a suitable vector for cloning of the *PA-IIL* gene and the lectin region of the *ppxA* gene as outlined in Sections 3.2 and 4.2. It was therefore chosen for the cloning of the lectin region of the Photopexin *B (ppxB/L)*. The pQE30 vector facilitates the translation of a fusion protein having an N-terminal 6HIS tag. The Photopexin B protein can be viewed as a 227 amino acid N-terminal protein fused to a smaller 113 amino acid C-terminal lectin. Since the C-terminal lectin can tolerate this large N-terminal fusion the addition of a 6HIS tag to the N-terminus of the lectin domain should not greatly affect the lectin activity.

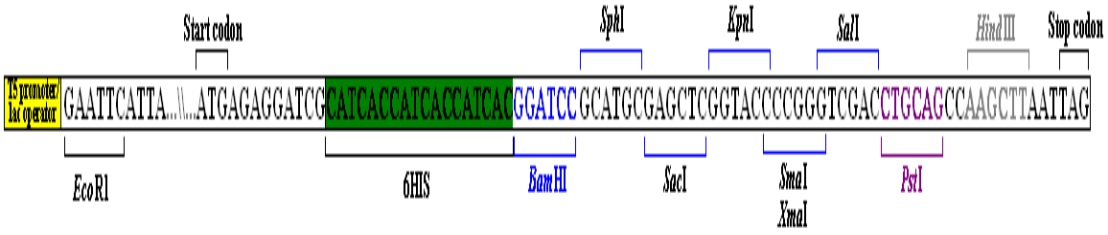


Figure 4.11: The MCS of the pQE30 vector.
Schematic outlining the positions of the relevant restriction sites, 6HIS tag and start and stop codons found within the MCS of the expression vector pQE30.

The primers ppxB/L-f and ppxB-r were designed for the amplification of the lectin region of the *ppxB* gene as a *BamH*I- *Pst*I 348 b.p fragment by PCR (as outlined in section 2.11.1). The primed region of the *ppxB* gene for which the primers were designed is shown in Figures 4.12 and 4.13.

ppxB/L-f: (GGATCC) ATG GCA GCA CAA ATA CCG G
*BamH*I

ppxB-r: (CTGCAG) TCA GGA CTG TGA CCA AAT TAA CAC
*Pst*I

Figure 4.12: Primers used for the amplification of the *ppxB/L* gene from *P. luminescens* TTO1 for cloning into the pQE30 vector.
The enzymatic restriction sites added to the *ppxB/L* gene are shown underlined and labelled. Mutated nucleotide for the addition of a start codon is highlighted in red.

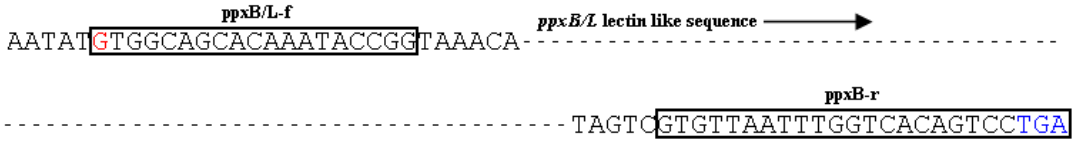


Figure 4.13: Primed region of the *ppxB/L* sequence from *P. luminescens* TTO1 for suitable amplification for cloning into the pQE30 vector.
The stop codon of *ppxB/L* is highlighted in blue. The base mutated by the primer for the introduction of a start codon is highlighted in red.

Genomic DNA was prepared from *P. luminescens* and used as template DNA for the amplification of the *ppxB/L* gene. The conditions used for the PCR are shown in Table 4.2 below. The PCR was carried out with the high fidelity *Taq* polymerase, Phusion*Taq*.

Table 4.2: PCR condition for the amplification of the *ppxB/L* gene for cloning to the pQE30 vector.
Primers ppxAB/L-f and ppxB-r used for amplification as described above

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	30 sec
No. of Cycles	50

The PCR product was cleaned using Gel/PCR DNA fragment extraction kit. Both the cleaned PCR product and the pQE30 vector were digested with *BamH*I/*Pst*I and the restricted products cleaned (as outlined in Section 2.8, Figure 4.14).

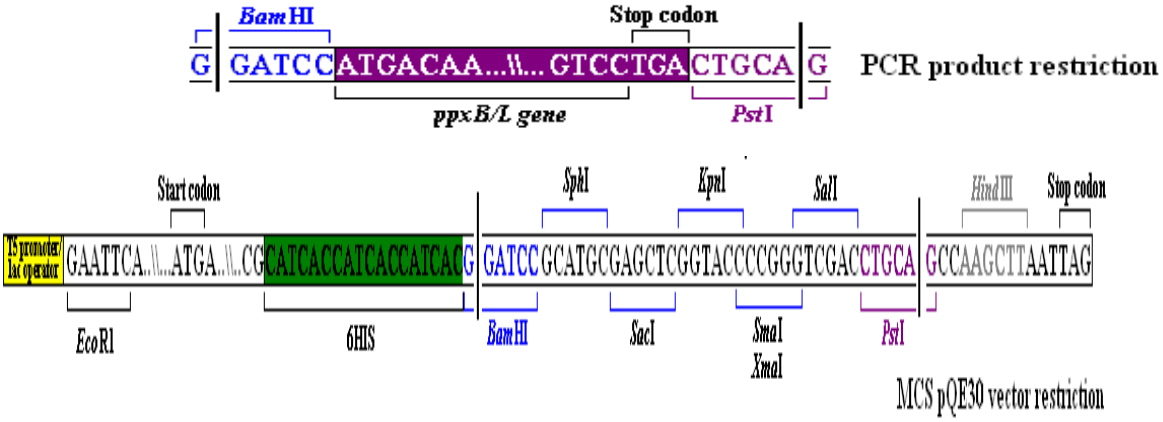


Figure 4.14: Schematic of *Bam*HI/*Pst*I restricted PCR product amplified by ppxB/L-f and ppxB-r and schematic of the MCS of the pQE30 vector restricted by *Bam*HI/*Pst*I
The point of restriction is indicated by a black line and break in the sequence.

The linearised pQE30 plasmid DNA was ligated with the purified ppxB/L *BamH*I-*Pst*I fragment. The ligated DNA was transformed to *E. coli* JM109 and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Candidate clones were selected, the plasmid DNA was isolated and then analysed by restriction analysis. Clones were screened by performing *BamH*I, *Pst*I and *BamH*I/*Pst*I double digests on plasmid DNA prepared from each clone. The restriction digests were visualised on 0.7% agarose gels. Screening in this manner resulted in the identification of the desired plasmid named pQE30.ppxB/L. Figure 4.15 below shows the results of the verifying restriction analysis on the plasmid.

Lane No.	Sample
1	1Kb Ladder
2	pQE30.ppxB/L uncut
3	pQE30.ppxB/L cut *Bam*HI
4	pQE30.ppxB/L cut *Pst*I
5	pQE30.ppxB/L cut *Bam*HI/*Pst*I

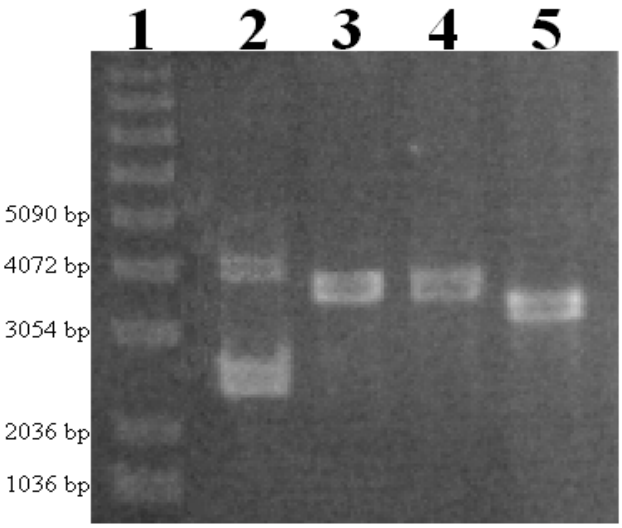


Figure 4.15: Analysis and restriction of the pQE30.ppxB/L vector.
The restriction digest of the pQE30.ppxB/L clone analysed on a 0.7% agarose gel (Section 2.7) showed that single restriction with either *BamH*I and *Pst*I produced a band corresponding to the expected (~3800 b.p) linearised plasmid. Double restriction with *BamHI/Pst*I excised a band of ~350 b.p resulting in a visible band shift lower of the plasmid which is indicative of the presence of the *ppxB/L* fragment within the clone indicating that the cloning of *ppxB/L* to the pQE30 vector was successful.

The presence of the *ppxB/L* gene within the pQE30 vector resulting in the pQE30.ppxB/L plasmid (Figure 4.16) was further confirmed by DNA sequencing. Translation of protein from this expression vector results in N-terminal 6HIS tagged ppxB/L.

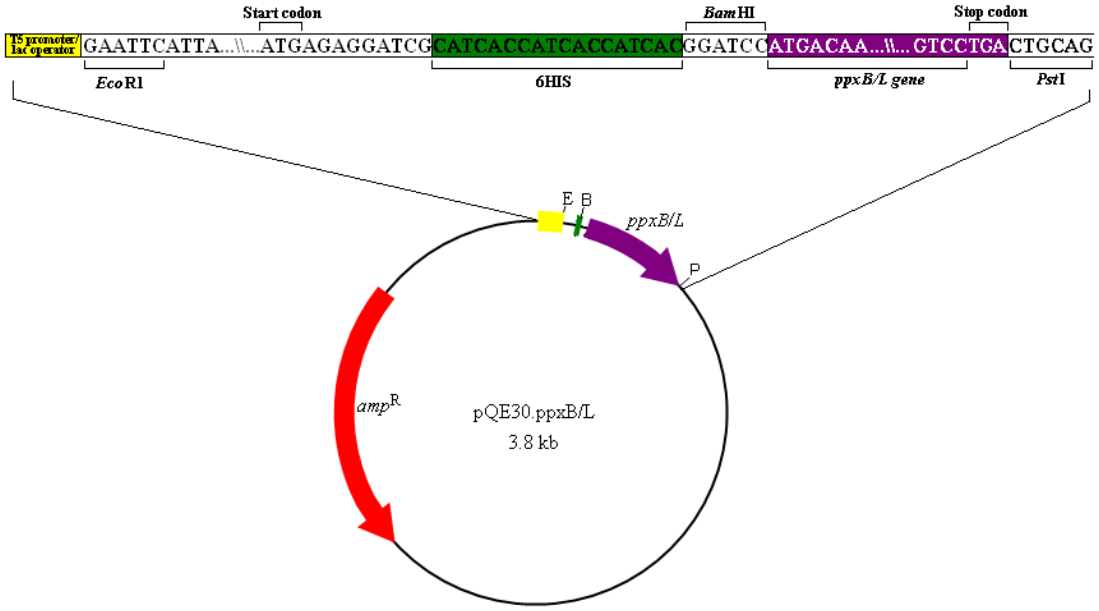


Figure 4.16: Schematic of pQE30.ppxB/L plasmid.
The emphasised region of the pQE30.ppxB/L vector outlines the relative positions of the *ppxB/L* gene, the relevant restriction sites, start and stop codon and 6HIS. Image generated using BVtech.

4.3.2 Addition of linker region

The failure to purify 6HIS tagged proteins by IMAC is commonly attributed to the inaccessibility of the tag due to it being buried as a result of protein fold or tightly bound via interaction with adjacent residues of the protein surface. The insertion of a linker peptide to distance the 6HIS tag from the protein has been shown to be a potential solution (Loughran *et al.* 2006). A linker was therefore introduced into the pQE30.ppxB/L vector to aid in protein purification via the 6HIS tag (as discussed in Section 4.5). The primers ppxB/L-link.f and ppxB/L-link.r were designed to insert a linker into the pQE30.ppxB/L.The primed region of pQE30.ppxB/L vector for which the primers were designed is shown in Figures 4.17 and 4.18.

ppxB/L-link.f: GGT GGT GGT GGA TCC ATG ACA AAT ACC GGT A
(5' Phosphorylation)

ppxB/L-link.r: ACC AGA GTG ATG GTG ATG GTG ATG GTG ATG CGA TCC
(5' Phosphorylation)

Figure 4.17: Primers used for the insertion of a linker sequence to the pQE30.ppxB/L vector.
The linker sequence, highlighted in orange, was dived between the forward primer and reverse primer. Phosphorylation of the 5' terminus of the primers facilitates ligation of the PCR product.

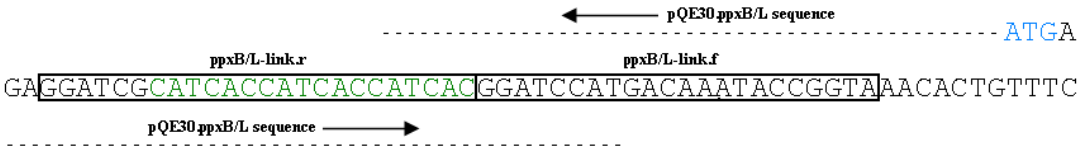


Figure 4.18: Primed region of the pQE30.ppxB/L plasmid for the addition of a linker region to the *ppxB/L* fragment.
The start codon of the *PA-IIL* gene is highlighted in blue, the 6HIS is highlighted in green.

Plasmid DNA was prepared from the pQE30.ppxB/L vector and used as template DNA for the addition of a linker to 6HIS ppxB/L. The conditions used for the PCR are shown in Table 4.3 below. The use of the high fidelity *Taq* polymerase, Phusion*Taq*, and elongated extension time, enabled full vector amplification.

Table 4.3: PCR conditions for the incorporation of a linker to the pQE30.ppxB/L vector.
The primers ppxB/L-link.f and ppxB/L-linkr.r were used for amplification as described above. Elongated extension time enables full vector amplification.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	2 min 30 sec
No. of Cycles	50

The PCR product (Figure 4.19) was cleaned using Gel/PCR DNA fragment extraction kit. The restriction enzyme *Dpn*I is effective against methylated DNA. Treatment of the PCR product with *Dpn*I leads to the cutting of the template pQE30.ppxB/L while the un-methylated PCR product remains intact (Section 2.12.1).

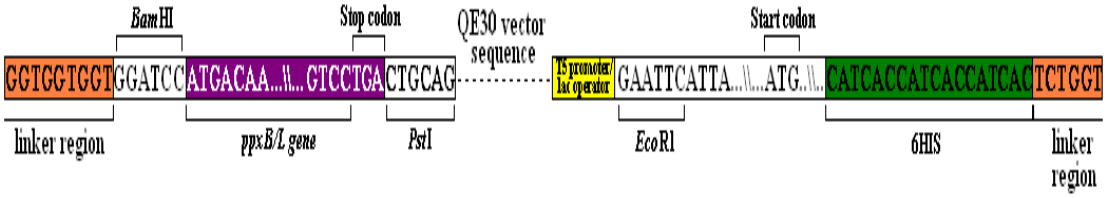


Figure 4.19: PCR product amplified by primers ppxB/L-link.f and ppxB/L-link.r, using the pQE30.ppxB/L vector DNA as a template.
Schematic outlines the positions of the *ppxB/L* gene, the relevant restriction sites, linker region, 6HIS tag and start and stop codons found within the MCS of the expression vector pQE30.ppxB/L.link.

Phosphorylation of the primers facilitated ligation of the PCR product to form a plasmid which was then called pQE30.ppxB/L.link (Figure 4.20). The ligated DNA was transformed to *E. coli* XL10-Gold and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Restriction analysis cannot differentiate between re-circularised template DNA and pQE30.ppxB/L.link as all restriction sites were identical.

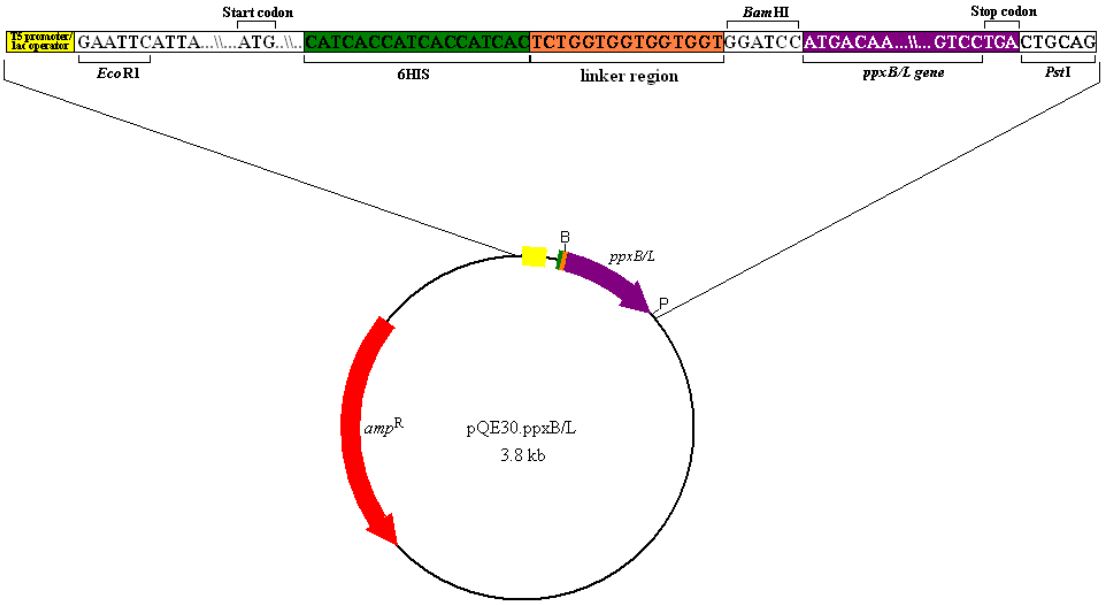


Figure 4.20: Schematic of the pQE30.ppxB/L.link plasmid
Emphasised region of the pQE30.ppxB/L.link vector outlines the relative positions of the *ppxB/L* gene, linker region, the relevant restriction sites, StrepII tag, start and stop codon. Image generated using BVtech.

DNA sequencing of the pQE30.ppxB/L-link plasmid revealed the linker was successfully incorporated into the plasmid. The sequencing also revealed the loss of a single nucleotide base (Figure 4.21). The loss of this nucleotide resulted in a frame shift and errors in protein translation.

```
Forward : ATGAGAGGATCGCATCACCATCACCATCACTCTGGTGGTGGTGGTGGATCCATGACAAATACCGGTAAACACTGTTTCCAGTT
Expected: ATGAGAGGATCGCATCACCATCACCATCACTCTGGTGGTGGTGGTGGATCCATGACAAATACCGGTAAACACTGTTTCCAGTT
Reverse : ATGAGAGGATCGCATCACCATCACCATCACTCTGGTGGTGGTGGTGGATCCATGACAAATACCGGTAAACACTGTTTCCAGTT

Forward : ACCTCACAATATAAAGCTTAGCCTGTCAGCCTATGGTAATACTGCCCACCAGCAAACGATAAAAGTGTATATTGACGATCAGT
Expected: ACCTCACAATATAAAGCTTAGCCTGTCAGCCTATGGTAATACTGCCCACCAGCAAACGATAAAAGTGTATATTGACGATCAGT
Reverse : ACCTCACAATATAAAGCTTAGCCTGTCAGCCTATGGTAATACTGCCCACCAGCAAACGATAAAAGTGTATATTGACGATCAGT

Forward : TAGTCGATACATTAATCAGCCAAGGTGCCAATAGTGTGTTGGGCTTTAAAAATTACTCATCGAGTACCGGCAAAGTCTGCATT
Expected: TAGTCGATACATTAATCAGCCAAGGTGCCAATAGTGTGTTGGGCTTTAAAAATTACTCATCGAGTACCGGCAAAGTCTGCATT
Reverse : TAGTCGATACATTAATCAGCCAAGGTGCCAATAGTGTGTTGGGCTTTAAAAATTACTCATCGAGTACCGGCAAAGTCTGCATT

Forward : GAAATGATAGGCGACGGCAAGCCCTGCAAACTCCGTTACGCCTATAACACCCTTGACGAAAA-CCGGGAACAGCCATTATCGG
Expected: GAAATGATAGGCGACGGCAAGCCCTGCAAACTCCGTTACGCCTATAACACCCTTGACGAAAAACCGGGAACAGCCATTATCGG
Reverse : GAAATGATAGGCGACGGCAAGCCCTGCAAACTCCGTTACGCCTATAACACCCTTGACGAAAA-CCGGGAACAGCCATTATCGG

Forward : TGCCAGCAATGGAGGCAATAATAATTACAATGACAGCATAGTCGTGTTAATTTGGTCACAGTCCTGA
Expected: TGCCAGCAATGGAGGCAATAATAATTACAATGACAGCATAGTCGTGTTAATTTGGTCACAGTCCTGA
Reverse : TGCCAGCAATGGAGGCAATAATAATTACAATGACAGCATAGTCGTGTTAATTTGGTCACAGTCCTGA
```

Figure 4.21: Sequence alignment of predicted ppxB/L linker sequence from GenBank with sequence from pQE30.ppxB/L-linker vector.
Stop and start codons highlighted in red, 6HIS tag highlighted in yellow and linker sequence highlighted in green. Loss of the single nucleotide highlighted in blue.

4.3.3 Addition of a single nucleotide to pQE30.ppxB/L.link

Primers, ppxB/L-linkfix-f and ppxB/L-linkfix-r, were designed to correct the loss of the nucleotide. The primed region of pQE30.ppxB/L.link vector for which the primers were designed is shown in Figures 4.22 and 4.23.

ppxB/L-linkfix.f: GCC ATT ATC GGT GCC AGC
(5' Phosphorylation)

ppxB/L-linkfix.r: TGT TCC CGG TTT TTC GTC AAG GG
(5' Phosphorylation)

Figure 4.22: Primers used for the insertion of a nucleotide base to the pQE30.ppxB/L.link plasmid.
The additional nucleotide to be inserted is highlighted in red. Phosphorylation of the 5' terminus of the primers facilitates ligation of the PCR product.

Figure 4.23: Primed region of the pQE30.ppxB/L.link construct for the addition of a nucleotide base to the *ppxB/L.link* fragment.
The region into which the nucleotide was inserted is underlined in red.

Plasmid DNA was prepared from the pQE30.ppxB/L.link vector and used as template DNA for the insertion of a nucleotide base to the *ppxB/L*.link fragment. The conditions used for the PCR are shown in Table 4.4 below. The use of the high fidelity *Taq* polymerase, Phusion*Taq*, and optimised elongated extension time, enabled full vector amplification.

Table 4.4: PCR conditions for the incorporation of a nucleotide to the pQE30.ppxB/L.link vector.
The primers ppxB/L-linkfix.f and ppxB/L-linkfix.r were used for amplification as described above. Elongated extension time enables full vector amplification.

PCR Condition	
Annealing Temp.	65°C
Annealing Time	20 sec
Extension Time	2 min 30 sec
No. of Cycles	50

The PCR product was cleaned using Gel/PCR DNA fragment extraction kit. The PCR product was restricted with the *Dpn*I restriction enzyme leading to the cutting of the template pQE30.ppxB/L.link while the PCR product remains intact (as outlined in

Section 2.12.1). Phosphorylation of the primers facilitated ligation of the PCR product to form a plasmid which was then called pQE30.ppxB/L.link. The ligated DNA was transformed to *E. coli* XL10-Gold and the transformation was plated on LBamp media (as outlined in Section 2.9.3). Restriction analysis cannot differentiate between re-circularised template DNA and pQE30.ppxB/L.linkfix as all restriction sites were identical. The presence of the correct sequence within the pQE30.ppxB/L.linkfix plasmid was confirmed by DNA sequencing. Translation of protein from this expression vector results in N-terminal 6HISlinker tagged ppxB/L.

4.4 Expression and purification of Photopexin A lectin region

4.4.1 Effect of *E. coli* host strain on the recombinant expression and purification of ppxA/L

Recombinant 6HIS N-tagged ppxA/L was expressed from the pQE30.ppxA/L plasmid from *E. coli* XL10-Gold using the standard procedures established for the expression and purification of 6HIS N-tagged PA-IIL (as outlined in Section 2.15). A pQE30.ppxA/L expression culture was prepared and cleared lysate sample was obtained as per the methodology described in Section 2.15.1. ppxA/L was purified by IMAC purification over Ni-NTA resin (Section 2.16.1). The elution profile of 6HIS N-tagged ppxA/L was determined using lysis buffer (Section 2.3) supplemented with increasing concentrations of imidazole up to 350 mM imidazole. The resulting profile is shown in Figure 4.24.

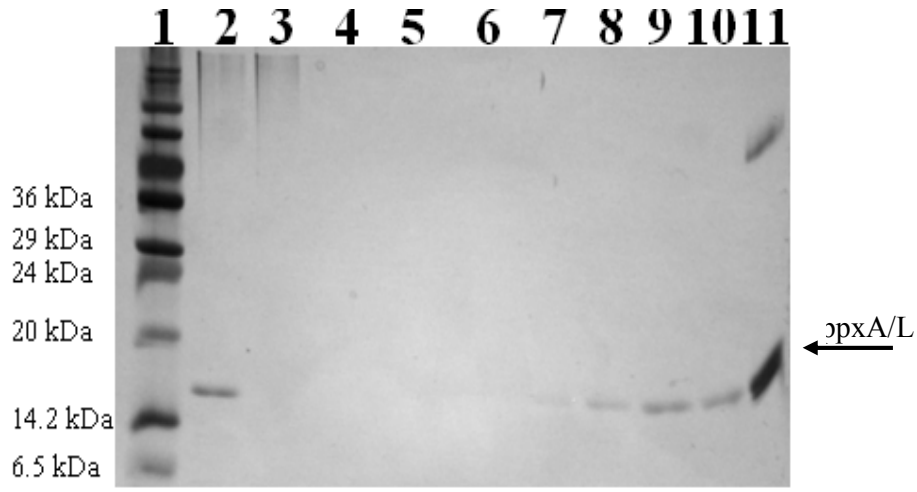


Figure 4.24: SDS PAGE analysis of wash optimisation of ppxA/L
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxA/L, expressed from pQE30.ppxA/L construct *E. coli* XL10-Gold, purification using IMAC (Section 2.16.1). Lane 1, Mr (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash; Lane 5, 40mM imidazole wash; Lane 6, 60mM imidazole wash, Lane 7, 80mM imidazole wash; Lane 8, 100mM imidazole wash; Lane 9, 120mM imidazole wash; Lane 10, 160mM imidazole wash; Lane 11, 350mM elution buffer.

From the SDS-PAGE analysis ppxA/L was found to remain bound to the column in the presence of imidazole washes up to 60 mM. Washing with a 30 mM imidazole was found to sufficiently remove all loosely bound proteins. The standardised procedure to purify ppxA/L was established as; two 10 ml 30 mM imidazole wash steps followed by elution with 5 ml 350 mM imidazole buffer. Relatively low yields of ppxA/L were purified from a 100ml culture (~0.2-0.3mg). Expression of ppxA/L in larger scale (500ml culture), using lysosome (1mg/ml) for improved lysis, was found to increase yields of ppxA/L up to 1mg as shown in Figure 4.25.

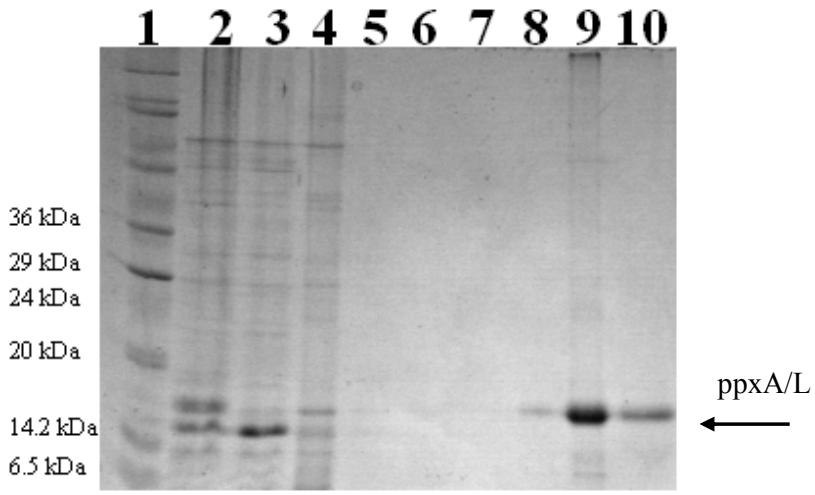


Figure 4.25: SDS-PAGE analysis of large scale ppxA/L purification.
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxA/L, expressed from pQE30.ppxA/L construct *E. coli* XL10-Gold, purification using IMAC (Section 2.16.1). Lane 1, Mr (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash 1; Lane 5, 30mM imidazole wash 10; Lane 6, 30mM imidazole wash 20, Lane 7, 30mM imidazole wash 30; Lane 8 - 10, 350mM imidazole elution buffer.

In both small and large scale, the yield of recombinant ppxA/L was roughly 20% of that obtained for recombinant PA-IIL under the same conditions. The pQE30.ppxA/L plasmid (Figure 4.15) was transformed into the *E. coli* strain KRX. KRX (Table 2.1) is an *E. coli* strain deficient in OmpT and OmpP proteases, purportedly allowing a higher recovery of heterologous recombinant protein. Protein expression levels in KRX were also reported to be as high as levels found for the 'leaky' expressor strain BL21 (DE3). A number of freshly transformed clones were randomly selected and 10ml expression cultures of each clone chosen were grown in LBamp broth containing 50µM IPTG at 37^{o}C overnight. Figure 4.26 shows the SDS-PAGE analysis of the soluble fractions of ppxA/L expression from *E. coli* KRX. Clear bands corresponding to an over expressed protein could be seen corresponding to the expected MW of 13,949 Da in all of the clones selected.

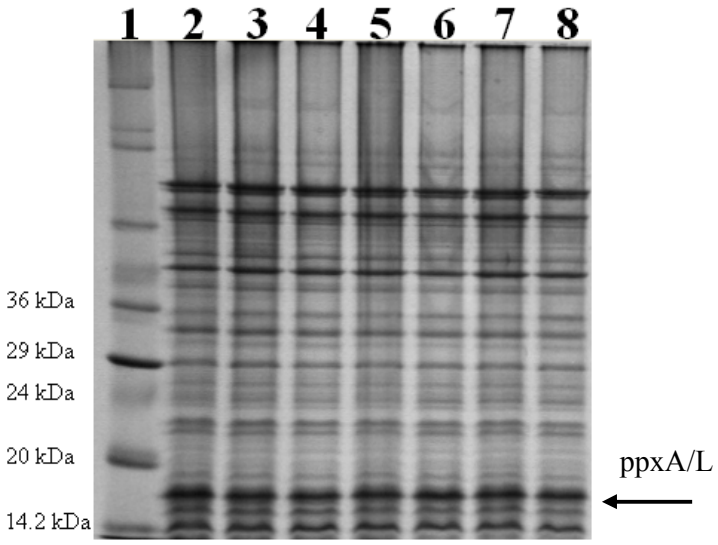


Figure 4.26: Expression cultures from the pQE30.pppxA/L plasmid from *E. coli* KRX.
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxA/L expressed from the pQE30.ppxA/L plasmid *E. coli* KRX. Lane 1, Mr (sizes given in Figure 2.8); Lane 2 -8, Cleared lysate of clone 1-7.

Expression cultures of 100 ml and 500 ml were prepared as described in Section 2.15. Cleared lysate samples were obtained from these expression cultures and analyzed by SDS-PAGE (Figure 4.27). This analysis found no over expression of protein in the larger 500 ml scale.

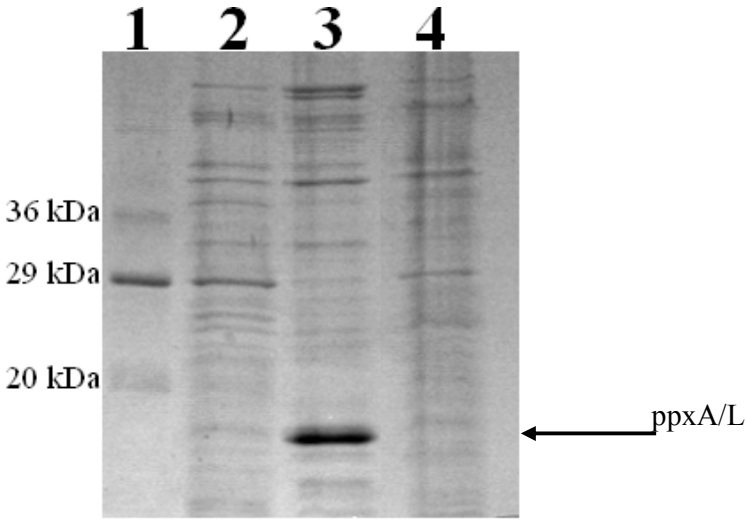


Figure 4.27: Expression of recombinant ppxA/L from pQE30.ppxA/L plasmid from *E. coli* KRX.
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxA/L, expressed from pQE30.ppxA/L plasmid *E. coli* XL10-Gold and KRX. Lane 1, Mr (sizes given in Figure 2.8); Lane 2, Cleared lysate 100 ml XL10-Gold culture; Lane 3, Cleared lysate 100 ml KRX culture; Lane 4, Cleared lysate 500 ml KRX culture.

Wada, K. *et al.* (1992) established the codon usage of *E. coli* from GenBank genetic sequence data. Codons that occurred in *E. coli* with a frequency of <1% were arbitrarily defined as rare codons (Table 4.5).

Table 4.5: Codons used by *E. coli* at a frequency of <1%.
The *E. coli* codon frequency was obtained by Wada, K. *et al.* (1992). Kane, J.F. (1995) identified a subset of rare codons (codons underlined into red) that led to translation problems when expressed from *E. coli.*

Amino acid	Rare Codons
Thr	ACA
Arg	AGG AGA CGA CGG
Leu	CTA CTC

Amino acid	Rare Codons
Ile	ATA
Ser	TCA AGT TCG TCC
Cys	TGT TGC

Amino acid	Rare Codons
Pro	CCC CCT CCA
Gly	GGG GGA

A study by Kane, J.F. *et al.* (1995) identified a subset of rare codons whose presence in large numbers may lead to transcriptional problems leading to a reduction in quantity and quality of protein synthesised. Analysis of the ppxA/L sequence found that the protein has a composition of 6% of rare codons found within the subset related to poor protein composition (Figure 4.28). Expression of rare codons can be improved by use of a host containing a plasmid with the appropriate tRNA. The *E. coli* strain Rosetta carries a plasmid which facilitates rare codon expression.

```
atgagaggatcgcatcaccatcaccatcacggatccatgaataccggtaaacactgcttt
 M  R  G  S  H  H  H  H  H  H  G  S  M  N  T  G  K  H  C  F
caattacctcaagtattcgatttggcctgactgcctacgtaaatagtgatgttcatcaa
 Q  L  P  Q  S  I  R  F  G  L  T  A  Y  V  N  S  D  V  H  Q
caatcaatcaatgtctatatcgatgaccgattggtcgataccttaaccggtaaaggaata
 Q  S  I  N  V  Y  I  D  D  R  L  V  D  T  L  T  G  K  G  I
agtcacataacagatgttaggacctacacatcaggtaccggcaaagtctgcatcgagatg
 S  H  I  T  D  V  R  T  Y  T  S  G  T  G  K  V  C  I  E  M
atagga gagggtaaacccggcaaactccgttatgcctataacactctcgaagcaaaacccc
 I  G  E  G  K  P  G  K  L  R  Y  A  Y  N  T  L  E  A  K  P
ggcacagccattatcggcgctaacaatggttccaatgacaattatgatgatagtgtggta
 G  T  A  I  I  G  A  N  N  G  S  N  D  N  Y  D  D  S  V  V
gtgttgaattggccattatcatga
 V  L  N  W  P  L  S  -
```

Figure 4.28: *E. coli* bias relative to the codons of recombinant ppxA/L
The nucleotide sequence coding for ppxA/L as found for pQE30.ppxA/L. The amino acids corresponding to codon with a frequency less then <1% as outlined in Table 4.5, are highlighted in green. The subset of codons considered rare in *E. coli* are highlighted red.

The pQE30.ppxA/L plasmid was transformed into the *E. coli* Rosetta strain (as outlined in Section 2.9.3). A number of cultures were prepared and cleared lysate samples were obtained by the methodology described in Section 2.15.1. Figure 4.29 shows the SDS-PAGE analysis of the cleared lysate.

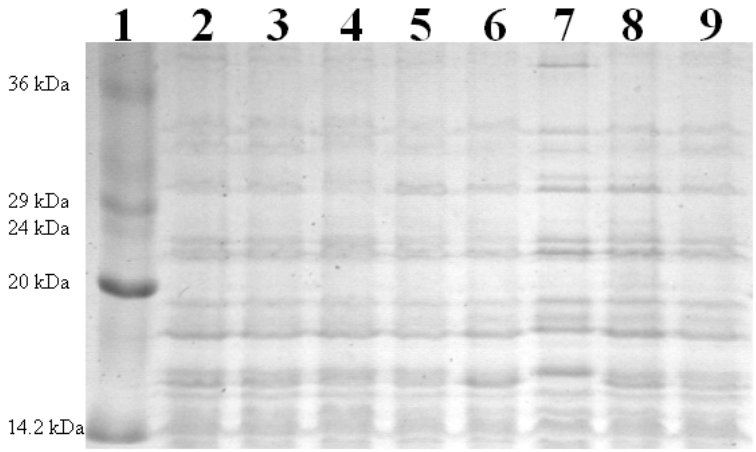


Figure 4.29: SDS-PAGE analysis of the expression of recombinant ppxA/L from *E. coli* Rossetta strain
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxA/L from the pQE30.ppxA/L plasmid in *E. coli* Rosetta. Lane 1, Mr (sizes given in Figure 2.8); Lane 2 -9, Cleared lysate of clone 1-8.

From the results above it was concluded that the use of Rosetta strain *E. coli* did not significantly improve expression of ppxA/L.

4.4.2 Effect of induction point on the expression of *ppxA/L*

Cultures of *E.* coli XL10-Gold and KRX containing the pQE30.ppxA/L expression plasmid were induced (with 50µM IPTG) at various times from OD_{600} of 0.6 to 1.2 to investigate the effect, if any, that time of induction had on expression. Cleared lysate samples for SDS-PAGE analysis were taken 24hrs after induction as shown in Figure 4.30.

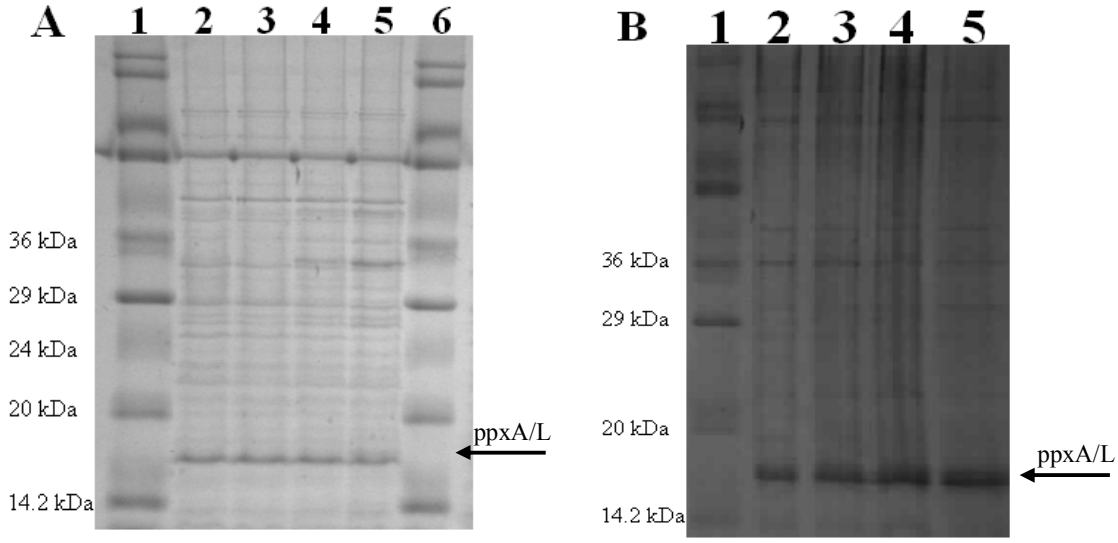


Figure 4.30: Effect of induction time on ppxA/L expression
Analysis of ppxA/L expression in *E. coli* XL10-Gold (A) and KRX (B) as outlined in Section 2.24, from pQE30.ppxA/L (Figure 4.5) by 15% SDS PAGE. Point of induction monitor by measuring OD at 600nm Lane 1, Mr (sizes given in Figure 2.8); Lane 2 OD_{600} of 0.6; Lane 3, OD_{600} of 0.8; Lane 4, OD_{600} of 1.0; Lane 5, OD_{600} of 1.2

Induction of pQE30.ppxA/L cultures after an OD_{600} 0.6 was found not to significantly increase the yield of ppxA/L expressed. The standard protocol for the expression and purification of recombinant ppxA/L to obtain sufficient quantities for future work was established to be largely similar to PA-IIL;

Expression: Growth of pQE30.ppxA/L *E. coli* KRX at 37^{o}C to OD_{600} 0.6, induction with 50 µM IPTG and expression at 30^{o}C overnight

Purification: 100ml x 2 cultures, 30 ml 30 mM imidazole wash, elution with 350 mM imidazole elution buffer.

Figure 4.31 represents a typical purification using the above standard procedures.

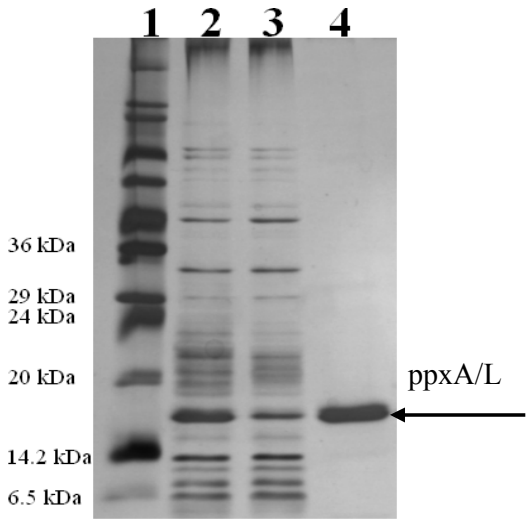


Figure 4.31: SDS analysis of the purification of 6HIS N-tagged ppxA/L by IMAC
Analysis by 15% SDS-PAGE, silver stained, of recombinant 6HIS N-tagged ppxA/L purification using IMAC (Section 2.16.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash; Lane 5, 350mM imidazole elution .

4.4.3 Biological activity of ppxA/L

The activity of ppxA/L was determined by ELLA (as outlined in Section 2.27). ppxA/L was found to bind to BSA. The binding specificity of ppxA/L is unknown. The lectins within the PA-IIL superfamily are predominately fucose or mannose binding lectins. Therefore the soluble sugars fucose and mannose were added to competitively inhibit the binding activity of ppxA/L to BSA. Both fucose and mannose were found to quench the binding activity (Figure 4.32), therefore the binding activity can be said to be sugar specific. ppxA/L was also stripped of Ca^{2+} ions through the addition of the chelating agent EDTA resulting in the loss of binding activity. This confirms that ppxA/L has a dependency on Ca^{2+} for binding as was found for PA-IIL.

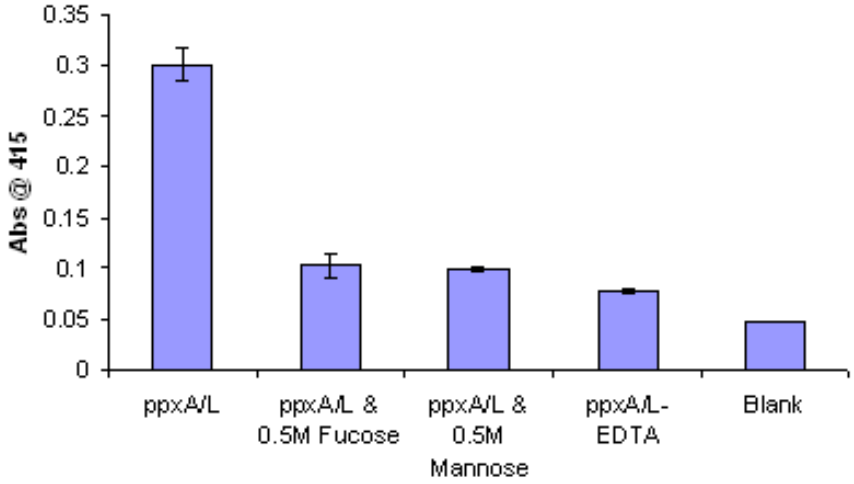


Figure 4.32: Histogram of ppxA/L activity assessed by ELLA
ELLA carried out as per Section 2.27. Histogram; PBS used in place of ppxA/L as blank control. 2.5% BSA placed in each well and 10 µg/ml of 6HIS N-tagged ppxA/L allowed to react with each well. The effect of the removal of the calcium ions from the lectin binding pocket by EDTA on lectin binding to BSA was evaluated. Removal of the calcium ion resulted in a loss of binding signal, confirming that Ca^{+} ions are required for binding. The binding signal was also found to be quenched in the presence of 0.5 M fucose and mannose. The binding signal was therefore attributed to sugar specific binding. (Error bars represent mean ± SD, n = 3).

4.4.4 Protein stability of purified recombinant ppxA/L

Similar to PA-IIL, ppxA/L was found to precipitate in solution after a short period of time when highly purified. For long term storage ppxA/L was lyophilised (see Section 2.21). For short term storage (> 1 day) ppxA/L was found to be most stable at room temperature.

4.5 Expression and purification of Photopexin B lectin region

Recombinant 6HIS N-tagged ppxB/L was expressed from the pQE30.ppxB/L plasmid (Figure 4.16) in the *E. coli* strain KRX as this strain was found to be the optimal expression strain for ppxA/L. A number of freshly transformed clones from pQE30.ppxB/L *E. coli* KRX were randomly picked. 10ml expression cultures of each clone chosen were grown in LBamp broth containing 50µM IPTG at 37^{o}C overnight. Figure 4.33 shows the SDS-PAGE analysis of the soluble fractions of ppxB/L expression from *E. coli* KRX. A clear band corresponding to an over expressed protein could be seen in lane 9 (clone 8), corresponding to the expected MW of 14,185 Da of ppxB/L.

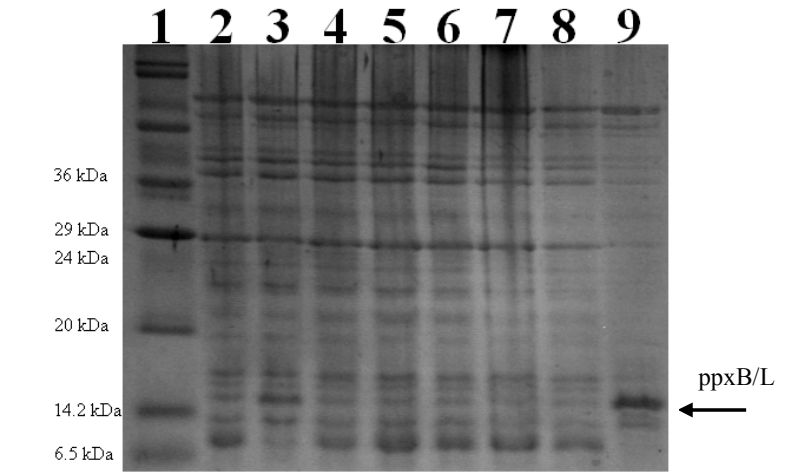


Figure 4.33: Expression cultures from pQE30.ppxB/L plasmid from *E. coli* KRX
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxB/L from the pQE30.ppxB/L plasmid in *E. coli* KRX. Lane 1, Mr (sizes given in Figure 2.8); Lane 2 -9, Cleared lysate of clone 1-8.

A 100ml expression culture of clone 8 was carried out as described in Section 2.15. The elution profile for the purification of ppxB/L was investigated with a series of washes containing increasing concentrations of imidazole (up to 350mM). The resulting profile is shown in Figure 4.34.

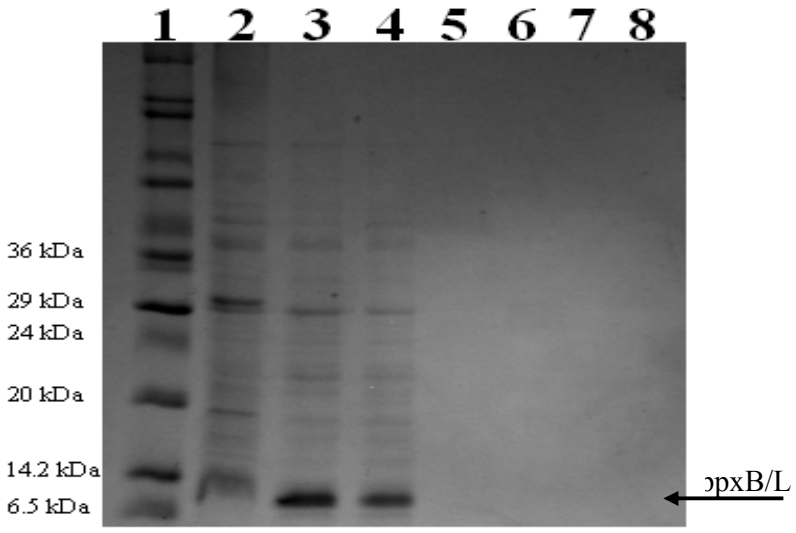


Figure 4.34: Purification of 6HIS N-tagged ppxB/L by IMAC
Analysis by 15% SDS-PAGE, of recombinant 6HIS N-tagged ppxB/L purification by IMAC (Section 2.16.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash; Lane 5, 40 mM imidazole wash; Lane 6, 60 mM imidazole wash; Lane 7-8, 350mM imidazole elution 1 and 2.

The elution profile indicates that ppxB/L is not binding to the Ni-NTA resin as all protein is seen to be in the flowthrough and wash fractions. In order to make the poly-his tag more accessible a linker was cloned into the pQE30.ppxB/L plasmid to form the

pQE30.ppxB/L.link plasmid (Figure 4.20). The pQE30.ppxB/L.link vector was transformed into the *E. coli* strain XL10-Gold. A number of clones were expressed and cleared lysate samples were obtained from each culture as outlined by the methodology described in Section 2.15.1. Figure 4.35 shows the SDS-PAGE analysis of the soluble fractions of ppxB/L expression from *E. coli* XL10-Gold. No clear bands corresponding to an over expressed protein could be seen which was attributed to expression of low protein yields.

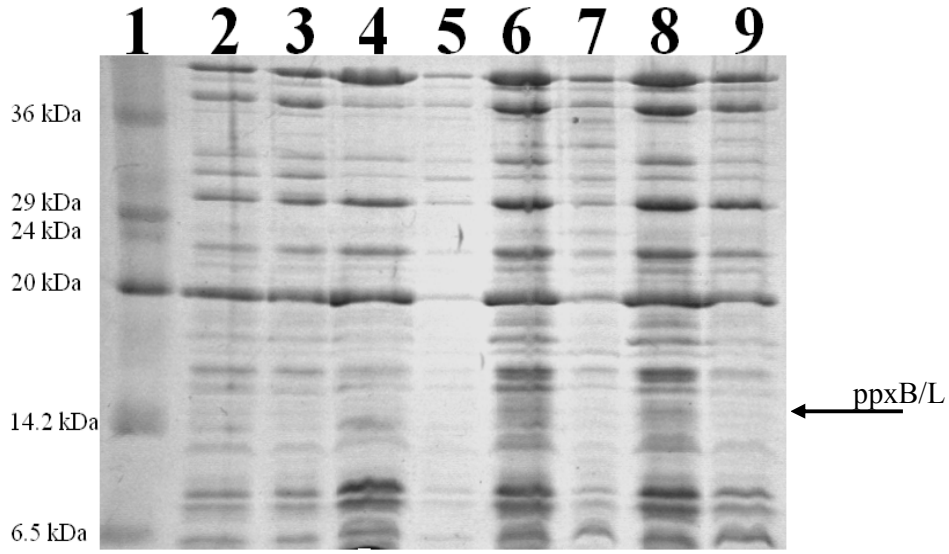


Figure 4.35: Expression cultures of pQE30.ppxB/L-link from *E. coli* XL10-Gold
Analysis by 15% SDS-PAGE of recombinant 6HIS N-tagged ppxB/L expressed from pQE30.ppxB/L.link construct *E. coli* XL10-Gold. Lane 1, Mr (sizes given in Figure 2.8); Lane 2 -9, Cleared lysate of clone 1-8

A 100ml expression culture of ppxB/L from the pQE30.ppxB/L.link vector was carried out and cleared lysate was obtained (Section 2.15) and purified over Ni-NTA resin as outlined in Section 2.16.1.

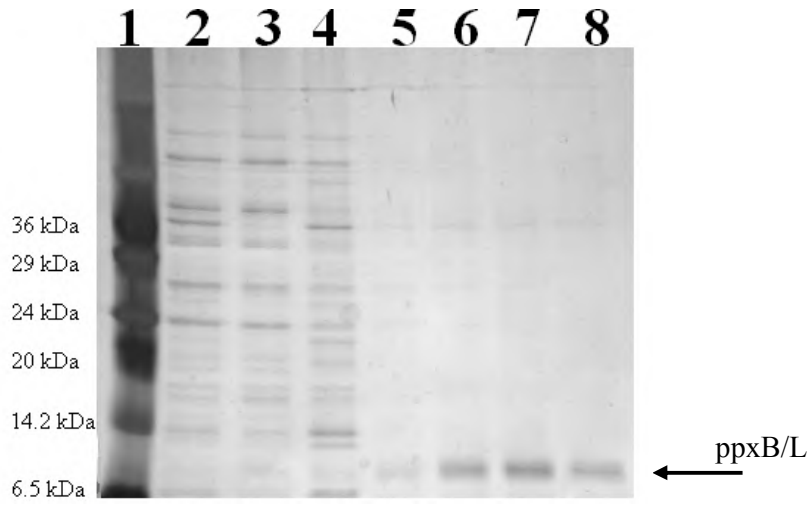


Figure 4.36: Purification of recombinant ppxB/L from the pQE30.ppxB/L.link construct
Analysis by 15% SDS-PAGE, silver stained, of recombinant 6HIS N-tagged ppxB/L-link purification using IMAC (Section 2.16.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash; Lane 5-8, 350mM imidazole elution 1-2.

A purified protein band is visible, although yield of protein appears very low, running between the 6.5 and 14.2 kDa markers (Figure 4.36) at approximately 8-10 kDa. The predicted molecular weight of 6HIS N-tagged ppxB/L is 14,185 Da respectively as deduced from the amino acid sequence. The expected protein size and the protein size found by SDS-PAGE gel analysis did not correlate. The pQE30.ppxB/L-link vector was DNA sequenced to confirm the sequence of the *ppxB/L* gene. DNA sequencing identified a frame shift in the pQE30.ppxB/L.link plasmid leading to the early introduction of a stop codon (Figure 4.21). The plasmid pQE30.ppxB/L.linkfix was cloned to correct the frame shift (see Section 4.3.3). The pQE30.ppxB/L.linkfix plasmid was transformed into the *E. coli* strains KRX and XL10-Gold (see Section 2.9.3). A number of freshly transformed clones of pQE30.ppxB/L.linkfix from each strain of *E. coli* were randomly selected. 10ml expression cultures of each clone chosen were grown in LBamp broth containing 50µM IPTG at 37^{o}C overnight. Figure 4.37 shows the SDS-PAGE analysis of the soluble fractions of ppxB/L expression from each *E. coli* strain.

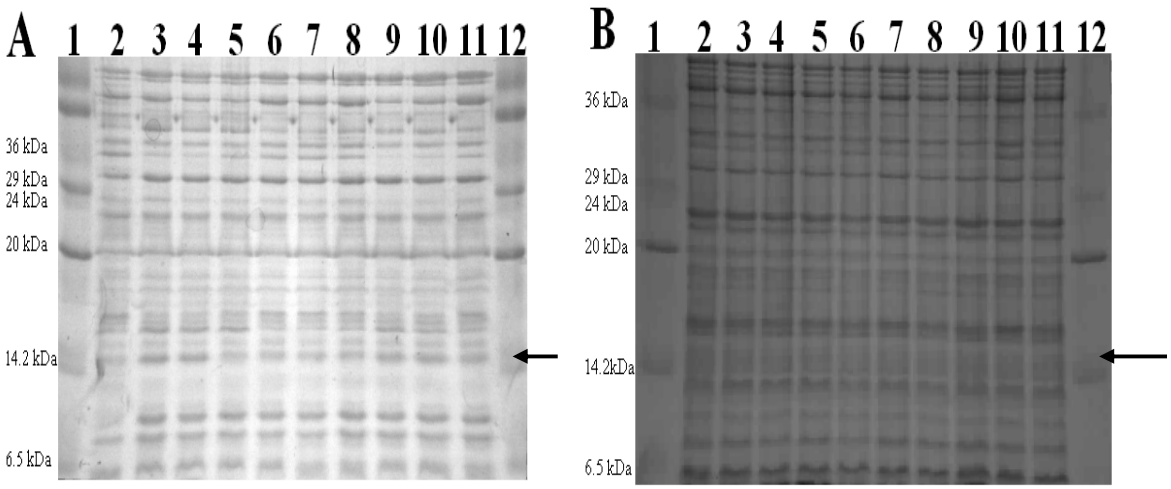


Figure 4.37: Expression of pQE30.ppxB/L.linkfix from *E. coli* strains KRX and XL10-Gold

Analysis by 15% SDS-PAGE of recombinant 6HIS-linker N-tagged ppxB/L expressed from pQE30.ppxB/L.linkfix plasmid from *E. coli* strain [A] KRX and [B] XL10-Gold. Lane 1, Mr (sizes given in Figure 2.8); Lane 2 -11, Cleared lysate of clone 1-10; Lane 12, Mr Ladder

No prominent band indicating the over expression of a protein of expected M.W of 13.87 kDa was found within the soluble fractions when ppxB/L.link was expressed from the pQE30.ppxB/L.linkfix plasmid in either of the *E. coli* strains KRX or XL10-Gold. The insoluble fractions of XL10-Gold pQE30.ppxB/L.linkfix cultures were analysed by SDS-PAGE for the presence of recombinant ppxB/L.link (Figure 4.38).

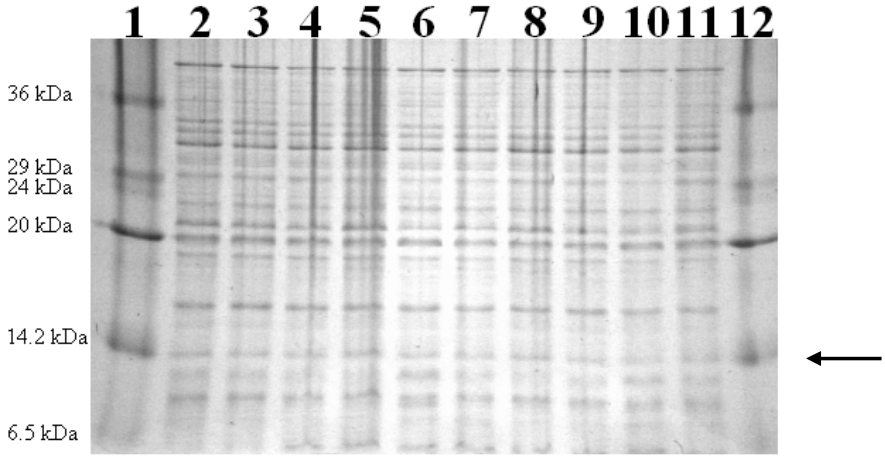


Figure 4.38: Expression of recombinant ppxB/L.link from the pQE30.ppxB/L.linkfix *E. coli* XL10-Gold insoluble fraction

Analysis by 15% SDS-PAGE of recombinant 6HIS-link N-tagged ppxB/L expressed from pQE30.ppxB/L.linkfix plasmid *E. coli* strain XL10-Gold, insoluble fraction. Lane 1, Mr (sizes given in Figure 2.8); Lane 2 -11, Cleared lysate of clone 1-10; Lane 12, Mr Ladder

No significant quantity of recombinant ppxB/L was found within the insoluble fraction. Recombinant ppxB/L-link was expressed from the pQE30.ppxB/L.linkfix plasmid from *E. coli* XL10-Gold using the standard procedures established for the expression and

purification of ppxA/L. Expression cultures were prepared as per the methodology described in Section 2.15.1. A cleared lysate sample was obtained from the pQE30.ppxB/L.linkfix expression culture as outlined in Section 2.15.2. The yields of ppxB/L obtained in this manner were not sufficient for future characterisation work.

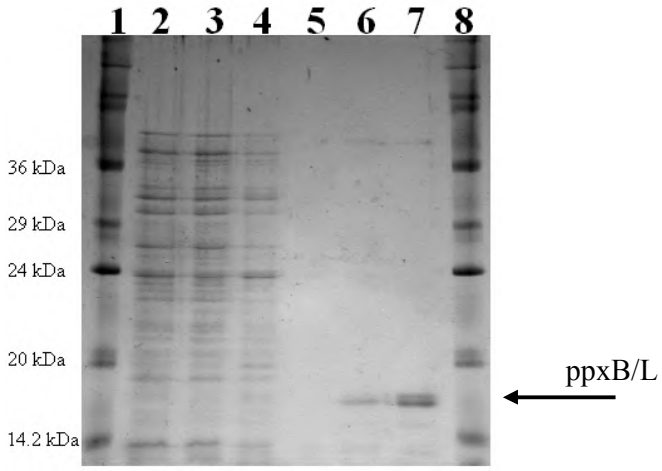


Figure 4.39: Purification of recombinant ppxB/L.link by IMAC
Analysis by 15% SDS-PAGE, silver stained, of recombinant 6HIS N-tagged ppxB/L.link purification using IMAC (Section 2.16.1). Lane 1, Mr Ladder (sizes given in Figure 2.8); Lane 2, Cleared lysate; Lane 3, flow-through; Lane 4, 30mM imidazole wash; Lane 5-8, 350mM imidazole elution 1-2.

4.6 Physiochemical characterisation of ppxA/L

To determine the quaternary structure of recombinant 6HIS N-terminally ppxA/L, size exclusion chromatography was employed using a manually poured Toyopearl HW-55S (Tosoh Biosciences, Germany) and a pre-packed 25ml Superdex™ 75 column (GE Healthcare) (as outlined in Section 2.25). Size exclusion chromatography separates proteins based on molecules hydrodynamic radius. Proteins of lower relative molecular mass are retarded within the porous resin beads and are thus separated from proteins of larger relative molecular mass which are excluded from smaller pores. The larger proteins pass through the packed column more rapidly than the smaller proteins and their exact size can be determined by comparison with a standard curve created using a set of known standards.

4.6.1 Determination of lectin size using a Toyopearl HW-55S column

Toyopearl HW-55S resin is a column matrix composed of a methyl-methacrylate and has no carbohydrate component (such as that found in Sepharose) and is therefore suitable for examination of carbohydrate binding molecules such as lectins. An HW55S Toyopearl resin size exclusion chromatography column was manually poured. It is not

uncommon for air-pockets or ineffective packing to occur during this process and as a result an additional validation step was carried out. The effectiveness of the packing procedure was validated by injecting a 0.5ml 0.1% (v/v) actone sample onto the column (Figure 4.40) and the column plate count (N) was then determined.

$$N = 5.54\left[\frac{V_e}{V_{1/2}}\right]^2$$

V_e - retention time
$V_{1/2}$ - peak width at half the peak height

Equation 4.1: Equation to determine theoretical plates (N) for validation of Toyoperal HW-55S column.

The values for V_e and $V_{1/2}$ were measured from the acetone elution profile below and imputed into the equation 4.1.

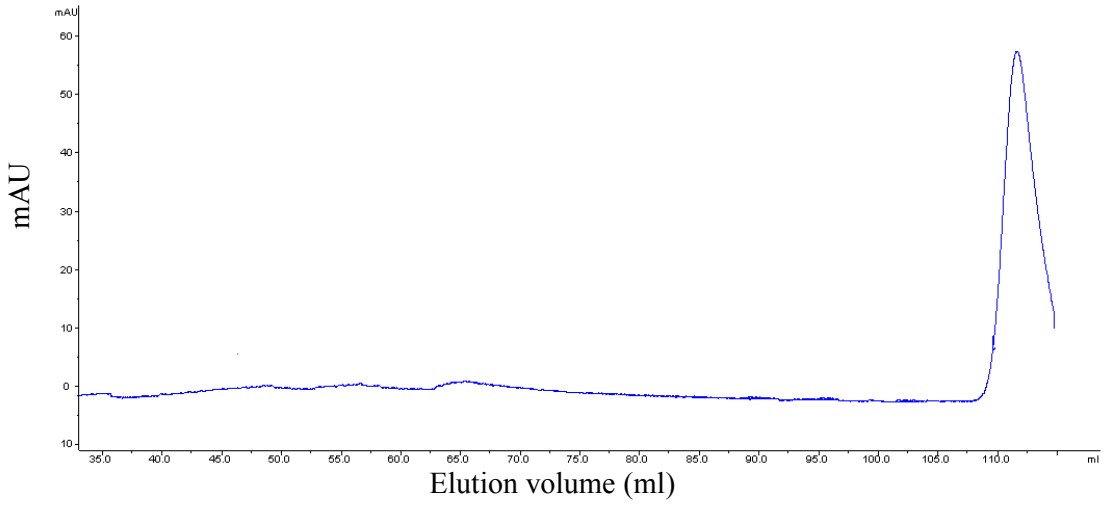


Figure 4.40: Determination of the column efficiency of the HW55S Toyopearl column.
The elution profile of acetone (0.1% w/v) using size exclusion chromatography with Toyopearl HW55S column. The OD_{280nm} of eluted acetone is plotted against the elution volume. Void volume (elution volume of acetone) was determined to be 111.66ml.

The number of theoretical plates was determined, using the equation outlined in Equation 4.1, to be 6962, which was within the operable range (3500 – 10000) for this resin (as outlined in the Toyopearl Instruction manual). The columns performance was checked regularly as manually poured columns are susceptible to compaction over time. Once validated a standard curve was created using five known standards; bovine serum albmin (BSA), ovalbumin, bovine carbonic anhydrase, cytochrome C and myoglobin.

The K_{av} value, determine by equation 3.1, was plotted against the log of the MW of each protein to create a standard curve (Figure 4.41). The molecular weight range of the standards encompassed the predicted molecular weight of monomeric and tetrameric ppxA/L.

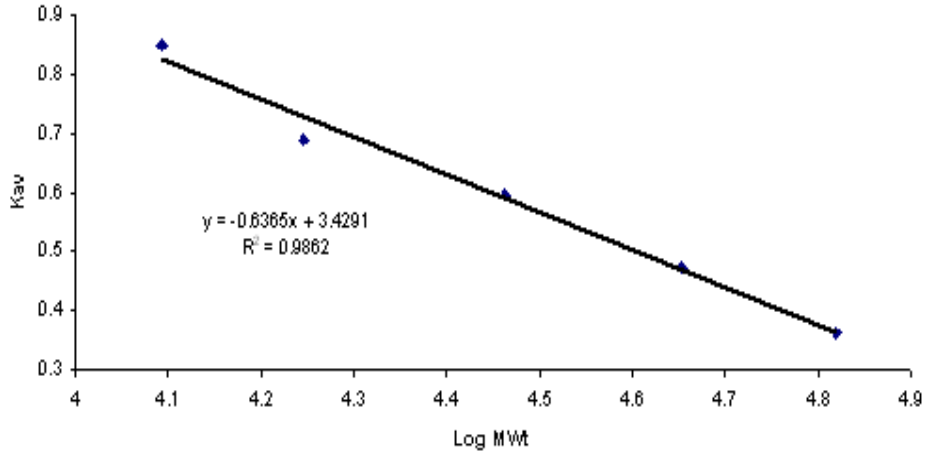


Figure 4.41: Development of size exclusion chromatography standard curve for the Toyopearl HW55S column.
The protein molecular weight standard curve was created using size exclusion chromatography, with the Toyopearl HW55S column. Void volume (V_o) of 42.42 ml was determined from the elution volume of blue dextran, and total volume (V_t) was 111.66 ml, determined with acetone. The elution volumes of protein standards are given in Table 4.5. These values were used to calculate the relative MW of all lectin samples that were eluted in this range. The K_{av} value was calculated using Equation 3.1.

Table 4.6: Construction of a protein molecular weight standard curve for the Toyopearl HW55S size exclusion chromatography column at pH 8.0 in PBS.

Standard	Elution volume (ml)	Expected MW	logMW	Kav	Actual Mw	% Error
Cytochrome C	74.2	12400	4.09	0.85	11339	8.56
Myoglobin	70.71	17600	4.25	0.69	20047	-13.90
Carbonic Anhydrase	68.25	29000	4.46	0.60	28353	2.23
Ovlbumin	64.62	45000	4.65	0.47	44256	1.65
BSA	60.85	66000	4.82	0.36	65699	0.46

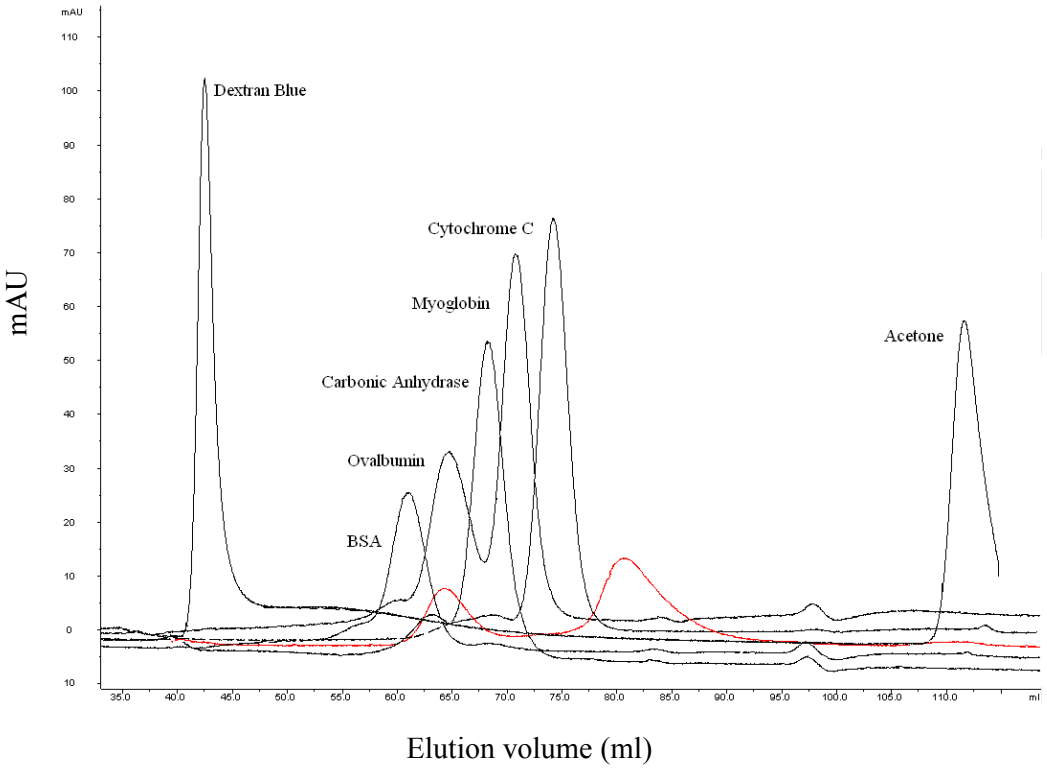


Figure 4.42: Elution of protein standards from the Toyopearl HW-55S SEC column
The elution profiles of standard proteins using size exclusion chromatography with the Toyopearl HW55S column shown in black, the elution profile of ppxA/L shown in red. The optical density at OD_{280nm} of eluted standard proteins was plotted against the elution volume. The elution volume for each standard protein was used in the construction of the stand curve (Figure 4.38).

Having successfully created a standard curve for the Toyopearl SEC column, with an R^2 value for the curve of 0.98, the relatitvie molecular mass of ppxA/L could be estimated. The elution profile of purified recombinant ppxA/L over the Toyopearl column consisted of two elution peaks, one of which lies outside of the range of the standard curve (Figure 4.42).

Table 4.7: Calculated molecular weight of ppxAL using a Toyopearl HW-55S SEC column

Lectin	Retention time (ml)	Theoretical monomer size (Da)	Calculated MW
ppxA/L 1st peak	64.24	13,949	46185.3
ppxA/L 2nd peak	80.68	13,949	2799.9

To determine if protein-carbohydrate binding was occurring between ppxA/L and the column matrix the lectin was applied to the column under normal conditions in the presence of the competing sugar, D-mannose, for which ppxA/L is predicted to bind. Figure 4.43 shows the elution profile of ppxA/L in replete and delete mannose conditions.

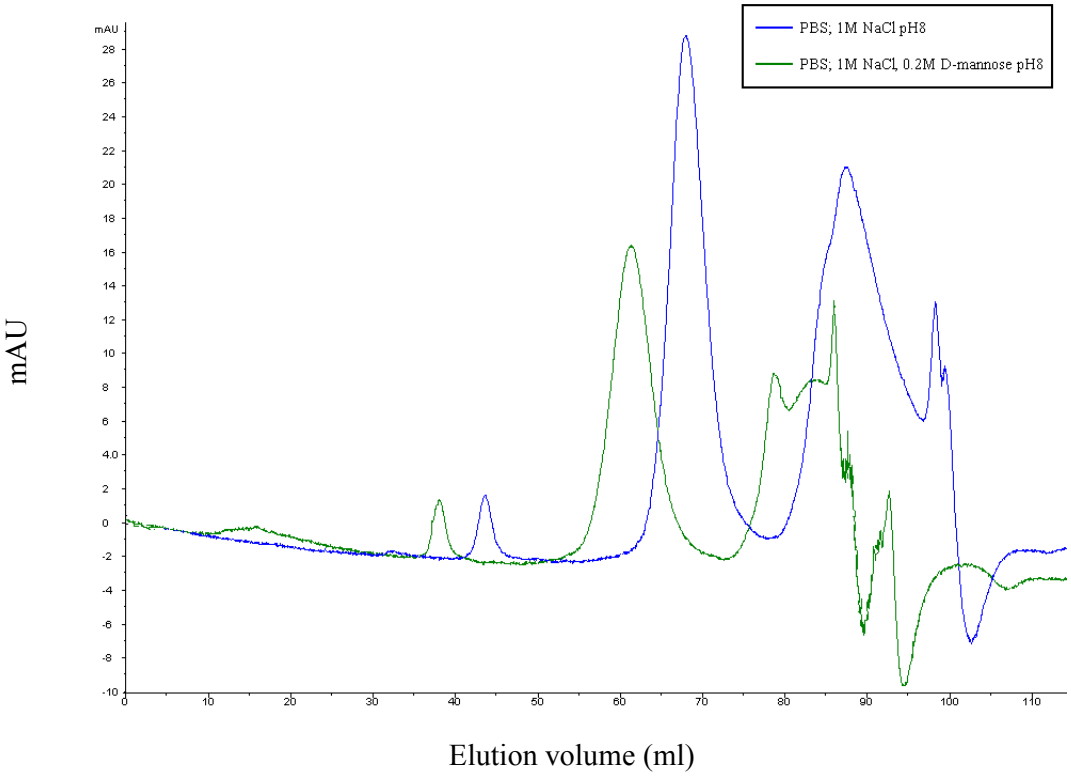


Figure 4.43: The adsorption OD_{280nm} of ppxA/L through the Toyopearl HW-55S column in the presence/absence of mannose.
The elution profile of ppxA/L is highlighted in blue, the profile of ppxA/L in 0.2M D-mannose is highlighted in green. Peak maxima are 67.69/61.69 and 86.92/78.31 respectively.

It was found that there was some variation between the elution volumes of ppxA/L in mannose replete and deplete conditions (Figure 4.43). The variation between the profiles is attributed to some compression of the resin. The column resin was packed using PBS. Alteration of the mobile phase, such as the addition of mannose, changes the properties of the mobile phase, e.g. viscosity of the buffer, and results in additional compression of the resin. Additional compression of the resin leads to a reduced bed height and reduced elution times. The variation of elution profiles therefore is not due to the inhibition of the sugar-specific interaction with the column matrix as the presence of the two peaks persists. The secondary peak present in the elution profile of ppxA/L does not appear to be due to sugar-specific interaction. Toyopearl HW55S was deemed

an unsuitable resin for the determination of the molecular weight of ppxA/L as it appeared as though the protein was non-specifically interacting with the resin. An alternative resin was investigated for suitability of determination of the molecular weight of ppxA/L.

4.6.2 Determination of lectin size using a Superdex™ 75 column

A second gel filtration matrix was used to examine the protein size of ppxA/L. Superdex is composed of a highly cross-linked agarose beads, to which dextran is covalently bonded and has no carbohydrate element. The prepacked Superdex™ 75 10/100 column from GE Healthcare, separates proteins based on their relative size and has a globular protein separation range of 3,000 – 70,000 Da. A molecular weight standard curve was constructed as before (Section 4.6.1).

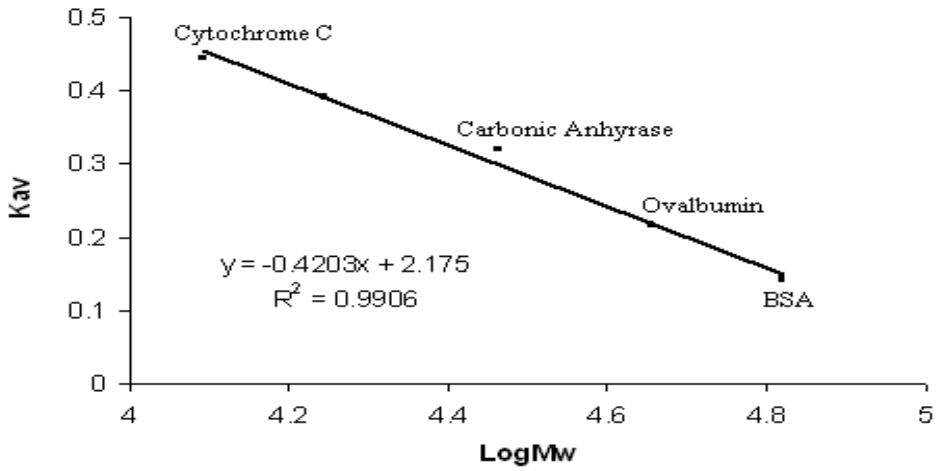


Figure 4.44: Development of size exclusion chromatography standard curve for the Superdex 75 high performance column.
The protein molecular weight standard curve was created using size exclusion chromatography, with the Superdex 75 column. Void volume (V_o) of 7.56 ml was determined from the elution volume of blue dextran, and total volume (V_t) was 19.13 ml, determined with acetone. The elution volumes of protein standards are given in Table 4.8. These values were used to calculate the relative MW of all lectin samples that were eluted in this range. The K_{av} values were calculated using Equation 3.1.

Table 4.8: Construction of protein molecular weight standard curve for the Superdex 75 high performance column at pH 7.8 in PBS.

Standard	Elution volume (ml)	MW	logMW	Kav	Actual Mw	% Error
BSA	9.19	66000	4.82	0.14	69132	4.75
Ovlbumin	10.07	45000	4.65	0.22	45574	1.23
Carbonic Anhydrase	11.26	29000	4.46	0.32	25942	10.54
Myoglobin	12.08	17600	4.25	0.39	17595	-0.03
Cytochrome C	12.71	12400	4.09	0.45	13057	5.30

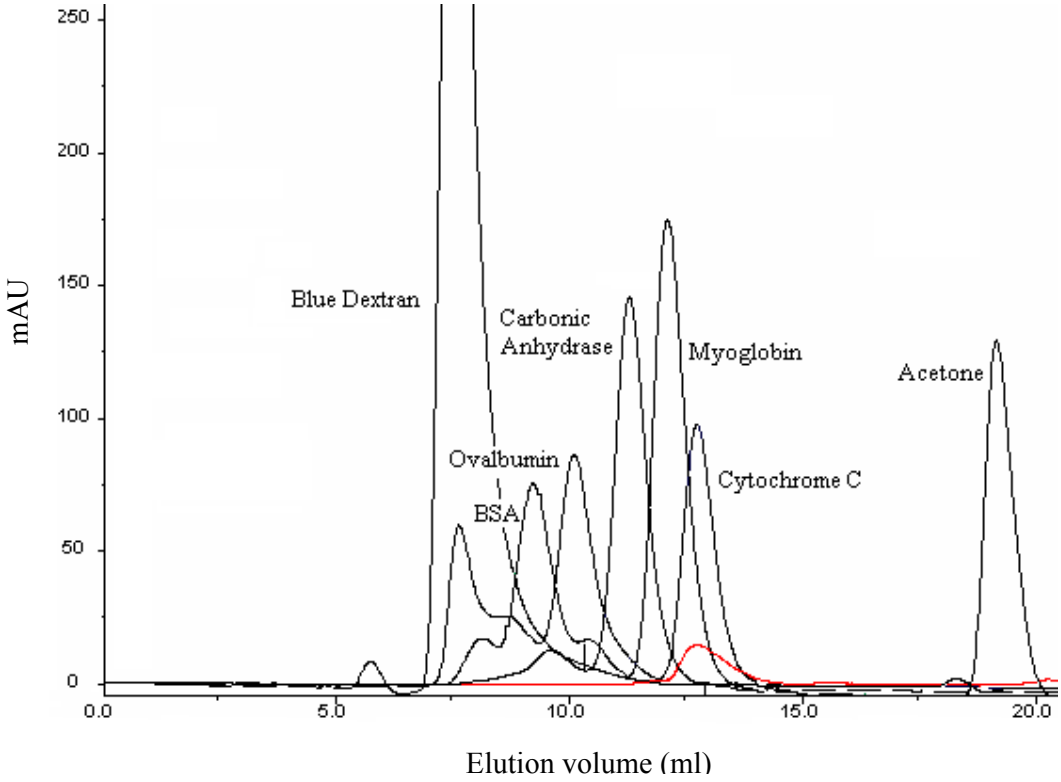


Figure 4.45: Elution of protein standards from the Superdex column
The elution profiles of standard proteins using size exclusion chromatography with the Superdex column shown in black, the elution profile of ppxA/L shown in red. The optical density at OD_{280nm} of eluted standard proteins was plotted against the elution volume. The elution volume for each standard protein was used in the construction of the stand curve (Figure 4.43).

Having successuflly created a standard curve for the Superdex column (Figure 4.46), the relatitvie molecular mass of ppxA/L could be estimated. As outlined in Section 2.25.2, 100 ul of a freshly purified ppxA/L was buffer exchanged to PBS and was analysed under native conditions to ascertain its multimeric structure. The peak heights of the elution profiles were converted into relative molecular masses using the standard curve created (see Figure 4.47).

Table 4.9: Predicted molecular weight of ppxA/L and number of subunits from Superdex75 column

Lectin	Retention time (ml)	Theoretical monomer size (Da)	Calculated Mw (Da)	No. of subunits
ppxA/L	12.63(±0.04)	13,949	13,560	0.97

The molecular weight calculated by measuring the average elution volume at which the maximum peak height was eluted from four separate runs found ppxA/L to be just below the predicted size of a monomer. It is theorized that as the standard curve was created using globular proteins it serves as only a rough estimate as to the size of non-globular shaped proteins, and ppxA/L can therefore be said to be approximately a monomer.

The lectin ppxA/L was found to produce a two peak elution profile representing a monomer and a multimer, when the lectin was run over the Superdex column in the presence of imidazole (350 mM from the elution buffer). Figure 4.46 shows the elution profile of ppxA/L in the presence of imidazole. When ppxA/L was purified and analysed by SDS PAGE a multimer band can be seen to be present which run at approx 55-66 kDa (Figure 4.24). This correlates with the molecular weight determined by size exclusion chromatography which was estimated at 63.7 kDa (see Table 4.10). Imidazole has been recognized as a potential new additive for the stabilization of proteins and has been demonstrated as a suitable catalyze for in vitro re-folding (Shi et al. 2007). Therefore the presence of imidazole is likely to be facilitating the formation of multimeric ppxA/L.

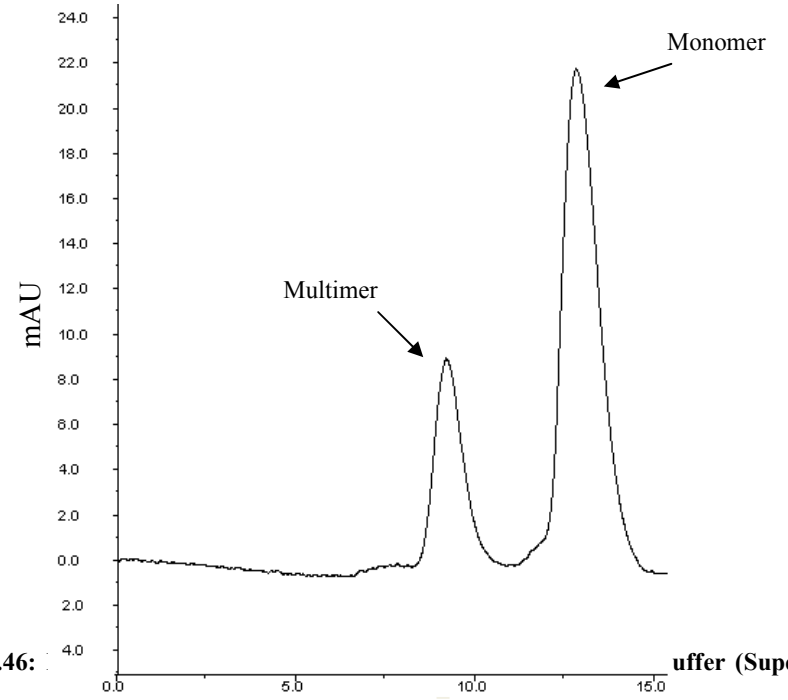


Figure 4.46: uffer (Superdex 75 column matrix)

The two peak elution profile of ppxA/L using SEC with the Superdex 75 column. The optical density of OD_{280nm} of the eluted protein plotted against the elution volume.

Table 4.10: Predicted molecular weight of ppxA/L in the presence of imidazole from a Sephadex 75 column

Lectin	Retention time (ml)	Theoretical monomer size (Da)	Calculated MW	No. of subunits
ppxA/L Peak 1	9.8	63,739	13,949	4.57
ppxA/L Peak 2	12.8	11,221	13,949	0.80

The calculated relative molecular masses obtained for both peaks of ppxA/L in the presence of imidazole are not accurate but an approximation. They were calculated using a standard curve created with a running buffer of PBS. In order to obtain accurate results a standard curve created with a running buffer containing imidazole would be required. Although the calculated relative molecular masses cannot be said to be exact it does imply that ppxA/L is capable of forming multimers.

4.7 Discussion

The pQE vectors from Qiagen were found to be suitable expression vectors for the expression of PA-IIL. The cloning and optimal expression of PA-IIL was predominately carried out using the expression vector pQE30. Therefore ppxA/L was cloned into the

pQE30 vector. Sequence confirmation of pQE30.ppxA/L plasmid found three points of nucleotide variation between the sequence obtained for pQE30.ppxA/L and the GeneBank sequence of ppxA. These single nucleotide polymorphisms (SNP) resulted in two amino acid changes within the ppxA/L protein. The third variation alters the codon from ACA to ACG, both codons for threonine. These SNP variations in amino acid sequence result in a threonine found in place of a seriene and an aspartic acid in place of a glutamic acid (Figure 4.8). Serine and threonine are both classified as nucleophilic in nature and differ in an additional CH_2 group present on threonine. Glutamic acid and aspartic acid are both classified as acidic in nature and differ in two additional CH_2 groups on glutamic acid (Figure 4.47).

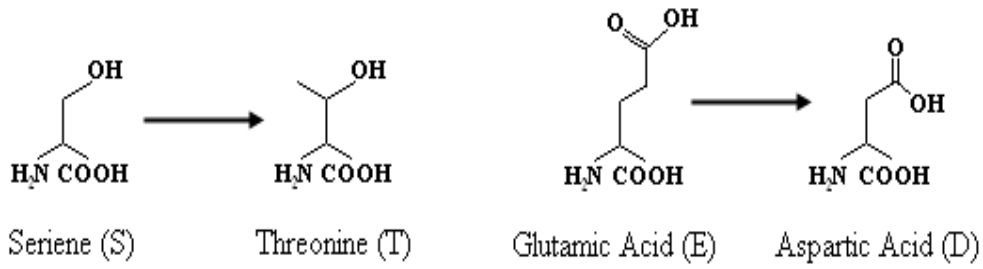


Figure 4.47: Schematic of amino acid structures found to vary between pQE30.ppxA/L and the Genebank ppxA sequence.
The structures of the amino acids found to vary between the sequence found for pQE30.ppxA/L and the sequence obtained for ppxA from GeneBank.

The amino acids found to vary were not present within the conserved calcium binding loop or the sugar binding loop and the amino acids found within the ppxA/L protein are similar in nature to those found within the GeneBank sequence of ppxA (Figure 4.47). It is therefore hypothesised that the variation of amino acid would not impact the binding activity or structure of ppxA/L to a significant degree. The variation between the sequences is likely due to minor genomic sequence variation which have been found in both human and bacterial genomes (Hirschhorn *et al,* 2002, Weissman *et al.* 2003). ppxB/L was also cloned into the pQE30 plasmid. Additional cloning steps were carried out to introduce a linker and for the insertion of a nucleotide base to facilitate correct expression and purification. The pQE30 vector facilitates the translation of a fusion protein having an N-terminal 6HIS tag. The lectin region within ppxA can be viewed as a 12.39 kDa protein with a 50.84 kDa protein fused to the N-terminus. The lectin region within ppxB can be viewed as a 12.24 kDa protein with a 25.42 kDa protein fused to the N-terminus. Therefore the fusion of 6HIS to the N-terminus of ppxA/L and ppxB/L was

not expected to affect the binding activity of the lectins. Despite some variation on the amino acid level the optimised conditions for the expression of PA-IIL in the *E. coli* expression system were utilised for the expression of the lectin regions of both ppxA and ppxB.

The pQE30.ppxA/L plasmid was transformed into the *E. coli* strain, XL10-Gold, the optimal expression strain identified for PA-IIL. ppxA/L was found to be expressed using this strain under the optimised conditions found for PA-IIL. An IMAC purification strategy was investigated to maximise the recovery of recombinant protein from the soluble cell extract while maintaining a high level of purity. ppxA/L, like PA-IIL, was found to stay bound to the column in the presence of 60mM imidazole and final protein was eluted with 350mM imidazole. The high level of purity obtained by using IMAC can be seen by the highly sensitive silver stain SDS-PAGE gel, where no impurities were detected (Figure 4.24). Some multimer formation does appear as a larger upper band on this gel. Although ppxA/L could be expressed and purified to a high level the yield of protein recovered was significantly lower than yields obtained for a similar expression of PA-IIL. To improve the protein yield the pQE30.ppxA/L plasmid was transformed to the *E. c*oli strain KRX. KRX is a protease deficient *E. coli* strain with expression levels reported to be as high as levels found for BL21 (DE3). The use of *E. coli* KRX did significantly improve expression of soluble ppxA/L (Figure 4.27, Lane 3). Attempts to upscale ppxA/L expression to 500 ml resulted in the loss of soluble protein (Figure 4.27, Lane 4). Expression of high yields of PA-IIL was found to result in a loss of soluble protein, attributed to aggregate formation. ppxA/L, like PA-IIL was found to inherently precipitate when purified. It was therefore hypothesised that the large scale expression of ppxA/L (from 500 ml) lead to high yields of protein production resulting in aggregate formation and loss of soluble protein.

The pQE30.ppxA/L plasmid was also transformed into the *E. coli* strain Rosetta. The frequencies with which different codons appear in genes in *E. coli* are different from those derived from other organisms. The frequency of specific codons is also reflected in the amount of specific tRNA, i.e. tRNA which recognizes rarely used codons is present in low amounts. Therefore genes that contain codons which are rare in *E. coli* may be inefficiently expressed. Amino acid sequence analysis of ppxA/L found the gene to be composed of 6% rare codons which may effect protein translation. The *E. coli* strain Rosetta carries a plasmid which facilitates rare codon expression. The use of Rosetta strain *E. coli* did not significantly improve expression of ppxA/L (Figure 4.29).

The poor protein expression was therefore not attributed to the presence of rare codons within ppxA/L. Further optimisation of expression parameters was therefore required.

The effect of induction time on protein expression was investigated at various points from OD_{600} 0.6 to 1.2. The induction of cultures after an OD_{600} of 0.6 did not significantly increase the yield of ppxA/L expressed. Subsequently 100ml *E. coli* KRX expression cultures of ppxA/L were pooled, typically two pooled, for purification to recover suitable protein yields for future work. The high level of purity obtained by using IMAC purification can be seen by the highly sensitive silver stain SDS-PAGE gel, Figure 4.31. ELLA found ppxA/L to be active.

The pQE30.ppxB/L plasmid was transformed into the *E. coli* strain, KRX, the optimal expression strain identified for ppxA/L. ppxB/L was found to be expressed using this strain and under the optimised PA-IIL conditions. When an IMAC purification strategy was investigated to maximise the recovery of recombinant protein from the soluble cell extract the protein did not bind to the Ni-NTA resin. The structure of Photopexin B is as yet unknown. The 6HIS tag may be unavailable due to steric hindrance of the protein folding. Addition of a linker may allow the 6HIS tag to become more available for binding. A linker was cloned into the pQE30.ppxB/L plasmid following the 6HIS tag and prior to the *ppxB/L* gene. The addition of a linker was found to improve binding of ppxB/L to the Ni-NTA resin so that protein could then be purified. The IMAC purification strategy established for ppxA/L and PA-IIL was used to purify ppxB/L. However, upon purification the molecular weight of the purified protein was found to be smaller than predicted. SDS-PAGE gel analysis shows the protein running between 6.5 kDA and 14 kDA marker. The predicted molecular weight of 6HIS N-tagged ppxB/L was calculated to be 14.2 kDa. Sequence confirmation of the clone found a nucleotide base to be missing leading to a frame shift and the premature presence of a stop codon. The missing nucleotide base is absent from a string of five A bases. Studies have revealed that nucleotide repeats generally contract during PCR because of the loss of repeat units (Clarke *et al.* 2001). The missing nucleotide was therefore attributed to an error during PCR. Primers were designed to correct the loss of the nucleotide and the resulting clone, pQE30.ppxB/L.linkfix, was confirmed by DNA sequencing. Protein expressed and purified over Ni-NTA resin from the pQE30.ppxB/L.linkfix plasmid was found to be, approximately, the correct molecular weight namely 14.19 kDa predicted from the amino acid sequence. However the yield of protein recovered from a 100 ml cell extract was found to be relatively low. Protein

was not found to be present in the insoluble fraction indicating that poor protein expression was not due to aggregation of protein in the insoluble fraction but due to low protein expression. The yield of expressed ppxB/L was insufficient to carry out further characterisation studies of the lectin. To improve the protein yield of ppxB/L further optimisation such as clonal selection is required. There was not sufficient time during the course of this study to carry out this work.

The relative molecular mass and possible quaternary structure of ppxA/L under native conditions was investigated by size exclusion chromatography using two columns with differing matrices. Initial size exclusion chromatography was carried out using the HW-55S Toyoperal matrix. Noticeable on the HW-55S traces of ppxA/L is the presence of two elution peaks, one of which lies outside the size range of the standard curve. The second peak was not attributed to contaminants as silver stained gels show ppxA/L to be relatively highly purified. Nor was it attributed to a lectin-carbohydrate interaction as ppxA/L was found to elute similarly from the column in the presence and absence of mannose. Toyopearl HW55S resin was deemed unsuitable for use with ppxA/L as presence of two peaks in the elution profile indicated the protein was interacting with the resin. An alternative matrix was used to investigate the native structure of ppxA/L. A gel filtration study using the Superdex 75 column was carried out. In PBS ppxA/L was found as a monomer (Table 4.9). The calculated molecular mass of ppxA/L (MW of ~ 13.56kDa) was consistently smaller than the molecular mass predicted from the amino acid sequence (MW of ~13.95kDa). It is theorized that as the standard curve was created using globular proteins it serves as only a rough estimate as to the size of non-globular shaped proteins such as ppxA/L. This feature was also found for the PA-IIL run over a Sepadex G-100 column. Of note is the formation of ppxA/L multimers in the presence of imidazole. This was also noted in the SDS-PAGE gel analysis (Figure 4.24). The fusion of a 50.84 KDa protein to the N-terminus of the lectin domain of Photopexin A in its native state suggests that no such multimer formation would be possible in nature. It may be theorized that the formation of multimers is an evolutionary relic of the protein as other members of the PA-IIL family, e.g. CV-IIL and RS-IIL (Sudakevitz *et al.* 2004), as well as PA-IIL have a native tetrameric structure (Mitchell *et al.* 2002).

5.0 Determination of carbohydrate binding protein specificity using the Enzyme linked lectin assay (ELLA)

5.1 Overview

The PA-IIL and ppxA/L lectins are described in Chapter 3 and 4. Characterisation of PA-IIL and ppxA/L with respect to glycan binding specificity was performed using a series of specificity analysis experiments. Traditionally, lectins are characterised using a hemagglutination inhibition assay, however the ELLA (Section 2.27) is proving a more useful method for glycobiologists as it provides more detailed information about both lectin specificity and affinity. Preliminary attempts to carry out the hemagglutination assay with PA-IIL failed to agglutinate rat erythrocytes. Further work was not carried out on this assay due to the lack of access to alternative erythrocytes and the unsuitability of the assay for the determination of the specificity of monomeric lectins, such as ppxA/L.

5.2 Enzyme Linked Lectin Assay

The ELLA is a modified ELISA in which the primary antibody has been replaced by a tagged lectin and the protein target is a glycoprotein. The ELLA technique is a simple microarray that has been proven to be suitable for the detection of specific carbohydrate units, either directly on immobilised glycoproteins or indirectly by inhibition with soluble sugars (McCoy, Varani and Goldstein 1983). In this study, the binding specificity of the lectins was determined by the ability of a lectin to bind to a set of known glycoproteins. The 6HIS affinity tag which was used to purify the protein was utilised in the ELLA as a target for a secondary antibody. Alternative purification tags, such as StrepII, may also act as targets for antibodies. The principle of a typical ELLA is outlined in Figure 5.1.

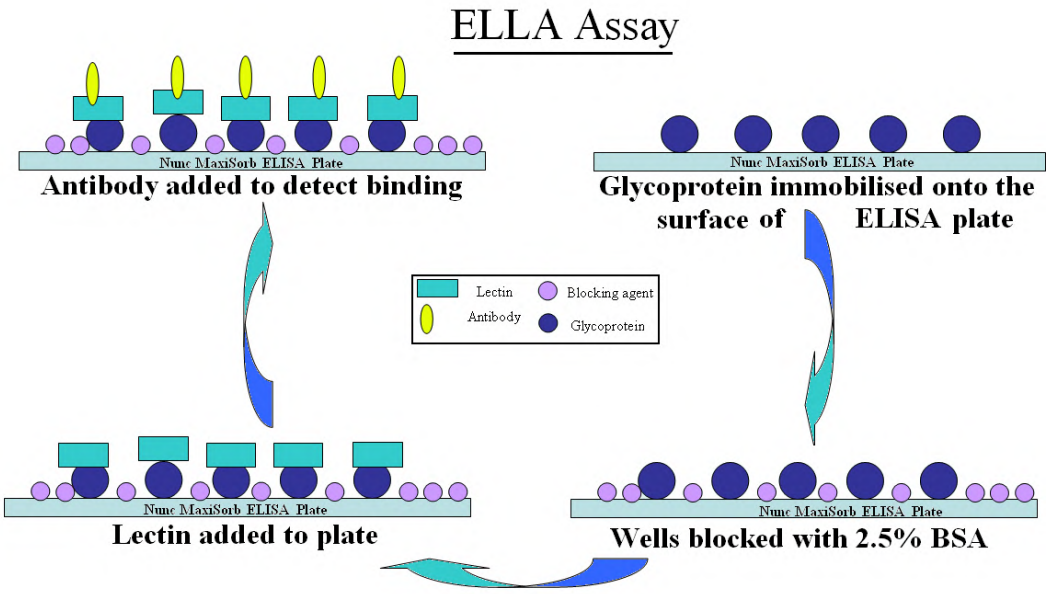


Figure 5.1: Schematic outlining steps involved during ELLA.
HRP-bound antibody facilitates the detection of binding. The addition of the HRP substrate, tetramethyl benzidine (TMB), to the HRP-bound antibody results in a colour change which can be measured.

Background signals and false positives may occur in ELLAs and are due to the non-specific interaction of reagents with each other, and also to their adherence to residual, non-saturated sites on the solid surface of the microwells. Addition of blocking reagents offers an efficient strategy for minimizing the non-specific background binding. After coating with a specific ligand the blocking agent is added to block any excess solid surface to avoid unspecific immobilisation of succeeding reagents. A typical blocking agent would be a neutral or inert macromolecule, large enough to establish a stable attachment to the surface, yet small enough to find its way between the previously bound ligand. Prior to investigating lectin interaction with known glycoproteins a number of different blocking agents were investigated to minimize signal-to-noise ratio and enable accurate detection of gylcans using lectins.

5.2.1 Bovine serum albumin (BSA) as a blocking agent

BSA is a un-glycosylated protein, commonly used as a blocking agent which has been shown to be an efficient reagent for saturating the surface of Nunc microwells (Steinitz 2000). However, BSA is purified from a complex protein source and therefore may contain interfering glycoproteins, such as globulins, causing false positives (Afrough, Dwek and Greenwell 2007). When microwells were saturated with 2.5% BSA and

probed with 10µg/ml PA-IIL as outlined in Section 2.27, PA-IIL was found to bind to BSA with a purity of ≥ 98%, regardless of supplier, (Figure 5.2).

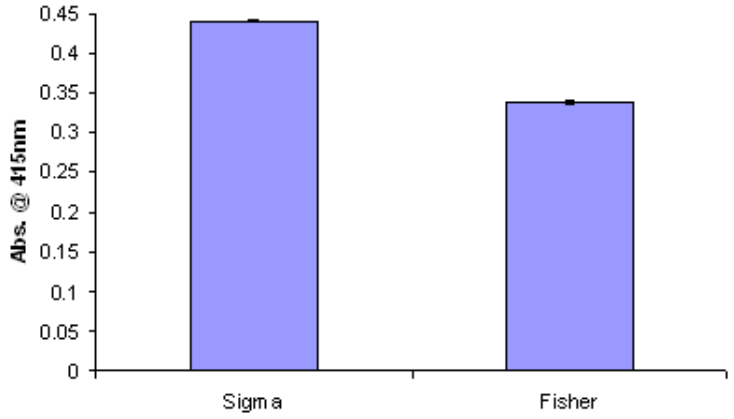


Figure 5.2: Comparison of recombinant PA-IIL binding to BSA solutions from differing suplliers
Comparison of the binding activity of PA-IIL to BSA from two manufactures, Sigma-Aldrich and Fisher, both Fraction V purifications of a purity of ≥ 98%. A control, lacking the addition of lectin, was run in tandem to the BSA. Assayed by the activity of a HRP-linked secondary antibody (OD_{415nm}). (Error bars represent mean ± SD, n = 3).

There was some variation in the binding signal between the two BSA batches. This was attributed to a greater efficiency of the purification treatment in one company over the other although this may also vary batch to batch. The binding signal of both BSA batches was deemed too high for use in an ELLA as the lectin-glycoprotein binding signal would be obscured by the lectin-blocking agent binding signal.

The sugar specificity of PA-IIL was reported to be L-Fuc > D-Man, L-Gal> D-Fru (Gilboa-Garber 1982). The binding of PA-IIL and ppxA/L to a BSA solution was investigated through the addition of soluble sugars, glycoproteins and EDTA during the ELLA. Invertase is a yeast glycoprotein glycosylated with a carbohydrate content of 50% (w/w), composed primarily of high mannose structures. Glucose oxidase is a glycoprotein glycosylated with a carbohydrate content of 16% (w/w) of which 80% (w/w) is composed of high mannose structures. PA-IIL binding has been shown to be dependent on the presence of Ca^{2+} (Mitchell *et al.* 2002). ppxA/L binding is hypothesised to be dependent on the presence of Ca^{2+} due to the high conservation of the calcium binding loop in ppxA/L. Stripping the lectins of Ca^{2+} using EDTA therefore results in loss of the carbohydrate specific binding of the lectins.

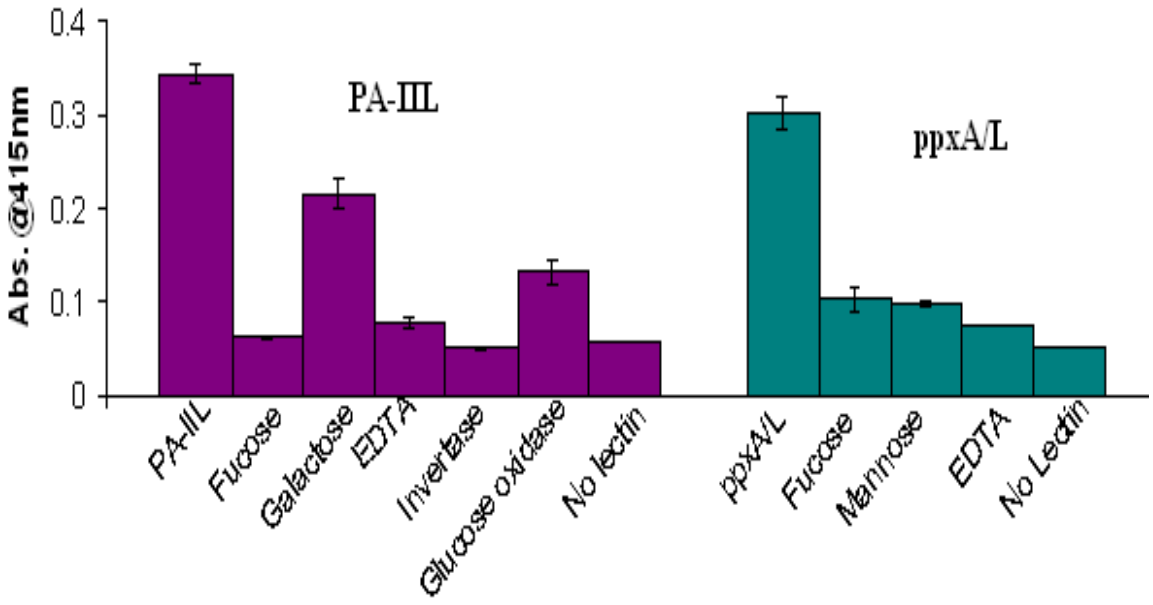


Figure 5.3: Determination of recombinant PA-IIL and ppxA/L binding to 2.5% BSA
Quantitative detection of lectin bound to BSA assayed by the activity of a HRP-linked secondary antibody (OD_{415nm}). The effect of the removal of the calcium ions from the lectin binding pocket by EDTA on lectin binding to BSA was evaluated. Loss of signal after treatment with EDTA indicates binding is sugar specific. The inhibition of PA-IIL binding signal to BSA by the addition of 0.3M fucose and galactose and 2mg/ml of invertase and glucose oxidase confirms that binding activity is sugar specific. The inhibition of ppxA/L binding to BSA by the addition of 0.5M fucose and mannose confirms that binding activity is sugar specific. (Error bars represent mean ± SD, n = 3).

If the binding signal of PA-IIL or ppxA/L to BSA was due to non-specific protein-protein interaction, treatment of the lectin with EDTA would not be expected to quench the binding signal as was observed (Figure 5.3). The degree of quenching of the binding signal of PA-IIL as shown in Figure 5.3, by the addition of soluble sugars and glycoproteins correlates with the sugar specificity determined by Gilboa-Garber, N. (1983). Due to the higher mannose structure content of invertase, it would be expected to quench the carbohydrate specific binding signal of PA-IIL more efficiently the glucose oxidase. Sugar specificity of ppxA/L is unknown, however the addition of both fucose and mannose significantly quenched the binding signal. Therefore PA-IIL and ppxA/L are binding to the BSA solution, or a contaminant carbohydrate within the BSA solution, due to sugar specific binding. In order to eliminate the problem caused by carbohydrate contamination a BSA solution was treated with sodium periodate prior to use (Section 2.21). Sodium periodate oxidises carbohydrate moties thereby preventing lectin binding.

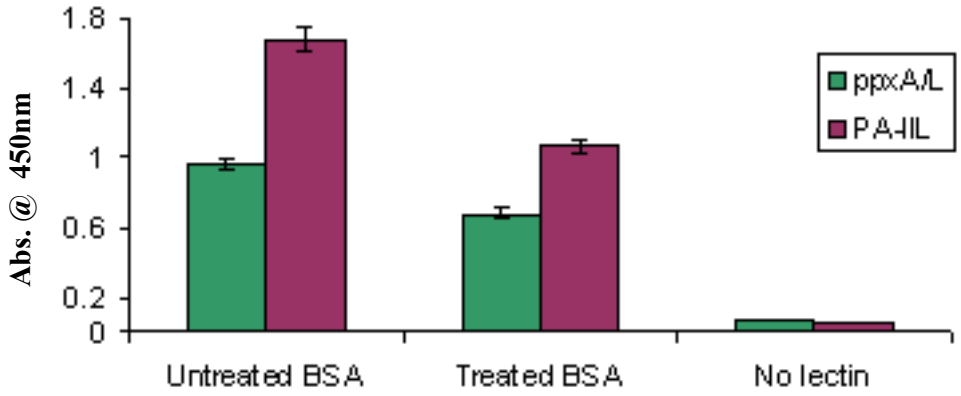


Figure 5.4: Comparison of the activity of recombinant PA-IIL and ppxA/L with untreated and sodium periodate treated 2.5% BSA in ELLA
Quantitative detection of lectin bound to untreated and sodium periodate treated 2.5% BSA assayed by the activity of a HRP-linked secondary antibody (OD_{450nm}). The sodium periodate treatment on lectin binding signal was compared. Lectin binding was reduced on treated BSA. (Error bars represent mean ± SD, n = 3).

Treatment of BSA with sodium periodate did reduce the binding signal obtained for PA-IIL-BSA and ppxA/L-BSA by 30% - 50% (Figure 5.4). However, halving of the binding signal did not cause a sufficient reduction to allow for the sensitive and accurate detection of specific protein-carbohydrate interactions.

The specific interaction of PA-IIL and ppxA/L with BSA, which was reduced but not eliminated by periodate treatment, makes BSA an unsuitable blocking reagent for analysis of these lectins.

5.2.2 Non-protein blocking agents

There are advantages to using non-protein blocking agents as they can replace complex mixture blocking agents, such as BSA, with a single pure compound. The use of pure compounds over complex mixtures may make an important contribution to standardization and since they are not purified from a protein source they are not contaminated by interfering proteins. Two such commonly used non-protein blocking agents are polyvinylpyrrolidone (PVP) and polyvinyl alcohol (PVA). PVP was used as a suitable blocking agent for immunochemical studies by Rodda, D.J. *et al.* (1994). PVA was used as a non-protein blocking agent on nitrocellulose membranes (Harvey 1991). The optimal conditions for the use of PVA in enzyme immunoassays (EIA) were established by Haycock, J.W. (1993). This study established PVA to be four fold better at reducing non-specific binding than PVP. It was theorized that the structure of PVA

allows for efficient binding to the hydrophobic polystyrene while more effectively exposing its hydrophilic substituents compared to PVP (Haycock 1993). Two commercial ELISA buffers were also examined as potential blocking agents; SYNBlock ELISA blocking buffer (AbD seroTec) and Carbo-Free ELISA blocking buffer (Vector Labs). When wells were saturated by each of the blocking agents and probed with 10µg/ml of PA-IIL as outlined in Section 2.27, PA-IIL was found to interact with all of the blocking agents to a high degree. Quenching of the binding signal of PA-IIL by the addition of fucose indicated that the binding signal for each of the non-protein blocking agents was due to sugar-specific binding (Figure 5.5).

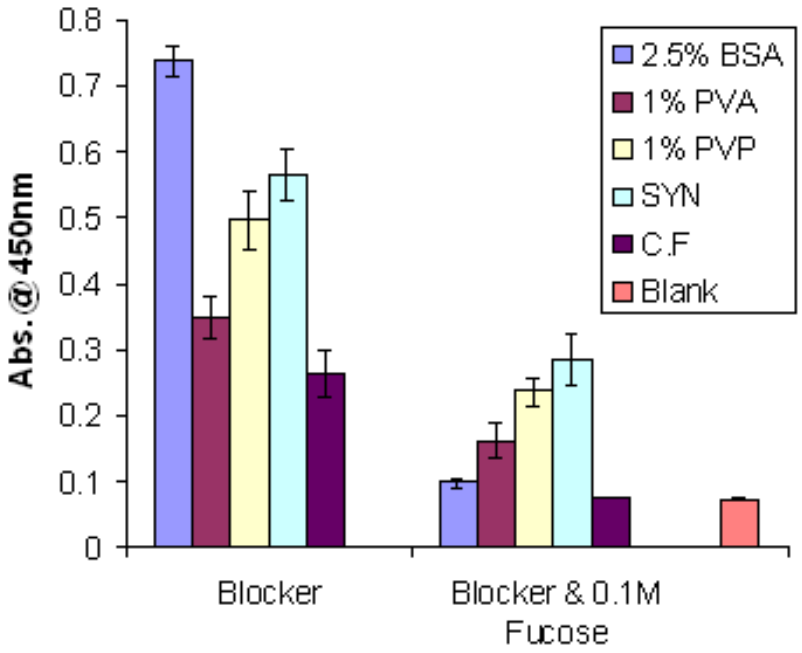


Figure 5.5: Comparison of blocking reagents in ELLA with recombinant PA-IIL
Quantitative detection of lectin bound to a number of blocking agents assayed by the activity of a HRP-linked secondary antibody (OD_{450nm}). The blocking agents investigated were 1% polyvinylpyrrolidone (PVP), 1% polyvinyl alcohol (PVA), 2.5% BSA (Sigma-Aldrich), SYNBlock ELISA blocking buffer (SYN, AbD seroTec) and Carbo-Free ELISA blocking buffer (C.F, Vector Labs). (Error bars represent mean ± SD, n = 3).

A number of the non-protein blocking agents were also probed with 10 µg/ml ppxA/L to establish if ppxA/L behaves in the same manner as PA-IIL (Figure5.6).

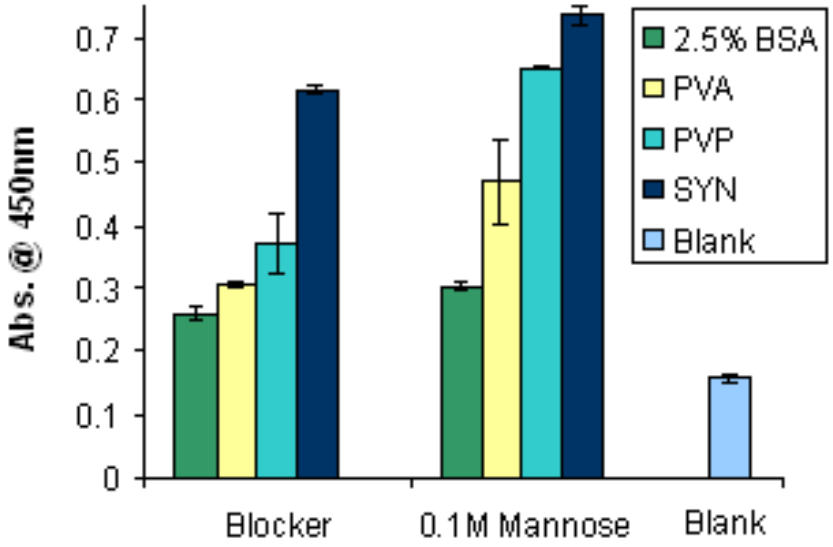


Figure 5.6: Comparison of blocking reagents in ELLA with recombinant ppxA/L
Quantitative detection of lectin bound to a number of blocking agents assayed by the activity of a HRP-linked secondary antibody (OD_{450nm}). The blocking agents investigated were 1% polyvinylpyrrolidone (PVP), 1% polyvinyl alcohol (PVA), 2.5% BSA (Sigma-Aldrich) and SYNBlock ELISA blocking buffer (SYN, AbD seroTec). (Error bars represent mean ± SD, n = 3).

The addition of mannose does not quench the binding signal of ppxA/L to any of the non-protein blocking agents implying that the binding is due to non-specific binding. However, the binding signal of ppxA/L to BSA, which was previously found to be quenched by the addition of sugar (Figure 5.3), was not quenched during this experiment. During the course of this work lyophilization was regularly utilised for long term storage of proteins. However over a period of time the lyophilizer was found to operate poorly and ultimately break down. When the lyophiliser was operating poorly it is believed to have resulted in heating of the protein when drying. The heating of the protein could lead to denaturation. This could result in unfolding of the protein and exposure of hydrophobic and hydrophilic sites leading to non-specific binding as observed for ppxA/L-BSA binding in Figure 5.6. Heating of PA-IIL would not result in the same degree of denaturation as PA-IIL has been shown to be a robust protein, tolerating temperatures up to 80^{o}C (Gilboa-Garber 1982). The activity of ppxA/L was further investigated using the alternative ELLA format, the HRP assay (see Section 2.26). In this format the lectin is bound to surface of an ELISA plate, HRP was then added to determine if the lectin could bind to this mannose bearing compound. If HRP can be bound by sugar-specific activity the addition of soluble sugars resulting in the quenching of binding signal would be indicative of lectin sugar preference. The

principle of a typical HRP assay is outlined in Figure 5.7, however since ppxA/L was previously found to bind to BSA no blocking was carried out. The lack of a blocking agent may result in a higher signal-to noise ratio but in this instance a qualitative result will suffice, i.e. the confirmation of the presence of non-specific binding, over a quantitative result so a higher signal-to-noise ratio can be tolerated.

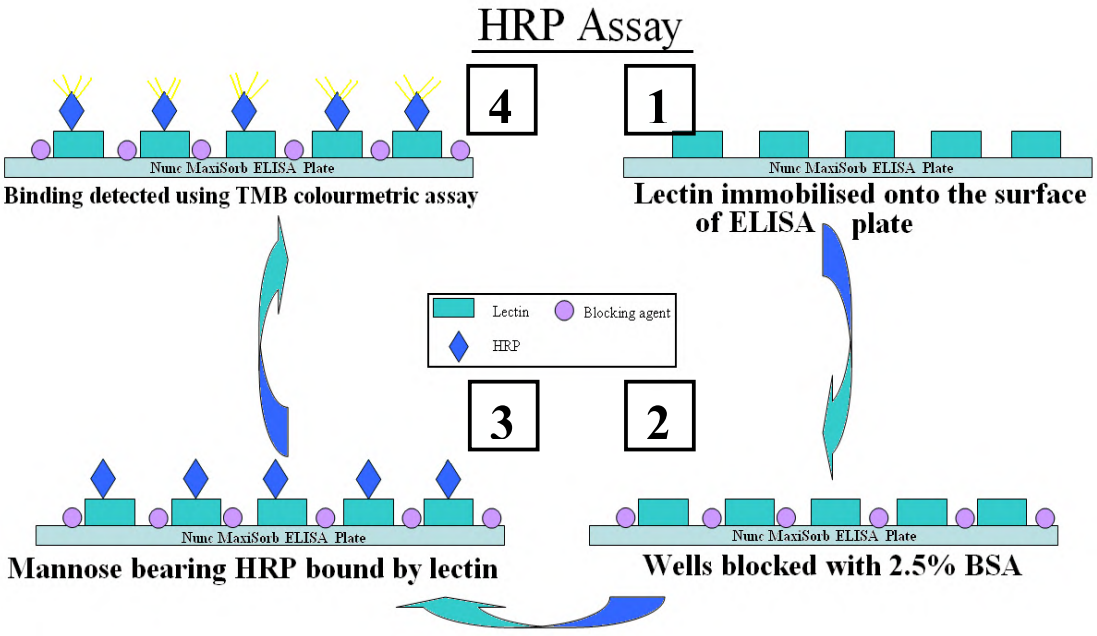


Figure 5.7: Schematic outlining steps involved during HRP assay

Since the binding specificity of ppxA/L is unknown the presence of non-specific binding was confirmed by attempting inhibition with a compound previously demonstrated to be bound by ppxA/L i.e. BSA (Figure 5.8).

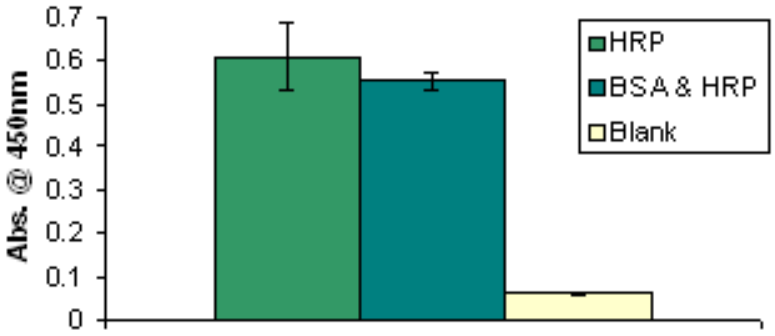


Figure 5.8: Inhibition of ppxA/L binding to HRP by 2.5% BSA
Quantitative detection of ppxA/L bound to a constant concentration of HRP in the presence of 2.5% BSA. The constant concentration of HRP was 10µg/ml. The addition of BSA did not quench the binding signal indicating that binding is due to non-specific interaction. (Error bars represent mean ± SD, n = 3).

The BSA solution preparation failed to quench the ppxA/L-HRP signal (see Figure 5.8) and non-specific binding was further confirmed by attempts to inhibit the interaction of ppxA/L to HRP by soluble sugars or glycoproteins (Figure 5.9 and 5.10). The binding signal failed to be quenched in all cases.

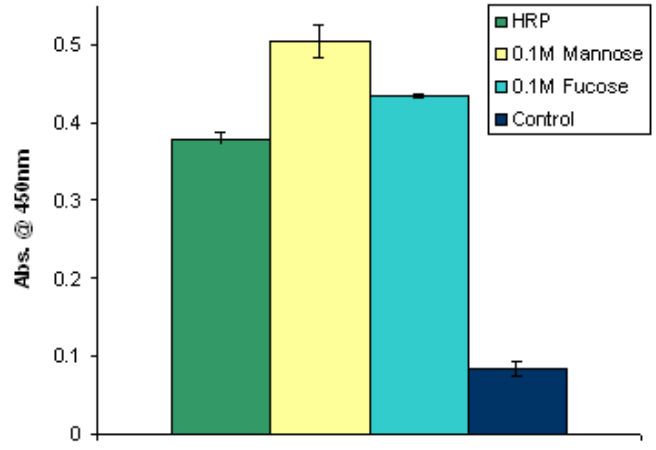


Figure 5.9: Inhibition of ppxA/L binding to HRP by soluble sugars
Quantitative detection of ppxAL bound to a constant concentration of HRP in the presence of various sugars. The constant concentration of HRP used was 10µg/ml. The addition of soluble sugars did not quench the binding signal confirming that binding is due to non-specific interaction. (Error bars represent mean ± SD, n = 3).

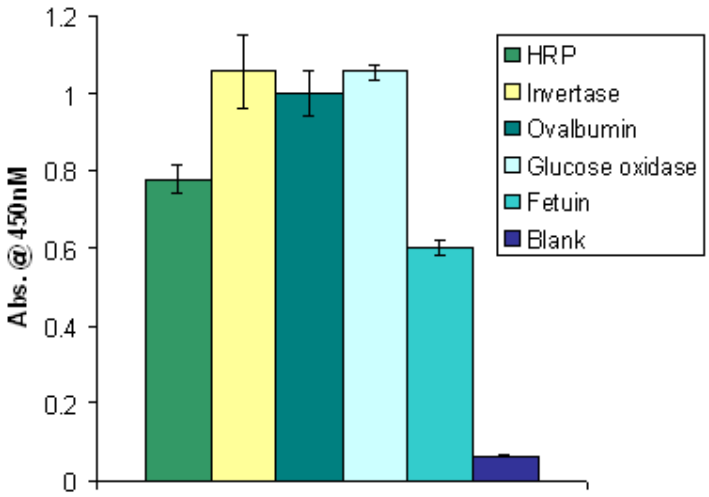


Figure 5.10: Inhibition of ppxA/L binding to HRP by glycoproteins
Quantitative detection of ppxAL bound to a constant concentration of HRP and inhibition with a constant concentration of various glycoproteins. The constant concentration of HRP used was 10µg/ml, while the constant concentration of glycoprotein was 1mg/ml. The addition of glycoprotein did not quench the binding signal confirming that binding is due to non-specific interaction. (Error bars represent mean ± SD, n = 3).

Due to the instability of ppxA/L it was necessary to lyophilize the protein directly after production. The loss of a facility to lyophilise the protein meant there was no means of storing the protein and therefore no further testing on blocking agents or binding to glycoproteins could be carried out.

5.3 Discussion

Of the common ELISA blocking agents investigated, (BSA, PVA, PVP, SYNBlock, Carbo-Free) none were found to be suitable for use in an ELLA with PA-IIL and ppxA/L binding. All of these blocking agents resulted in high signal-to-noise ratio, leading to false positives and loss of the ability to discern subtle signal variation. No definitive comment can be made on the binding specificity of ppxA/L due to the complications introduced by faulty equipment. Without the ability to lyophilize ppxA/L the protein could not be suitably stored and therefore extensive studies of binding affinity could not be carried out. Preliminary results indicated that ppxA/L bound to BSA, and may therefore exhibit a similar binding pattern to blocking reagents as found for PA-IIL. Although other blocking agents may be investigated, there are a number of alternative approaches to investigation of lectin binding affinities known that may be preferential to ELLA.

Calorimetry has evolved over several hundred years, giving rise to modern techniques. Development of microelectronics and advancements in instrument design has radically improved the sensitivity of calorimetry techniques. Isothermal titration calorimetry (ITC) is one such technique employed to investigate the thermodynamics of ligand macromolecule interaction. A typical ITC experiment is carried out by the stepwise addition of one of the reactants, e.g. ligand, into the reaction cell of the calorimeter containing the other reactant, e.g. protein (Figure 5.11).

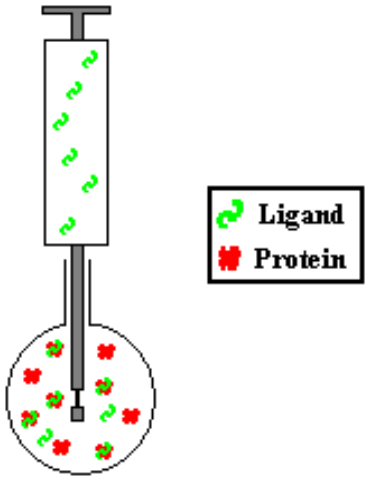


Figure 5.11: Illustration of the configuration of an isothermal titration calorimetry (ITC) reaction cell
The cell is filled with protein (red), while the injection syringe which also stirs the solution to insure proper mixing is filled with a ligand solution (green). At specific time intervals a small volume of the ligand solution in injected into the cell, resulting in the characteristic titration heat effects.

When substances, such as proteins, bind heat is generated or absorbed proportional to the amount of ligand that binds to the protein. Measuring heat fluctuations allows accurate determination of binding constants, reaction stoichiometry, enthalpy and entropy (Leavitt, Freire 2001). Using ITC the complete thermodynamic profile of molecule interaction can be determined in a single experiment. A thorough review by Dam, T.K. *et al.* (2002), outlines numerous studies of plant and animal lectins that utilised ITC for the determination of thermodynamic properties such as binding affinity. Perret, S. *et al.* (2005) used ITC analysis to examine the binding of PA-IIL to human milk oligosaccharides. The ITC values obtained for ligand affinity were found to be in agreement with the specificity range established by ELLA. This result indicates that ITC is a suitable alternative to the use of ELLA for carbohydrate binding determination. The effect of point mutation of the amino acids found within the sugar binding loop of

PA-IIL was also investigated using ITC (Adam *et al.* 2007). The point mutations were shown to alter the binding affinity and sugar preference of the lectin (see Section 1.6). ITC does not require the use of a blocking agent making it an attractive option for the analysis of PA-IIL and homologs to PA-IIL. The primary disadvantage of ITC is that protein activity obtained in solution. ELLA results indicate protein activity on a solid surface. For diagnostic and industrial application lectins will most likely be used on a solid surface. ITC could facilitate initial characterisation of binding specificity and binding mode of a lectin which may potentially aid in the identification of a suitable blocking agent for optimisation of a sensitive ELLA.

Surface plasmon resonance (SPR) is an optical phenomenon that provides a non-invasive, label-free means of monitoring binding interactions between an injected analyte, such as a glycoprotein, and an immobilized biomolecule, such as a lectin, in real time. Although there are several SPR-based systems (Schuck 1997, van der Merwe, Barclayt 1997) the most widely used system is BIAcore, GE HealthCare (Raghavan, Bjorkman 1995). The BIAcore system contains a sensor microchip, a laser light source emitting polarized light, an automated fluid-handling system and a diode-array position-sensitive detector as shown in Figure 5.12. The BIAcore chip consists of a glass support, an overlaid gold film as SPR is based on the peculiar properties of the interaction of polarized light with thin metal films of aluminium, silver or gold films (Turbadar 1959) and a matrix to which biomolecules can be coupled. Proteins can be immobilized to various kinds of surface via diffusion, adsorption/absorption and covalent cross-linking. Attachment of proteins via any of these methods leads to attachment of proteins to the surface in a random fashion which can lead to alteration of native protein confirmation, reduction in activity and make binding sites inaccessible to probes. Attachment via affinity tags, such as the 6HIS tag utilised in this work, facilitates uniform attachment to the surface. The basis of SPR detection is;

- Lectin immobilised on a sensor chip in a layer in contact with a thin metal film, typically gold
- Glycoprotein in solution is passed over the layer
- Glycoprotein is retained by interacting with the lectin leading to an increase in organic matter leading to a proportional increase in refractive index which is observed as a shift in the resonance angle.

The layer close to the metal film is monitored in real time facilitating determination of both the rate and the equilibrium binding constants.

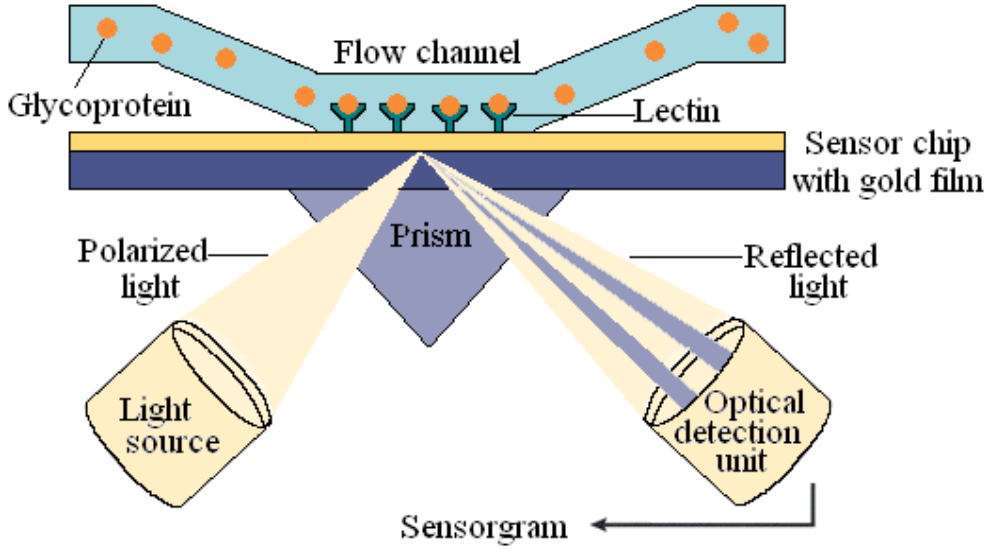


Figure 5.12: Schematic of the Surface Plasmon Resonance (SPR) system found in the BIAcore system

SPR has previously been demonstrated as a valuable tool for the characterisation of lectins and in particular the binding domain (Vornholt, Hartmann and Keusgen 2007). Shinohara, Y. *et al.* (1997), demonstrated that the affinity constants calculated from SPR experiments with immobilized lectin were in good agreement with those obtained by using classical methods. SPR has some advantage over ITC as data on translation states can be obtained. This information can be used for the characterisation of the mechanism of complex formation. ITC provides no direct information on the translational states of protein binding. SPR also consumes less reagents, however SPR assays are significantly longer then those for ITC experiments, involving, for example optimisation of immobilization conditions which can also consume significant amounts of reagents (Wear, Walkinshaw 2006).

Frontal affinity chromatography (FAC) can be used to characterise the binding of lectins. The original principle of FAC was derived by Kasai, K. *et al* (1975). Advances in technology have since lead to automated FAC systems (Nakamura-Tsuruta, Uchiyama and Hirabayashi 2006). A number of systems enabling sensitive monitoring of eluted glycan have been teamed with FAC; radioisotope (RI) (Ohyama *et al.* 1985), mass spectrometry (MS) (Schriemer *et al.* 1999) and fluorescence detection (FD). RI requires radio labelling of glycan which can pose a hazard to human health and

MS requires modification of glycans with appropriate alykyl reagent necessary to increase ionization efficiency in MS. FAC-FD with pyridylaminated (PA)-glycans offers the greatest ease of use. The principle of this technique is (outline in Figure 5.13);

- Diluted fluorescently labelled glycan solution is continuously applied to a lectin-immobilized column
- If the glycan has no affinity (blue) its elution front (V_0) is detected immediately
- If the glycan has affinity (red) with the immobilised lectin its elution front (V) is retarded

The retardaration is due to repeated interaction with the immobilized lectin as it passes through the column. Excess glycan is passed over the column until the glycan concentration reaches a 'plateau' that corresponds to the initial concentration [A_0].

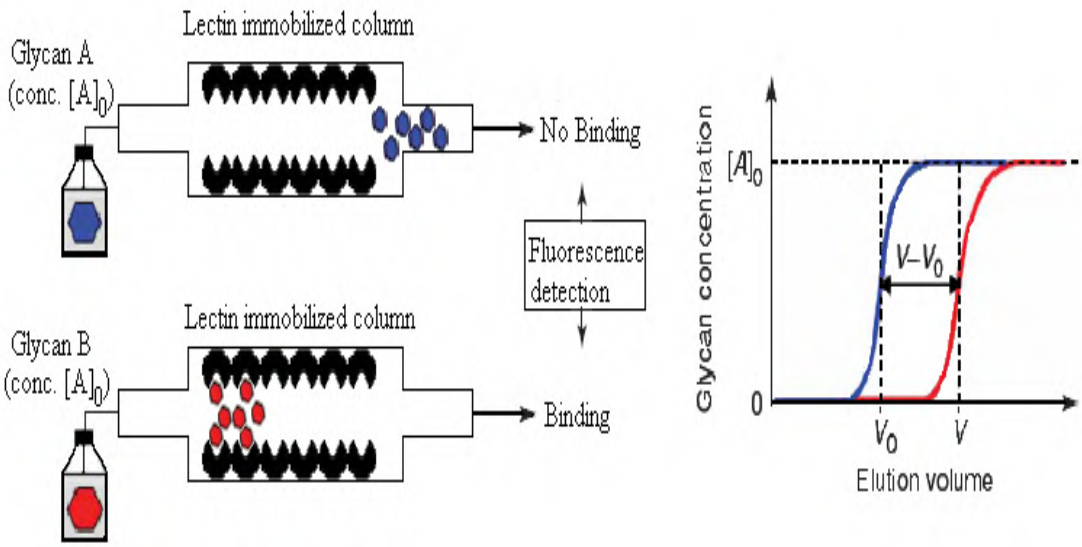


Figure 5.13: Schematic of Frontal Affinity Chromatography (FAC) – Fluorescence detection (FD) system.
[A_0] is the initial glycan concentration expressed in M.

Proteins of differing dissociation constants pass through the affinity column at different rates.

The FAC equation for the determination of the dissociation constant;

$$K_d = \frac{B_t}{(V - V_0)} - [A]_0$$

Equation 5.1: Basic equation of FAC to determine the dissociation constant (K_d) of lectins
Where B_t is an effective ligand content (expressed in mol) of the lectin immobilised column, V and V_0 are elution front volumes of analyte and a control substance respectively, and $[A_0]$ is the initial concentration of analyte.

The advantages of FAC-FD system include (Hirabayash 2008);

- High throughput
- Accurate and reproducible
- Simplicity of apparatus

The disadvantages of FAC-FD system include (Hirabayash 2008);

- Immobilization of lectin to matrix might result in modification (reduction) of binding
- Relative large amount of lectin required even on minature column
- Crude samples cannot be analysed
- Determination of B_t, only a small repertoire of saccharide derivatives have a UV-sensitive group.

An alternative format of the typically used FAC-FD, whereby a carbohydrate is immobilized on a column and a fluorescently tagged lectin passed over the column, could also be potentially used to determine the binding specificity of the lectins. If the lectin has no affinity to the immobilized carbohydrate its elution front would be detected immediately. If the lectin has affinity with the immobilised carbohydrate its elution front would be retarded. Since lectins can easily be recombinantly tagged by GFP (Waldo *et al.* 2004) this format is not limited to the small repertoire of saccharide derivatives with a UV-sensitive group that traditional FAC-FD is limited to. A potential disadvantage of this format is that it consumes more reagent, i.e. lectin, however the development of a high yield recombinant expression system would overcome this problem and the addition of a fusion protein, such as GFP, may affect the activity of a lectin. The purification of untagged PA-IIL over immobilized carbohydrate column of mannose agarose (see Section 3.8.1.1, Figure 3.55) demonstrates the principle of this format.

Traditionally the hemagglutination assay was used to determine lectin specificity. The assay is based upon lectins ability to agglutinate red blood cells in a U-bottomed well as opposed to red blood cells falling out of suspension by gravity. Therefore the assay is dependant on the lectin having a high specificity for the sugars displayed on red blood cells. The sugars displayed on red blood cells are species specific. The minimum amount of lectin required for agglutination of a fixed quantity of red blood cells is termed one hemagglutination unit (HU). Once HU is established inhibition assays can be carried out. The higher the affinity of the sugar for the lectin, the lower the concentration required to inhibit the agglutination reaction.

PA-IIL was demonstrated to agglutinate rabbit, rat and sheep erythrocytes and did not react with cow erythrocytes (Zinger-Yosovich *et al.* 2006). Rat blood cells were the only erythrocytes available constantly for this work. Preliminary attempts to carry out hemagglutination assay of PA-IIL failed to agglutinate the erythrocytes. The hemagglutination assay was not carried out using ppxA/L as the lectin was found to be monomeric (see Section 4.6.2). Agglutination requires the presence of multiple binding sites in the lectin molecule.

6.0 The PA-IIL superfamily and future work

6.1 The growing PA-IIL superfamily

Searches of sequencing databases by bioinformatic analysis for proteins displaying similarities to the PA-IIL lectin of *P. aeruginosa* identified a number of homologous sequences to PA-IIL, such as the characterised RS-IIL from *Ralstonia Solanacearum* and CV-IIL from *Chromobacterium violaceum* (as discussed in Section 1.6) and uncharacterised homologous proteins from *Photorhabdus luminescens* and *Photorhabdus asymbiotica*. In this study two of the uncharacterised homologues were investigated, namely the lectin regions of Photopexin A and B. Additional uncharacterised homologues of the lectin PA-IIL in *P. luminescens* and *P. asymbiotica* were identified by a protein sequence query against the NCBI database using BLAST. The search reveals the two Photopexins and another 3 open reading frames in the unannoted genome of *Photorhabdus luminescens* and 2 open reading frames in the unannoted genome of *Photorhabdus asymbiotica* with significant sequence similarity (28%-33%) to PA-IIL. Nominally, functional annotation can be reliably transferred between proteins that share 30% sequence identity or above, confidence increasing with identity. These hypothetical genes were designated Plu4230, Plu4231 and Plu4238 from *P. luminescens* and P.asy1 and P.asy2 from *P. asymbiotica.*

6.1.1 PA-IIL homologues from *P. luminescens*

The uncharacterised hypothetical genes from *P. luminescens* would encode:

Plu4238 – 320 amino acid protein with a calculated MW of 34.8kDa

Plu4231 – 333 amino acid protein with a calculated MW of 36.77kDa

Plu4230 – 312 amino acid protein with a calculated MW of 34.35kDa

Multiple protein alignment reveals a C-terminal domain with strong homology to PA-IIL, in particular across the amino acids of the calcium binding loop (see Figure 6.1).

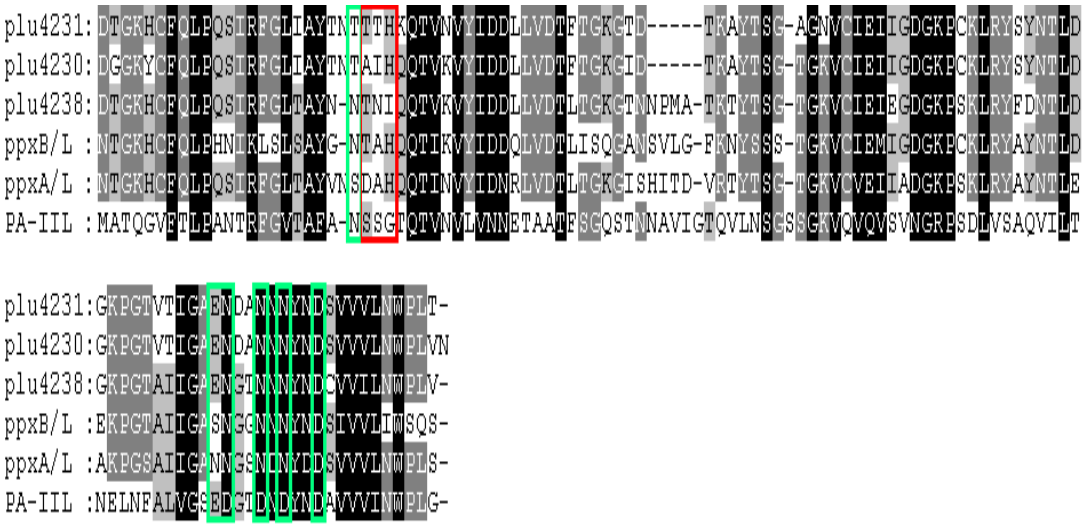


Figure 6.1: Sequence alignment of the lectin-like C-terminal domains of the homologues from *P. luminescens* TT01 and PA-IIL from *P. aeruginosa*.
Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

There is significant sequence similarity of the proteins within the *P. luminescens* strain (Figure 6.2), however within the sugar binding loop the amino acids are not as highly conserved between all the proteins (see Figure 6.1).

	Plu4231	Plu4230	Plu4238	ppxB/L	ppxA/L
PA-IIL	32%	33%	37%	26%	33%
	56%	57%	62%	56%	61%
	6	6	3	1	2
Plu4231	—	90%	75%	61%	67%
	—	91%	78%	71%	80%
	—	0	4	4	3

Figure 6.16: Homologous lectins; *P. aeruginosa* PA-IIL, lectin domains of *P. luminescens* plu4230, plu4231, plu4238, ppxA/L and ppxB/L. Percentage identity –black, percentage similarity – red, gap character – blue. Statistic report created using ClustalW and GenDoc.

The percent identity (Figure 6.2, black numbering) refers to the percent of amino acids in the alignment which are identical between the proteins. The percent similarity (Figure 6.2, red numbering) refers to the percent of amino acids in the alignment which are not identical between proteins but which have similar biochemical properties and are considered interchangeable in many situations. The gap character refers to padding or null characters to match homologous residues indicated by – in the alignment (see Figure 6.1). The alignment of PA-IIL to the homologous protein identified from *P.*

luminescens found lectin domains with exact identity of 26 to 37% and similarities of 56 to 62%. Reviews of comparative protein modelling have found that high-accuracy comparative models are based on more than 50% sequence identity to their templates. Medium-accuracy comparative models are based on 30 to 50% sequence identity. Low accuracy comparative models are based on less than 30% (Sanchez, Sali 1998). The identity of each of the Plu4230, Plu4231 and Plu4238 proteins were found to have an identity greater then 30% to PA-IIL (Figure 6.2). Therefore, a 3D structure model of each of the unknown proteins was created using the crystal structure of PA-IIL as a template. The amino acid sequence for each of the proteins was submitted to SWISS-MODEL web server (Section 2.14) which provides an automated protein homology-modelling service (Schwede *et al*. 2003). The 3D structures were predicted using the PA-IIL PDB file 1OUX as a template. The binding pocket from each of the 3D-models was compared to the binding pocket of PA-IIL (Figure 6.3).

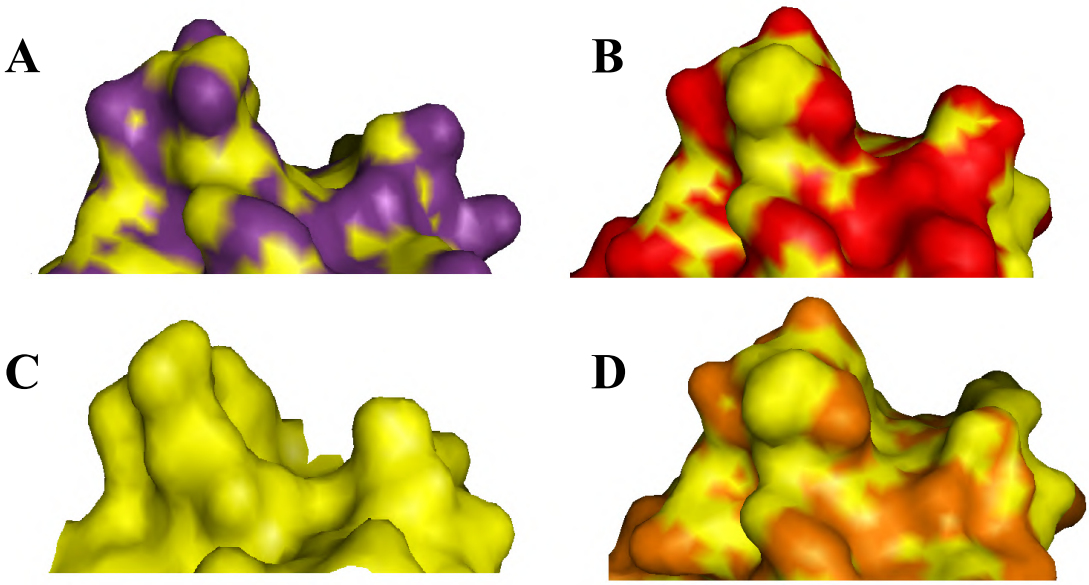


Figure 6.3: Plu4230 (orange), Plu4231 (red) and Plu4238 (purple) 3D-model of the binding pockets compared to the binding pocket of PA-IIL (yellow).
3D diagram of the binding pocket of the lectin domain of Plu4230 (orange), Plu4231 (red) and Plu4238 (orange), which was predicted by SWISS-MODEL server (Section 2.14), taking the structure of PA-IIL (yellow) as a template. Image generated using Pymol.

The variation of the amino acids within the sugar binding loop does appear to have an impact on the formation of the binding pocket. This impact may lead to variation of sugar binding preference of each of the proteins, as was observed for the binding pocket of RS-IIL. The topology of the binding pocket of RS-IIL accommodates the O6 of mannose, resulting in strong hydrogen binding of mannose and the RS-IIL preference of

mannose over fucose. In the binding pocket of PA-IIL there is no space for the O6 of mannose which is therefore directed towards the surface, away from the protein, resulting in mannose adopting a different orientation and ultimately resulting in weaker binding for mannose than found for RS-IIL-mannose binding (Sudakevitz *et al*. 2004).

A BLAST of N terminal regions of Plu4230, Plu4231 and Plu4238 against the protein sequence database indicated sequence similarity of Plu4231 to matrix metalloproteinase (MMP) as was found for Photopexin A and B from *P. luminescens*. MMP are a family of structurally related, zinc and calcium dependent proteins that have been associated with the break down of connective tissue in pathogenic conditions, in wound healing and pregnancy (Borkakoti 1998). The N-terminal domain of Photopexin A and B were assigned a 'haemopexin-like' motif by Crennell, S. *et al.* (2000) based on protein alignments and protein modelling. Crennell, S. *et al.* (2000) speculated that ppxA and ppxB are involved in host surface attachment or the binding of haem-like iron containing compounds scavenged from the insect host. The significant sequence homolog between the N-terminal domain of Photopexin B and the N-terminal domain of Plu4231 could suggest a similar protein function (Figure 6.4).

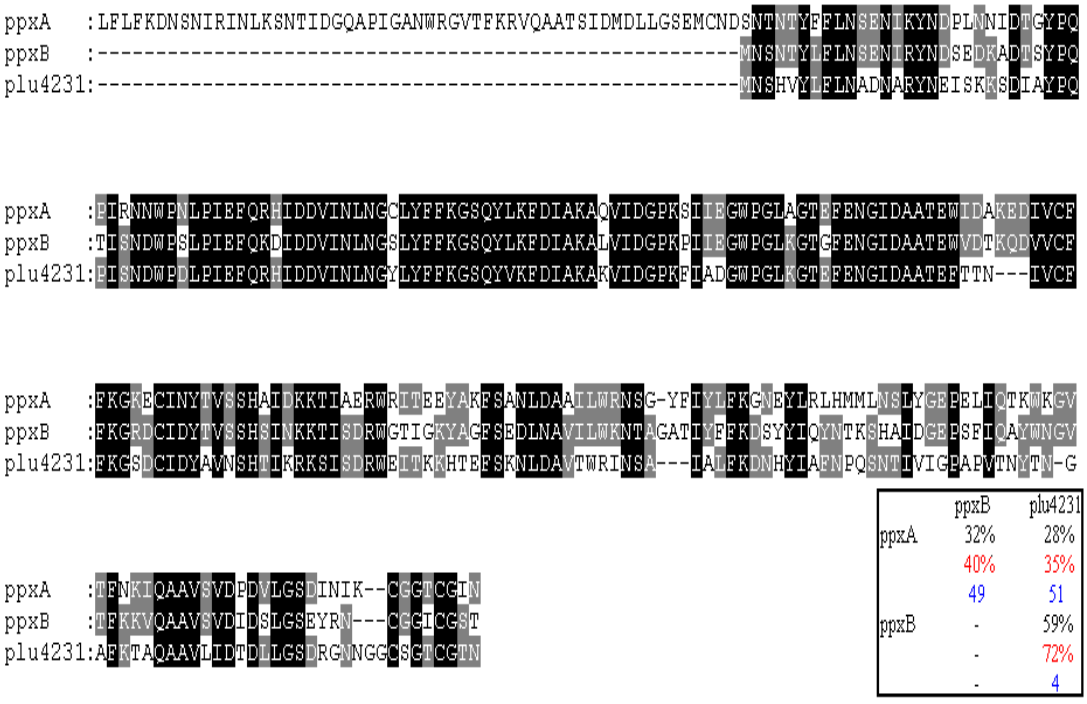


Figure 6.4: Sequence alignment of the N-terminal domains of PA-IIL homologs ppxA, ppxB and Plu4231, from *P. luminescens* TT01. Image and statistic reported generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey. Percentage identity –black, percentage similarity – red, gap character – blue.

No protein of significant sequence similarity to Plu4230 or Plu4238 was identified by a sequence database search although significant similarity was found between the proteins (Figure 6.5). Therefore a hypothesis for the putative biological role of the protein could not be formulated as there is no conserved protein or conserved motif upon which to base the hypothesis.

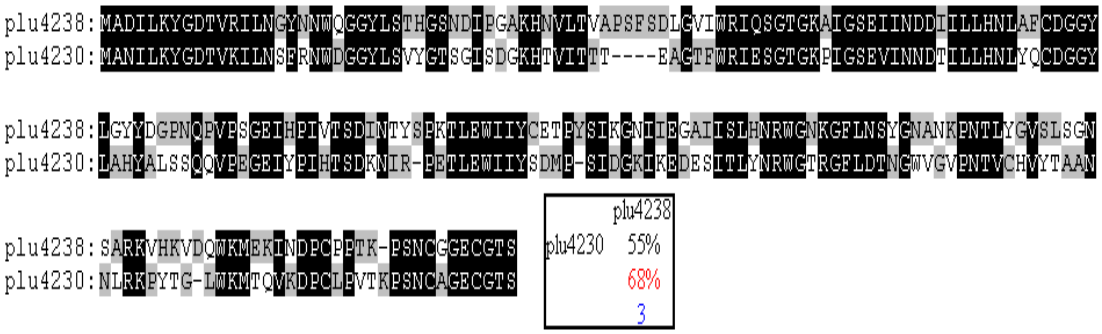


Figure 6.5: Sequence alignment of the N-terminal domains of PA-IIL homologs Plu4230 and Plu4238, from *P. luminescens* TT01. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

Ultimately what is required is the purification and X-ray crystallography structural studies of each of these proteins, as only limited information can be obtained from a predicted model base on sequence homology.

6.1.2 PA-IIL homologues from *P. asymbiotica*

A BLAST search of the NCBI sequencing database identified two open reading frames in the recently sequenced genome of *Photorhabdus asymbiotica* as having significant sequence similarity to PA-IIL. There are 3 recognised *Photorhabdus* species currently, *P. luminescens* and *P. temperate* that form symbiotic relationships with nematodes that infect insects and *P.asymbiotica*. *P. asymbiotica* has only been isolated from human clinical samples and as yet an invertebrate vector has yet to be identified. *Photorhabdus* species are gram negative, motile, facultatively anaerobic and rob shaped bacteria (Peel et al. 1999). The hypothetical genes from *P. asymbiotica* would encode:

Pasy1 – 340 amino acid protein with a calculated MW of 37.7kDa

Pasy2 – 363 amino acid protein with a calculated MW of 40.39kDa

Multiple protein alignment reveals a C-terminal domain with homology to PA-IIL in particular across the amino acids between the calcium binding loop. A higher sequence

homology was found for the homologs of *P. asymbiotica* and the Photopexins of *P. luminescens* then with PA-IIL (see Figure 6.6 and 6.7).

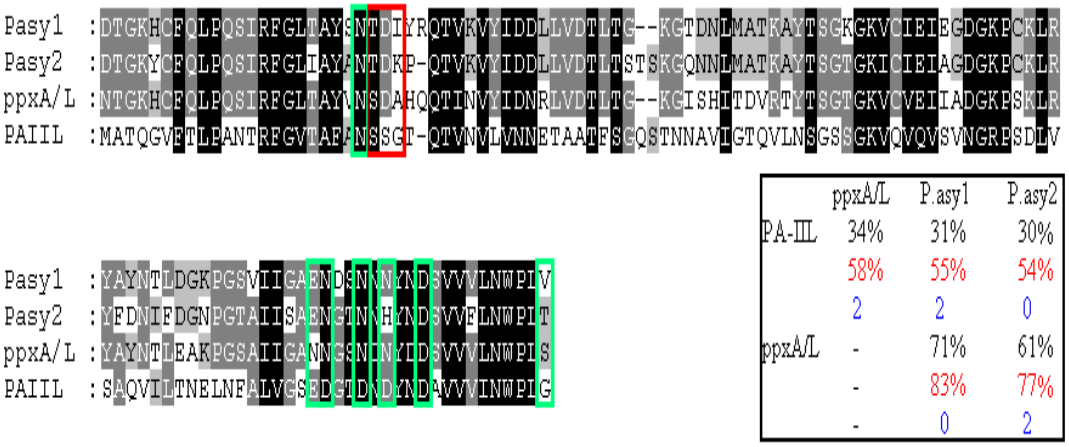


Figure 6.6: Sequence alignment of the lectin-like C-terminal domains of the homologues from *P. asymbiotica*.
Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

There is significant sequence similarity of the proteins within the *P. luminescens* strain, however within the sugar binding region the amino acids are not as highly conserved between the proteins (as highlighted in red in Figure 6.6).

	PA-III	ppxA/L	ppxB/L	P.asy2
	29%	67%	57%	76%
P.asy1	51%	77%	72%	83%
	8	6	8	2
	28%	57%	54%	-
P.asy2	51%	72%	68%	-
	6	7	9	-

Figure 6.7: Homologous lectins; *P.aeruginosa* PA-IIL, lectin domain of *P. luminescens* ppxA/L and ppxB/L and lectin region of *P. asymbiotica* P.asy1 and P.asy2.
Percentage identity –black, percentage similarity – red, gap character – blue. Statistic report created using ClustalW and GenDoc.

Sequence alignment of the protein sequence of PA-IIL to the lectin domains found within *P. asymbiotica* found that the proteins have an exact identity of 28-29%, with similarity of 51% (Figure 6.7). A greater number of null characters (Figure 6.7, blue numbering) were necessary to align the homologous residues of *P. asymbiotica* and PA-IIL and the lowest percentage identity between the proteins indicate that the

homologous from the *P. asymbiotica* are the most distally related members of the PA-IIL family. The lectin domain of P.asy1 and P.asy2 were found to have an identity slightly lower then 30% to PA-IIL. Therefore, it should be noted that a 3D structure model of each of the unknown proteins created using the crystal structure of PA-IIL as a template, is likely to contain errors. The amino acid sequence for each of the proteins was submitted to SWISS-MODEL web server (see Section 2.14) (Schwede et al. 2003). The 3D structures were predicted using the PA-IIL PDB file 1OUX as a template. The binding pocket from each of the 3D-models was compared to the binding pocket of PA-IIL and the results shown in Figure 6.8.

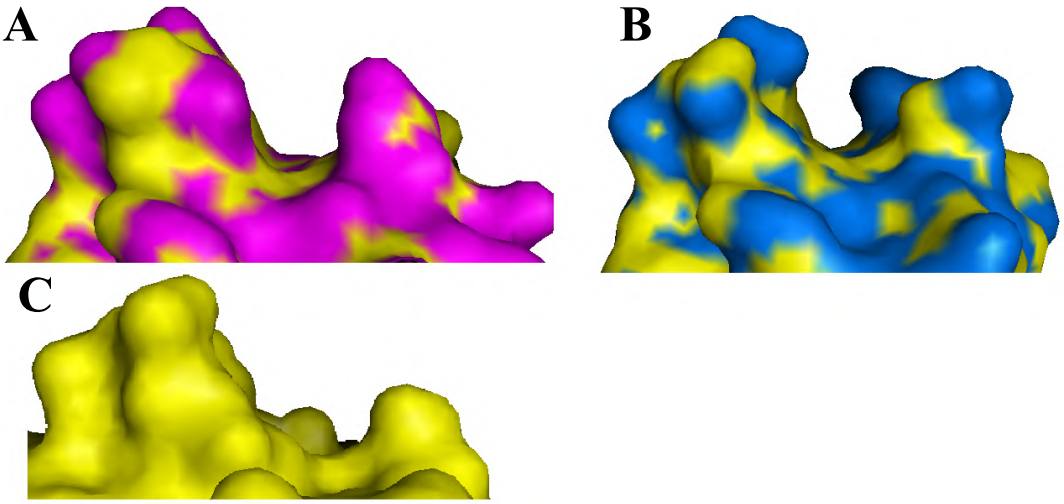


Figure 6.8: P.asy1 (pink) and P.asy2 (blue) 3D-model of the binding pockets compared to the binding pocket of PA-IIL (yellow).
3D diagram of the binding pocket of the lectin region of P.asy1 (pink) and P.asy2 (blue), which was predicted by SWISS-MODEL server (Section 2.14), taking the structure of PA-IIL (yellow) as a template. Image generated using Pymol.

The variation of the amino acids within the sugar binding loop does appear to have an impact on the formation of the binding pocket. This impact may lead to variation of sugar binding preference of each of the proteins, as was observed for the binding pocket of RS-IIL.

A BLAST of N terminal regions of P.asy1 and P.asy2 against the protein sequence database indicated sequence similarity to matrix metalloproteinase (MMP) as was found for Photopexin A and B from *P. luminescens*. The significant sequence homolog between the N-terminal domain of Photopexin B and the N-terminal domains of both P.asy1 and P.asy2 indicate a similar protein function, i.e. the hemaopexin motif

assigned to the N-terminal domain of ppxB may be assigned to the N-terminal domains of P.asy1 and P.asy2 (Figure 9.6). Again, as with the PA-IIL homologues from *P. luminescens,* only limited information can be obtained from a predicted model base on sequence homology Therefore what is ultimately required is the purification and X-ray crystallography structural studies of each of these proteins.

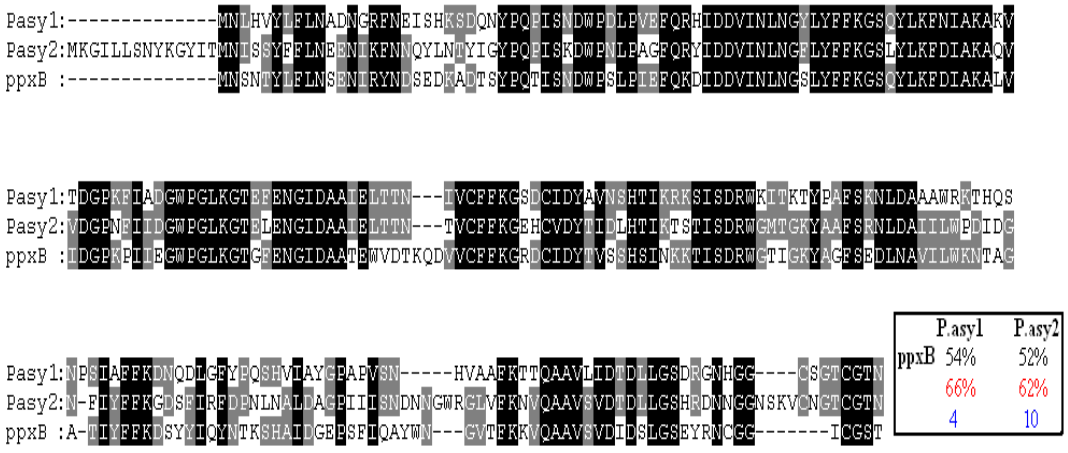


Figure 6.9: Sequence alignment of the N-terminal domains of PA-IIL homologs P.asy1 and P.asy2, from *P. asymbiotica* and the N terminal domain of Photopexin B from *P. luminescens*. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

6.2 Overview of the potential of the lectins of the PA-IIL superfamily

To date all homologues of PA-IIL identified were proteins with a highly conserved calcium binding loop and diversity within the amino acids of the sugar binding loop (Figure 6.10). The protein crystal structure of PA-IIL revealed an unusual protein/carbohydrate interaction formed by the two Ca^{+} ions, the amino side chain that co-ordinates the Ca^{2+} ions and the sugar. This was thought to be the basis for the relatively strong binding affinity found for PA-IIL (Mitchell *et al*. 2002). It was theorized, due to the high degree of conservation of the amino acids of the calcium binding loop, that homologues of PA-IIL bind in a similar manner to PA-IIL and therefore potentially have a strong binding affinity in the range of PA-IIL (Ka 1.5 $X10^{6}M^{-1}$, (Garber *et al*. 1987)). The homologs CV-IIL, RS-IIL and BC-IIL (*Burkholderia cenceopica* lectin II) has been shown to bind in the same manner via Ca^{2+}

ions by Pokorná, M. *et al*. (2006), Sudakevitz, D. *et al.* (2004) and Lameignere, E. *et al*. (2008) respectively.

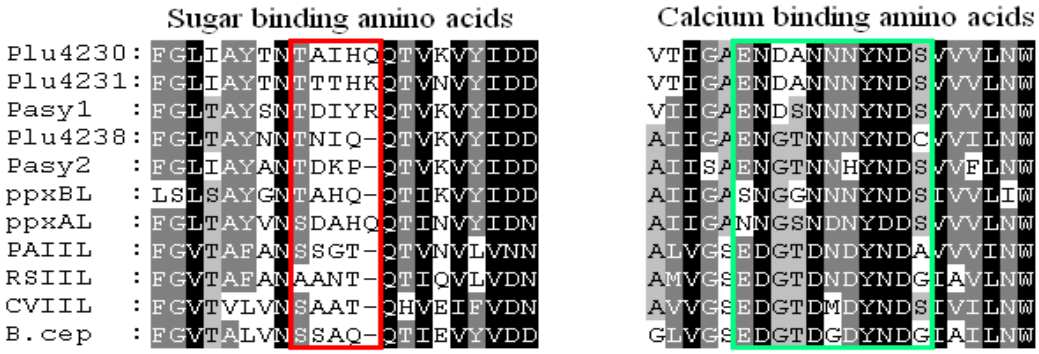


Figure 6.10: The variation/ found within the sugar binding loop and the conserved amino acids of a number of PA-IIL homologues.
B.cep – fucose lectin II *Burkholderia cenocepacia* AU1054, CV-IIL – fucose lectin II *Chromobacterium violaceum*, RS-IIL – mannose lectin II *Ralstonia Solanacearum,* PA-IIL – fucose lectin II *Pseudomonas aeruginosa*, ppxA/L, ppxB/L, Plu4230, Plu4231 and Plu4238 – hypothetical proteins *Photorhabdus luminescens,* P.asy1 and P.asy2 hypothetical proteins *Photorhabdus asymbiotica*. Amino acids involved in monosaccharide binding outlined in red and amino acids involved in calcium binding outlined in green. Image generated using ClustalW align and Genedoc. Identical residues highlighted in black and similar in grey.

A number of bacteria identified as containing homologues to PA-IIL are of interest as they have a high economic and social impact. There is a high mortality rate in immunodeficient individuals associated with infection by *C. violaceum* (Alves de Brito *et al.* 2004), *P. aeruginosa* (Bodey *et al.* 1983) and *Burkholderia spp* (Mahenthiralingam, Vandamme 2005) which has significant social impact, particularly in specific sections of society such as sufferers of Cystic Fibrosis (CF) and Acquired Immune Deficiency Syndrome (AIDS). *P. asymbiotica* has also been associated with infection of soft tissue of humans and disseminated bacteraemic infections. *P. asymbiotica* has been isolated from human clinical specimens from the USA and Australia (Gerrard *et al*. 2004). *R. solanacearum* is a plant pathogen that has a significant economic impact due to its extensive host range of over 200 plant species including economically important hosts such as potatoes, tomatoes and tobacco (Hayward 1995)

A common means of adhesion, expressed by numerous bacteria, are surface lectins that combine with complementary carbohydrates present on the host cell surfaces as an initial stage of infection (Sharon, Lis 1989). This indicates that lectins may serve as virulence factors for these organisms. Detailed characterisation of the combining site of the bacterial surface lectins is important not only for gaining a better insight into the nature of the interaction between bacteria and cell surfaces, but also for the design of

more effective inhibitors of adhesion to surfaces and thus for the prevention of infection. Therefore, characterisation of the binding specificity of proteins of the PA-IIL superfamily may contribute to a better understanding of the mechanism of infection of a number of organisms and potential anti-adhesion therapies by the addition of carbohydrate analogues.

The use of bacterial lectins in analytical platforms is more desirable then lectins from a eukaryotic source, such as plant and animal lectins. Eukaryotic lectins are often complex molecules that require post-translational modification (PTM). Therefore, *E. coil* expression systems often can't be utilised to produce them as *E. coli* lacks the ability to carry out complex PTMs such as glycosylation. Lectins from prokaryotic sources, such as the PA-IIL superfamily, do not require PTM and can therefore exploit the *E. coli* system for the expression of high yields of recombinant protein. A broad range of cloning and expression vectors for recombinant protein expression from the *E. coli* are available enabling flexible control of protein expression, e.g. addition of fusion proteins to facilitate one-step-purification. The relatively inexpensive production of large quantities of recombinant lectin that can be highly purified produced via the *E. coli* expression system can then be utilised for the production of lectin microarrays (Figure 6.11).

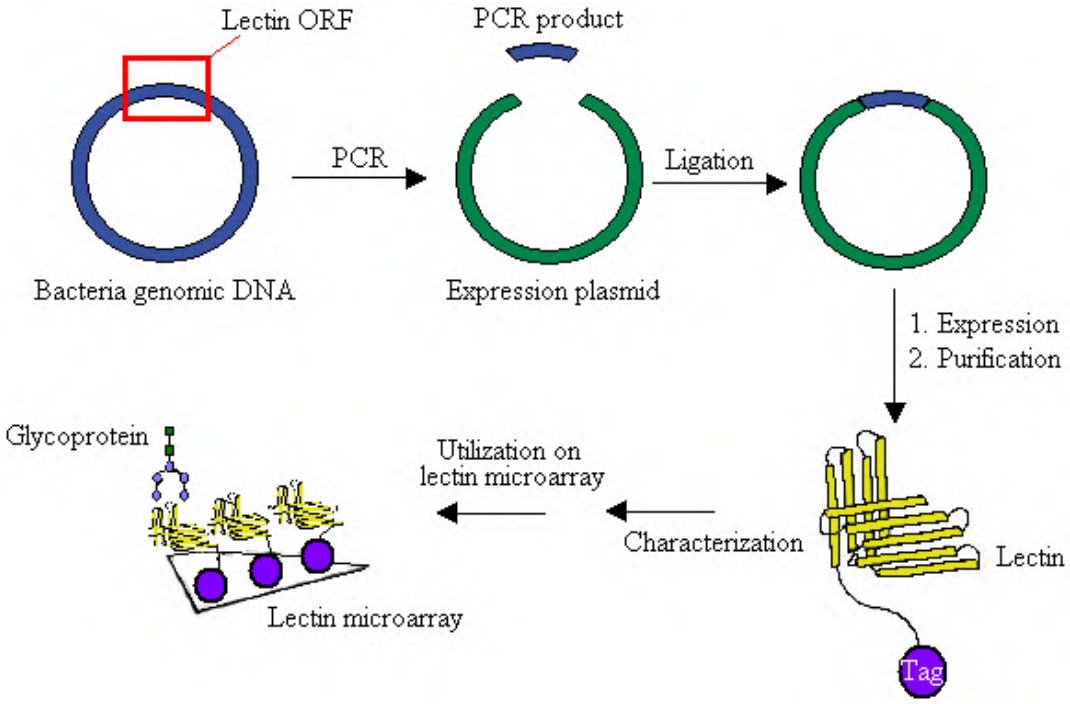


Figure 6.11: Fabrication of a recombinant lectin microarray.

A lectin microarray enables fast, sensitive and high throughput profiling of the sugar structures of a protein glycan composition with subtle differences. Previously they have been used for microbial typing (Hsu, Pilobello and Maha 2006) and differential profiling, e.g. differentiating between normal and cancerous epithelia (Matsudaa *et al*. 2008). An ideal lectin microarray would consist of a high density array with a diverse panel of immobilized lectins in order to achieve a distinct and characteristic analysis of a glycome. A limitation to the development of lectin microarrays is the lack of commercial availability of lectins that diversely recognize unique sugar structures. The uncharacterised lectins of the PA-IIL superfamily are potentially novel lectins of prokaryotic origin, theoretically with a relatively high binding affinity due to the predicted unique protein/carbohydrate binding formed by Ca^{2+}. Due to their prokaryotic origin these lectins can be cloned and expressed via the *E. coli* expression system, to produce high yields of protein and purified to a high degree by the addition of an affinity fusion tag. These fusion tags may also be used for directed orientation-specific immobilisation to an activated surface of a microfluidic platform.

6.3 Summary

In this work, protocols for the cloning, expression and purification of the *PA-IIL* gene from *P. aeruginosa* were optimised. Due to the significant sequence homology of PA-IIL and the PA-IIL homologues it was hypothesised that the established protocols would be applicable for the cloning, expression and purification of the uncharacterised homologues. The establishment of general protocols for investigation of the PA-IIL superfamily was predicted to streamline the characterisation of these lectins.

E. coli is one of the most commonly used hosts of recombinant expression as it facilitates protein expression by its relative simplicity, rapid high density growth rate on inexpensive substrates and well characterized genetics which allow for large scale protein production (Baneyx 1999). PA-IIL was successfully cloned into a number of pET vectors and expressed in the *E. coli* system in a number of studies, Loris, R. *et al.* (2003), Sabin, C. *et al.* (2006) and Tielker, D. *et al.* (2005). In all studies the protein was purified over mannose agarose as described by Gilboa-Garber, N. *et al.* (1983). Similarly the characterised members of the PA-IIL superfamily, RS-IIL, CV-IIL and BclA, were cloned into *E. coli* expression vectors and were purified over mannose agarose (Lameignere *et al.* 2008, Sudakeuitz, Datcoff and Gilboa-Garber 2002, Zinger-

Yosovich *et al*. 2006). To purify a lectin over a resin displaying an immobilised oligosaccharide, such as mannose-agarose, requires knowledge of the specificity of the lectin. The requirement of prior knowledge of lectin specificity to facilitate purification limits the ability to investigate previously uncharacterised novel lectins. Utilising the affinity tag, 6HIS, facilitates the purification of a lectin regardless of the binding specificity of the lectin. The ability to highly purify novel lectins aids in the structural and functional characterisation of these proteins. Histidine tagging is also advantageous as it combines a number of desirable features such as small size, interaction with a relatively inexpensive matrix (Ni-NTA resin) which can tolerate multiple regeneration cycles and a high binding affinity of 6HIS to the matrix. Therefore the use of the *E. coli* expression system and 6HIS affinity tagging enables characterisation of novel lectins and the large scale production of highly purified lectins that can then be used for the development of diagnostic lection microarrays and other platforms. To data no work has been published on the use of the affinity tag 6HIS for the purification of PA-IIL or any of its homologs.

The protocols established by this work for the expression of recombinant 6HIS tagged PA-IIL were successfully transferred to the expression and purification of the lectin domain of Photopexin A and B. The yield of each lectin expressed did vary. Sufficient yields of PA-IIL and ppxA/L were purified for subsequent characterisation studies. Some common traits were found for purified PA-IIL and ppxA/L: inherent instability in solution leading to precipitation, preferred temperature for storage in solution at 25^{o}C and sugar specific interaction with BSA solutions. Sufficient yields of ppxB/L were not obtained from the expression and purification of ppxB/L for characterisation studies. There are several options regarding further optimisation of the expression of ppxB/L to enhance protein yield, such as a 96 well plate clonal selection, expression from various strains of *E. coli* or media composition (LB was the only media used for expression). The protocols established in this work for the expression of PA-IIL could be used as a basis for the production of the remaining uncharacterised lectins of *P. luminescens* and *P. asymbiotica*. The next step of this work would be to establish a suitable method for the characterisation of these lectins.

Data from this work indicates that the ELLA may not be a suitable method for the investigation of sugar binding specificities of the PA-IIL superfamily and a number of alternative methods have been discussed (see Section 5.3).

A search of the literature found that the binding affinity of PA-IIL and homologs of PA-IIL were investigated by a number of methods. The binding specificity of PA-IIL was originally determined by Gilboa-Garber, N. *et al.* (1983). These results were confirmed by Mitchell, E. *et al.* (2005) by repetition of the haemagglutination assay and also by ITC. This study confirmed the order of sugar preference of PA-IIL but found the binding affinity to be one order of magnitude lower (Ka; 1.5 x $10^6 M^{-1}$) by ITC. A study by Sudakevitz, D. *et al.* (2004) successfully determined the binding specificity of the RS-IIL lectin and compared the binding to that of PA-IIL by the haemagglutination assays using human erythrocytes. The binding specificity of CV-IIL was also initially determined using the haemagglutination assay. The binding specificity of CV-IIL was compared to that of RS-IIL and PA-IIL using erythrocytes from a number of sources (human, cow, rabbit, rat and sheep). A study by Pokorna, M. *et al.* (2006) further investigated the binding mode of CV-IIL by ELLA and ITC. The results for CV-IIL were compared to that of PA-IIL. BSA was used as the blocking agent for the ELLA, however the data was not published in this paper and it was therefore not possible to compared background noise obtained by this group when using BSA to the results obtained during this work. Most recently the binding specificity of BclA was determined by the SPR based BIAcore system (Lameignere *et al*. 2008). Specific sugar was immobilised on the surface of a BIAcore chip and binding to the sugar by BclA was monitored. Although ITC and BIAcore have been proven as suitable methods for the investigation of the PA-IIL superfamily both methods require specialised equipment and in the case of BIAcore, specialised training. FAC whereby an oligosaccharide is immobilized onto a column and a GFP labelled lectin passed over the column potentially offers a relatively easy and inexpensive method for the characterisation of lectin binding. By GFP labelling lectins, protein elution from an oligosaccharide column can easily be monitored by a fluorescent detector or by visualisation using a UV light source. However this method has not been established and requires further investigation.

Future work should also include both directed and random mutation of the specific sugar-binding residues/domains of PA-IIL. It was previously established by Adam, J. *et al.* (2007) that point mutation of these amino acids can alter the binding specificity of PA-IIL (see Section 1.6). Characterisation of the binding specificity of the homologues of PA-IIL may lead to the identification of a specific lectin of interest. It was theorized that the cloning of the binding site from one homologue of interest to the

PA-IIL molecule could be carried out by point mutation, in order to develop a larger range of lectins with diverse specificities. Therefore an array of diverse specificities may be created within the basic framework of one lectin molecule. Ultimately this may lead to the creation of a uniform recombinant lectin microarray where each well contains the same core lectin molecule, each having differing glycan specificity. The advantage to such a lectin-array would be having the same optimal conditions, e.g. pH, temperature, co-factor requirement, buffer composition, for each lectin. Also of interest would be the random mutagenesis of the specific sugar-binding residues to incorporate very different residues beyond those identified within the sugar binding loop of PA-IIL homologs for the creation of novel lectins. Recombinant lectins of the PA-IIL superfamily, as demonstrated in this work, can be produced to a high level of purity through the *E. coli* expression system and one-step-purification via a fusion protein. Introduction of immobilised tags into the protein structure also facilitates the orientation-specific immobilisation to an activated surface creating a more homogenous distribution of lectin on an array surface.

In conclusion, the characterisation of the lectins of the PA-IIL superfamily and other novel lectins is increasingly important for the development of accurate, high throughput determination of glycosylation patterns of molecules. Expansion of our tools for the comprehensive glycoprofiling is essential for the increasingly relevant field of glycobiology. Compiling the necessary means of deciphering the glycoprofile is increasingly important for both basic research and the biopharmaceutical industry for the comprehensive determination of therapeutic biomolecules. It is also important for clinical diagnostics for the identification of disease states and creation of anti-adhesive therapeutics for bacterial infection. The identification or creation of diverse recombinant lectins will facilitate the production of the next generation of lectin microarrays.

References

-A-

Adam, E.C., Schumacher, D.U. and Schumacher, U. (1997) "Cilia from a cystic fibrosis patient react to the ciliotoxic *Pseudomonas aeruginosa* II lectin in a similar manner to normal control cilia - a case report.", *J. Laryngol. Otol.* **8** (111): 760-762.

Adam, J., Pokorná, M., Sabin, C., Mitchell, E., Imberty, A. and Wimmerová, M. (2007) "Engineering of PA-IIL lectin from *Pseudomonas aeruginosa* – Unravelling the role of the specificity loop for sugar preference." *BMC Struct. Biol.* **7**: 36-48.

Afrough, B., Dwek, M.V. and Greenwell, P. (2007) "Identification and elimination of false-positives in an ELISA-based system for qualitative assessment of glycoconjugate binding using a selection of plant lectins." *BioTechniques,* **43** (4): 458-464.

Agrawal, B.B. and Goldstein, I.J. (1965) "Specific binding of concanavalin A to cross-linked dextran gels." *Biochemical J.* **96** (3): 23-25.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) "Basic local alignment search tool." *J. Mol. Biol.* **215** (3): 403-410.

Alves de Brito, C.F., Carvalho, C.M.B., Santos, F.R., Gazzinelli, R.T., Oliveira, S.C., Azevedo, V.A. and Teixeira, S.M.R. (2004) "*Chromobacterium violaceum* genome: molecular mechanisms associated with pathogenicity." *Genet. Mol. Res.* **3** (1):148.

Ambrosi, M., Cameron, N.R. and Davis, B.G. (2005) "Lectins: tools for the molecular understanding of the glycocode." *Org. Biomol. Chem.* **3** (9):1593-1608.

Andrews, P. (1964) "Estimation of the molecular weights of proteins by Sephadex Gel-Filtration." *Biochem. J.* **91** (2): 222-233.

Arora, S.K., Ritchings, B.W., Almira, E.C., Lory, S. and Ramphal, R. (1998) "The *Pseudomonas aeruginosa* flagellar cap protein, FliD, is responsible for mucin adhesion." *Infect. Immun.* **66** (3):1000-1007.

Ashwell, G. and Morell, A.G. (1974) "The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins." *Adv. Enzymol. Relat. Area Mol. Biol.* **41**: 99-128.

Aub, J.C., Sanford, B.H. and Cote, M.N. (1965) "Studies on reactivty of tumor and normal cells to a wheat germ agglutinin." *PNAS- Biol. Sciences,* **54** (2): 396-399.

-B-

Bachhawat, K., Thomas, C.J. and Amutha, B. (2001) "On the stringent requirement of mannosyl substitution in mannooligosaccharides for the recognition by garlic (*Allium sativum*) Lectin." *J. Biol. Chem.* **276** (23): 5541-5546.

Bairoch, A. and Apweiler, R. (1997) "The SWISS-PROT protein sequence data bank and its supplement TREMBL." *Nucleic Acids Res.* **25** (1): 31-36.

Bajolet-Laudinat, O., Girod-de Bentzmann, S., Tournier, J.M., Madoulet, C., Plotkowski, M.C., Chippaux, C. and Puchelle, E. (1994) "Cytotoxicity of *Pseudomonas aeruginosa* internal lectin PA-I to respiratory epithelial cells in primary culture." *Infect. Immun*. **62** (10): 4481-4487.

Ball, G., Durand, R., Lazdunski, A. and Filloux, A. (2002) "A novel type II secretion system in *Pseudomonas aeruginosa.*" *Mol. Microbiol.* **43** (2): 475-485.

Baneyx, F. (1999) "Recombinant protein expression in *Escherichia coli*", *Current Opinion in Biotechnology,* **10** (5): 411-421.

Benson, D.A., Boguski, M., Lipman, D.J. and Ostell, J. (1996) "GenBank." *Nucleic Acids Res.* **24** (1): 1-5.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N. and Bourne, P.E. (2000) "The Protein Data Bank." *Nucleic acids Res.* **28** (1): 235-242.

Berrow, N.S., Bussow, B., Coutard, B., Diprose, J., Ekberg, M., Folkers, G.E., Levy, N., Lieu, V., Owens, R.J., Peleg, Y., Pinaglia, C., Quevillon-Cheruel, S., Salim, L., Scheich, C., Vincentellic, R. and Bussoh, D. (2006) "Recombinant protein

expression and solubility screening in *Escherichia coli*: a comparative study." *Acta Crystallogr. D* **62**: 1218-1226.

Birnboim, H.C. and Doly, J. (1979) "A rapid alkaline extraction procedure for screening recombinant plasmid DNA." *Nucleic Acids Res.* **7** (6):1513-1523.

Bodey, G.P., Bolivar, R., Fainstein, V. and Jadeja, L. (1983), "Infections caused by *Pseudomonas aeruginosa.*" *Rev. Infect. Dis.* **5** (2): 279-313.

Borkakoti, N. (1998), "Matrix metalloproteases: variations on a theme." *Prog. Biophys. Mol. Biol.* **70**, 1: 73-94.

Bowen, D., Rocheleau, T.A., Blackburn, M., Andreev, O., Golubeva, E., Bhartia, R. and ffrench-Constant, R.H. (1998) "Insecticidal toxins from the bacterium *Photorhabdus luminescens*." *Science,* **280** (5372): 2129-2132.

Brazilian National Genome Project Consortium. (2003) "The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability." PNAS-*Biol. Sci.* **100** (20).

Brewer, C.F., Brown, R.D. and Koenig, S.H. (1983) "Metal ion binding and conformational transitions in concanacalin A: a structure-funtion study." *J. Biomol. Struct. Dyn.* **1** (4): 961-997.

Brint, J.M. and Ohman, D.E. (1995) "Synthesis of multiple exoproducts in *Pseudomonas aeruginosa* is under the control of RhlR-RhlI, another ser of regulators in strain PAO1 with homology to the autoinducer-responsive LuxR-LuxI family." *J. Bacteriol.* **177** (24): 7155-7163.

Bujard, H., Gentz, R., Lanzer, M., Stueber, D., Mueller, M., Ibrahimi, I., Haeuptle, M.T. & Dobberstein, B. (1987) "A T5 promoter-based transcription-translation system for the analysis of proteins *in vitro* and *in vivo*" in *Methods in Enzymology,* ed. W. Ray, Academic Press. 416-433.

-C-

Chemani, C., Imberty, A., de Bentzmann, S., Pierre, M., Wimmerová, M., Guery, B.P. and Faure, K. (2009) "Role of LecA and LecB in *Pseudomonas aeruginosa* induced lung injury and effect of carbohydrate ligand." *Infect. Immun.* **77** (5): 2065-2075.

Clarke, L.A., Rebelo, C.S., Gonçalves, J., Boavida, M.J. and Jordan, P. (2001) "PCR amplification introduces errors into mononucleotide and dinucleotide repeat sequences." *J. Clin. Pathol.* **54**: 351-353.

Cole, J.L. and Hansen, J.C. (1999) "Analytical ultracentrifugation as a contemporary biomolecular research tool." *J. Biomol. Tech.* **10** (4):163-176.

Crennell, S.J. (2000) "The predicted structure of Photopexin from *Photorhabdus* shows the first haemopexin-like motif in prokaryotes." *FEMS Microbiol. Lett.* **191**: 139-144.

-D-

Dam, T.K. and Brewer, C.F. (2002) "Thermodynamic study of lectin-carbohydrate interactions by isothermal titration calorimetry." *Chem. Rev.* **102** (2): 387-429.

DeLano, W.L. (2002) "The PyMol molecular graphics system." DeLano Scientific, San Carlos, CA, USA

Denny, T.P. (2000) "*Ralstonia solanacearum* – a plant pathogen in touch with its host." *Trends Microbiol.* **8** (11): 486-489.

Devine, J.H., Shadel, G.S. and Baldwin, T.O. (1989) "Identification of the operator of the lux regulon from the *Vibrio fischeri* strain ATCC7744", *PNAS-Biol. Sci.* **86** (15): 5688-5692.

Duchaud, E., Rusniok, C., Frangeul, L., Buchrieser, C., Givaudan, A., Taourit, S., Bocs, S., Boursaux-Eude, C., Chandler, M., Charles, J.F., Dassa, E., Derose, R., Derzelle, S., Freyssinet, G., Gaudriault, S., Medigue, C., Lanois, A., Powell, K., Siguier, P.,

Vincent, R., Wingate, V., Zouine, M., Glaser, P., Boemare, N., Danchin, A. and Kunst, F. (2003) "The genome sequence of the entomopathogenic bacterium *Photorhabdus luminescens*." *Nat. Biotechnol.* **21** (11): 1307-1313.

Dubin, P.L. and Principi, J.M. (1989) "Failure of universal calibration for size exclusion chromatography of rod like macromolecules versus random coils and globular proteins." *Macromol.* **22** (4): 1891-1896.

-E-

Eberl, L. and Tümmler, B. (2004) "*Pseudomonas aeruginosa* and *Burkholderia cepacia* in cystic fibrosis: genome evolution, interactions and adaptation." *Int. J. Med. Microbiol.* **294** (2-3): 123-131.

Egland, K.A. and Greenberg, E.P. (1999) "Quorum sensing in *Vibrio fischeri*: elements of the *luxi* promoter", *Mol. Microbiol.* **31** (4): 1197-1204.

Elgavish, S. and Shaanan, B. (1998) "Structures of the *Erythrina corallodendron* lectin and of its complexes with mono- and disaccharides." *J. Mol. Biol.* **277** (4): 917-932.

-F-

Fisher, J., Klein, P.J., Vierbuchen, M., Skutta, B., Vhlenbruck, G. and Fisher, R. (1984) "Characterisation of glycoconjugates of lectin binding sites in normal alimentarty tract as well as in benign and malignant gastic neoplasms." *J. Histochem. Cytochem.* **32** (7): 681-689.

Frank, D.A. (1997) "The exoenzyme S regulon of *Pseudomonas aeruginosa*." *Mol Microbiol.* **26** (4): 621-629.

Fuqua, W.C., Winans, S.C. and Greenberg, E.P. (1994) "Quorum sensing in bacteria: the LuxR-LuxI Family of cell density-responsive transcriptional regulators." *J. Bacteriol.* **176** (2): 269-275.

-G-

Gambello, M.J. and Iglewski, B.H. (1991) "Cloning and characterization of the *Paeudomonas aeruginosa* lasR gene, a transcriptional activator of elastase expression." *J. Bacteriol.* **173** (9): 3000-3009.

Gambello, M.J., Kaye, S. and Iglewski, B.H. (1993) "LasR of *Pseudomonas aeruginosa* is a transcriptional activator of the alkaline protease gene (apr) and an enhancer of exotosin A expression." *Infect. Immun.* **61** (4): 1180-1184.

Garber, N., Guempel, U., Gilboa-Garber, N. and Royle, R.J. (1987) "Specificity of the fucose-binding lectin of *Pseudomonas aeruginosa*." *FEMS Microbiol. Lett.* **48** (3): 331-334.

Gekkot, K. and Timasheff, S.N. (1981) "Mechanism of protein stabilization by glycerol: Preferential hydration in glycerol-water mixtures." *Biochemistry,* **20**: 4667-4676

Gemeiner, P., Mislovičová, D., Tkáč, J., Švitel, J., Pätoprstý, V., Hrabárová, E., Kogan, G. and Kožár, T. (2009) "Lectinomics II. A highway to biomedical/clinical diagnostics." *Biotechnol. Adv.* **27**: 1-15.

Gerrard, J., Waterfield, N., Vohra, R. and ffrench-Constant, R. (2004) "Human infection with *Photorhabdus asymbiotica:* an emerging bacterial pathogen." *Microb. Infect.* **6** (2): 229-237.

Geyer, H. and Geyer, R. (2006) "Strategies for analysis of glycoprotein glycosylation." *Biochim. Biophys. Acta - Proteins and Proteomics,* **1764** (12): 1853-1869.

Gilboa-Garber, N. (1972) "Inhibition of broad spectrum hemagglutinin from *Pseudomonas seruginosa* by D-galactose and its derivates." *FEBS Lett.,* **20** (2): 242-244.

Gilboa-Garber, N. and Mizrahi, L. (1973) "Peroxidase attachment to sepharose mediated by bacterial hemagglutinin of *Pseudomonas aeruginosa*." *Biochim Biophys Acta (BBA) - Protein Structure,* **317** (1): 106-113.

Gilboa-Garber, N., Mizrahi, L. and Garber, N. (1977) "Mannose-binding hemagglutinins in extracts of *Pseudomonas aeruginosa*." *Can. J. Microbiol.* **55** (9): 975-981.

Gilboa-Garber, N. (1982) "*Pseudomonas aeruginosa* lectins." in *Methods in Enzymology* Academic Press, 378-385.

Gilboa-Garber, N., Katcoff, D.J. and Garber, N.C. (2000) "Identification and characterization of *Pseudomonas aeruginosa* PA-IIL lectin gene and protein compared to PA-IL." *FEMS Immunol. Med. Microbiol.* **29** (1): 53-57.

Glick, J. and Garber, N. 1983, "The intracellular localization of *Pseudomonas aeruginosa* lectins." *J. Gen. Microbiol.* **129** (10): 3085-3090.

Gupta, D., Cho, M., Cummings, R.D. and Brewer, C.F. (1996) "Thermodynamics of carbohydrate binding to Galectin-1 from chinese hamster ovary cells and two mutants. A comparison with four galactose-specific plant lectins." *Biochemistry,* **35** (48): 15236-15243.

-H-

Hanahan, D. (1985) "Techniques for transformation of *E. coli.*" in *DNA Cloning: A Practical Approach*, ed. D.M. Glover, Press Oxford: 109-135.

Hardman, A.M., Stewart, G.S.A.B. and Williams, P. (1998) "Quorum sensing and the cell-cell communication dependent regulation of gene expression in pathogenic and non-pathogenic." *Antonie van Leeuwenhoek,* **74** (4): 199-210.

Harvey, M.A. (1991) "Optimization of nitrocellulose membrane-based immunoassays." in Ed. S. Schleicher Keene.

Hassett , D.J., Ma, J., Elkins, J.G., McDermott, T.R., Ochsner, U.A., West, S.E.H., Huang, C., Fredericks, J., Burnett, S., Stewart, P.S., McFeters, G., Passador, L. and Iglewski, B.H. (1999) "Quorum sensing in *Pseudomonas aeruginosa* controls expression of catalase and superoxide dismutase genes and mediates biofilm susceptibility to hydrogen peroxide." *Mol. Microbiol.* **34** (5): 1082-1093.

Haycock, J.W. (1993) "Polyvinylpyrrolidone as a blocking agent in immunochemical studies." *Anal. Biochem.* **208** (2): 397-399.

Hayward, A.C. (1991) "Biology and epidemiology of bacterial wilt caused by *Pseudomonas solanacearum.*" *Annu. Rev. Phytopathol.* **29**: 65-87.

Hirabayash, J. (2008) "Concept, strategy and realization of lectin-based glycan profiling." *J. Biochem.* **144** (2): 139-147.

Hirakata, Y., Furuya, N., Tateda, K., Matsumoto, T. and Yamaguchi, K. (1995) "The influence of exo-enzyme S and proteases on endogenous *Pseudomonas aeruginosa* bacteraemia in mice." *J. Med. Microbiol.* **43**: 258-261.

Hirschberg, C.B. (2001) "Golgi nucleotide sugar transposrt and leukocyte adhesion deficiency II." *J. Clin. Invest.* **108** (1): 3-6.

Hirschhorn, J.N., Lohmueller, K., Byrne, E. and Hirschhorn, K. (2002) "A comprehensive review of genetic association studies." *Genet. Med*. **4** (2): 45-61.

Hsu, K.L., Pilobello, K.T. and Maha, L.K. (2006), "Analyzing the dynamic bacterial glycome with a lectin microarray approach." *Nat. Chem. Biol.* **2** (3): 153-157

Hsu, K.L., Gildersleevec, J.C. and Mahal, L.K.L. (2008) "A simple strategy for the creation of a recombinant lectin microarray." *Mol. BioSyst.* **4**: 654-662.

Huang, L., Hollingsworth, R.I., Castellani, R. and Zipser, B. (2004) "Accumulation of high-molecular-weight amylose in Alzheimer's disease brains." *Glycobiology,* **14** (5): 409-416.

Huet, C.H. and Bernadac, A. (1974) "Peroxidase binding to cell-bound Concanavalin A." *Exp. Cell Res.* **89** (2): 429-431.

-I-

Imberty, A., Wimmerova, M., Mitchell, E.P. and Gilboa-Garber, N. (2004) "Structures of the lectins from *Pseudomonas aeruginosa*: insights into the molecular basis for host glycan recognition." *Microb. Infect.* **6** (2): 221-228.

Inoue, H., Nojima, H. and Okayama, H. (1990) "High efficiency transformation of *Escherichia coli* with plasmids." *Gene,* **96** (1): 23-28.

-J-

Johansson, E.M.V., Crusz, S.A., Kolomiets, E., Buts, L., Kadam, R.U., Cacciarini, M., Bartels, K., Diggle, S.P., Camara, M., Williams, P., Loris, R., Nativi, C., Rosenau, F., Jaeger, K.E., Darver, T. and Reymond, J. (2008) "Inhibition and dispesion of *Pseudomonas aeruginosa* biofilms by glycopeptide dendrimers targeting the fucose-specific lectin LecB." *Chem. Biol.* **15**: 1249-1257.

Jones, S., Yu, B., Bainton, N.J., Birdsall, M., Bycroft, B.W., Chhabra, S.R., Cox, A.J., Golby, P., Reeves, P.J. and Stephens, S. (1993) "The lux autoinducer regulates the production of exoenzyme virulence determinants in Erwinia carotovora and *Pseudomonas aeruginosa*." *EMBO,* **12** (6): 2477-2489.

Joosten, J.A.F., Loimaranta, V., Appeldoorn, C.C.M., Haataja, S., Maate, F.A.E., Liskamp, R.M.J., Finne, J. and Pieters, R.J. (2004) "Inhibition of *Streptococcus suis* adhesion by dendritic galabiose compounds at low nanomolar concentration." *J. Med. Chem.* **47** (26): 6499-6508.

Jørgensen, F., Bally, M., Chapon-Herve, V., Michel, G., Lazdunski, A., Williams, P. and Stewart, G.S.A.B. (1999) "RpoS-dependent stress tolerance in *Pseudomonas aeruginosa.*" *Microbiology,* **145**: 835-844.

-K-

Kane, J.F. (1995) "Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coil.*" Curr. Opin. Biotechnol. **6** (5): 494-500.

Kasai, K. and Ishii, S. (1975) "Quantitative analysis of affinity chromatography of trypsin." *J. Biochem.* **77** (1): 261-264.

Katsuyama, T. and Spicer, S.S. (1978) "Histochemical differentiation of complex carbohydrates with variants of the ConA-horseradish peroxidase method." *J. Histochem. Cytochem.* **26** (4): 233-250.

Kearns, D.B., Robinson, J. and Shimkets, L.J. (2001) "*Pseudomonas aeruginosa* exhibitis directed twitching motility up phosphatidylethanolamine gradients." *J. Bacteriol.* **183** (2): 763-767.

Kolomiets, E., Swiderska, M.A., Kadam, R.U., Johansson, E.M.V., Jaeger, K.E., Darbre, T. and Reymond, J. (2009) "Glycopeptide dendrimers with high affinity for the fucose-binding lectin LecB from *Pseudomonas aeruginosa.*" *ChemMedChem,* **4** (4): 562-569.

-L-

Lameignere, E., Malinovska, L., Slavikova, M., Duchaud, E., Mitchelle, E.P., Varrot, A., Šedo, S., Imberty, A. and Wimmerova, M. (2008) "Structural basis for mannose recognition by a lectin from opportunistic bacteria *Burkholderia cenocepacia.*" *Biochem. J.* **411**: 307-318.

Laemmli, U.K. (1970) "Cleavage of structural proteins during the assembly of the head of bacteriophate T4", *Nature,* **227**: 680-685.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007) "Clustal W and Clustal X version 2.0." *Bioinformatics,* **23** (21): 2947-2948.

Latifi, A., Winson, M.K., Foglino, M., Bycroft, B.W., Stewart, G.S.A.B., Lazdunski, A. and Williams, P. (1995) "Multiple homologues of LuxR and LuxI control expression of virulence determinants and secondary metabolites through quorum sensing in *Pseudomonas aeruginosa* PAO1." *Mol. Microbiol.* **17** (2): 333-343.

Latifi, A., Foglino, M., Tanaka, K., Williams, P. and Lazdunski, A. (1996), "A hierarchical quorum-sensing cascade in *Pseudomonas aeruginosa* links the

transcriptional activators LasR and RhlR (VsmR) to expression of the stationary-phase sigma factor RpoS." *Mol. Microbiol.* **21** (6): 1137-1146.

Leavitt, S. and Freire, E. (2001) "Direct measurement of protein binding energetics by isothermal titration calorimetry." *Curr. Opin. Struct. Biol.* **11** (5): 560-566.

Lerrer, B. and Gilboa-Garber, N. (2001) "Interactions of *Pseudomonas aeruginosa* PA-IIL lectin with quail egg white glycoproteins." *Can. J. Microbiol.* **47** (12): 1095-1100.

Lesman-Movshovich, E. and Gilboa-Garber, N. (2003) "*Pseudomonas aeruginosa* lectin PA-IIL as a powerful probe for human and bovine milk analysis." *J. Dairy Sci,* **86** (7): 2276-2282.

Lis, H. and Sharon, N. (1998) "Lectins: Carbohydrate-specific proteins that mediate cellular recognition." *Chem. Rev.* **98** (2): 637-674.

Liu, P.V. (1966) "The roles of various fractions of *Pseudomonas aeruginosa* in its Ppthogenesis II. Effects of lecithinase and protease." *J. Infect. Dis.* **116** (1): 112-116.

Loris, R., Tielker, D., Jaeger, K.E. and Wyns, L. (2003) "Structural basis of Carbohydrate recognition by the lectin LecB from *Pseudomonas aeruginosa*." *J. Mol. Biol.* **331** (4): 861-870.

Loughran, S.T., Loughran, N.B., Ryan, B.J., D'Souza, B.N. and Walls, D. (2006) "Modified His-tag fusion vector for enhanced protein purification by immobilized metal affinity chromatography." Anal. Biochem. **1**: 148-150.

Lyczak, J.B., Cannon, C.L. and Pier, G.B. (2002) "Lung infections associated with Cystic Fibrosis." *Clin. Microbiol. Rev.* **15** (2): 194-222.

-M-

Mahenthiralingam, E., Baldwin, A. and Vandamme, P. (2002) "*Burkholderia cepacia* complex infection in patients with cystic fibrosis." *J. Med. Microbiol.* **51**: 533-538.

Mahenthiralingam, E. and Vandamme, P. (2005) "Taxonomy and pathogenesis of the *Burkholderia cepacia* complex." *Chron. Respir. Dis.* **2**:209-217.

Majumder, S. (2006) "Search for fucose binding domains in recently sequenced hypothetical proteins using molecular modeling techniques and structural analysis. " *Glycoconj. J.* **23**: 251-257.

Makrides, S.C. (1996) "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*", *Microbiological Reviews,* **60** (3): 512-538

Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning, A Laboratory Manual. Cold Spring Harbour Laboratory,* CHSL.

Matsudaa, A., Kunoa, A., Ishidab, H., Kawamotod, T., Shodab, J. and Hirabayashi, J. (2008), "Development of an all-in-one technology for glycan profiling targeting formalin-embedded tissue sections." *Biochem. Biophys. Res. Commun.* **370** (2): 259-263.

McCoy, J.P., Varani, J. and Goldstein, I.J. (1983) "Enzyme-Linked Lectin Assay (ELLA): Use of alkaline phosphatase-conjugated *Griffonia simplicifolia* B4 lsolectin for the detection of α-D-galactopyranosyl end groups." *Analytical Biochemistry,* **130**: 437-444.

Mishra, N.K., Kulhanek, P., Snajdrova, L., Martin, P., Imberty, A. and Koca, J. (2008) "Molecular dynamics study of *Pseudomonas aeruginosa* lectin-II complexed with monosaccharides." *Proteins.* **72**: 382-392.

Mitchell, E., Houles, C., Sudakevitz, D., Wimmerova, M., Gautier, C., Perez, S., Wu, A.M., Gilboa-Garber, N. and Imberty, A. (2002) "Structural basis for oligosaccharide-mediated adhesion of *Pseudomonas aeruginosa* in the lungs of cystic fibrosis patients." *Nat. Struct. Biol.* **9** (12): 918-921.

Mitchell, E.P., Sabin, C., Snajdrova, L., Pokorna, M., Perret, S., Gautier, C., Hofr, C., Gilboa-Garber, N., Koca, J., Wimmerova, M. and Imberty, A. (2005) "High Affinity Fucose Binding of *Pseudomas aeruginosa* lectin PA-IIL: 1.0A Resolution

crystal structure of the complex combined with thermodynamics and computational chemistry approaches." *Proteins.* **58**: 735-746.

Morimoto, M., Saimoto, H., Usui, H., Okamoto, Y., Minami, S. and Shigemasa, Y. (2001) "Biological activities of carbohydrate-branched chitosan derivatives." *Biomacromolecules,* **2** (4): 1133-1136.

Mullis, K.B. and Faloona, F.A. (1987) "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." *Meth. Enzymol.* **155**: 335-350.

-N-

Nakamura-Tsuruta, S., Uchiyama, N. and Hirabayashi, J. (2006) "High-throughput analysis of lectin-oligosaccharide interactions by automated Frontal Affinity Chromatography." *Meth. Enzymol.* **415**: 311-325.

Newburg, D.S., Ruiz-Palacios, G.M. and Morrow, A.L. (2005) "Human milk glycans protect infants against enteric pathogens." *Annu. Rev. Nutr.* **25**: 37-58.

Ng, K.K. and Weis, W.I. (1998) "Coupling of prolyl peptide bond isomerization and Ca^{2+} binding in a C-type mannose-binding protein." *Biochemistry,* **37** (51): 17977-17989.

Nicas, T.I. and Iglewski, B.H. (1985) "The contribution of exoproducts to virulence of *Pseudomonas aeruginosa.*" *Can. J. Microbiol.* **31** (4): 387-392.

Nicholas, K.B., Nicholas, H.B., Jr. and Deerfield II, D.W. (1997) "GeneDoc: Analysis and visualization of genetic variation." *European Molecular Biology Network News,* **4**: 1-4.

Nilsson, J., Stahl, S., Lundeberg, J., Uhlén, M. and Nygren, P. (1997) "Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins." *Protein Expr. Purif.* **11** (1): 1-16.

-O-

Ochsner, U.A., Koch, A.K., Fiechter, A. and Reiser, J. (1994) "Isolation and characterization of a regulatory gene affecting rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa.*" *J. Bacteriol.,* **176** (7): 2044-2054

Ochsner, U.A. and Reiser, J. (1995) "Autoinducer-mediated regulation of rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa.*" *PNAS-Biological Sciences,* **92**: 6424-6428.

Ohyama, Y., Kasai, K., Nomoto, H. and Inoue, Y. (1985) "Frontal affinity chromatography of ovalbumin glycoasparagines on a concanavalin A-sepharose column." *J. Biol. Chem.* **260** (10): 6882-6887.

-P-

Passador, L., Cook, J.M., Gambello, M.J., Rust, L. and Iglewski, B.H. (1995) "Expression of *Pseudomonas aeruginosa* virulence genes requires cell-to-cell communication." *Science,* **260** (5111): 1127-1130.

Pavlovskis, O.R. and Gordon, F.B. (1972) "*Pseudomonas aeruginosa* Exotoxin: Effect on cell cultures." *J. Infect. Dis.* **125** (6): 631-636.

Pearson, J.P., Pesci, E.C. and Iglewski, B.H. (1997) "Roles of *Pseudomonas aeruginosa* las and rhl quorum-sensing systems in control of elastase and rhamnolipid biosynthesis genes." *J. Bacteriol.* **179** (18): 5756-5767.

Peel, M.M., Alfredson, D.A., Gerrard, J.G., Davis, J.M., Robson, J.M., McDougall, J.J., Scullie, B.L. and Akhurst, R.J. (1999), "Isolation, identification, and molecular characterization of strains of from infected humans in Australia." *J. Clin. Microbiol.* **37**, (11): 3647-3653.

Perret, S., Sabin, C., Dumon, C., Pokorna, M., Gautier, C., Galanina, O.E., Ilia, S., Bovin, N.V., Nicaise, M., Desmadril, M., Gilboa-Garber, N., Wimmerova, M., Mitchell, E.P. and Imberty, A. (2005) "Structural basis for the interaction of human

milk oligosaccharides and the bacterial lectin PA-IIL of *Pseudomonas aeruginosa.*" *Biochem. J.* **389** (2): 325-332.

Pesci, E.C., Pearson, J.P., Seed, P.C. and Iglewski, B.I. (1997) "Regulation of las and rhl quorum sensing in *Pseudomonas aerugionsa.*" *J. Bacterio.* **179** (10): 3127-3132.

Pilobello, K.T., Krishnamoorthy, L., Slawek, D. and Mahal, L.K. (2005) "Development of a lectin microarray for the rapid analysis of protein glycopatterns." *ChemBioChem,* **6** (6): 985-989.

Pilobello, K.T. and Mahal, L.K. (2007) "Deciphering the glycocode: the complexity and analytical challenge of glycomics." *Curr. Opin. Chem. Bio.* **11** (3): 300-306.

Pokorná, M., Cioci, G., Perret, S., Rebuffet, E., Kostlánová, N., Adam, J., Gilboa-Garber, N., Mitchell, E.P., Imberty, A. and Wimmerová, M. (2006) "Unusual entropy-driven affinity of *Chromobacterium violaceum* lectin CV-IIL toward fucose and mannose." *Biochemistry,* **45** (24): 7501-7510.

Porath, J., Carlsson, J., Olsson, I. and Belfrage, G. (1975) "Metal chelate affinity chromatography, a new approach to protein fractionation." *Nature,* **258** (5536): 598-599.

Preston, M.J., Fleiszig, S.M., Zaidi, T.S., Goldberg, J.B., Shortridge, V.D., Vasil, M.L. and Pier, G.B. (1995) "Rapid and sensitive method for evaluating *Pseudomonas aeruginosa* virulence factors during corneal infections in mice." *Infect. Immun.* **63** (9): 3497-3501.

-R-

Raghavan, M. and Bjorkman, P.J. (1995) "BIAcore: a microchip-based system for analyzing the formation of macromolecular complexes." *Structure,* **3** (4): 331-333.

Ramphal, R., Koo, L., Ishimoto, K.S., Totten, P.A., Lara, J.C. and Lory, S. (1991) "Adhesion of *Pseudomonas aeruginosa* pilin-deficient mutants to mucin." *Infect. Immun.***59** (4): 1307-1311.

Rodda, D.J. and Yamazaki, H. (1994) "Poly (Vinyl Alcohol) as a blocking agent in enzyme immunoassays." *Immunol. Invest.* **23** (6 and 7): 421-428.

Rust, L., Pesci, E.C. and Iglewski, B.H. (1996) "Analysis of the *Pseudomonas aeruginosa* elastase (lasB) regulatory region." *J. Bacteriol.* **178** (4): 1134-1140.

-S-

Sabin, C., Mitchell, E.P., Pokorna, M., Gautier, C., Utille, J.P., Wimmerova, M. and Imberty, A. (2006) "Binding of different monosaccharides by lectin PA-IIL from *Pseudomonas aeruginosa*: Thermodynamics data correlated with X-ray structures." *FEBS Lett.* **580** (3): 982-987.

Saiman, L., Ishimoto, K., Lory, S. and Prince, A. (1990) "The effect of piliation and exoproduct expression on the adherence of *Pseudomonas aeruginosa* to respiratory epithelial monolayers." *J. Infect. Dis.* **161** (3): 541-548.

Saiman, L., Cacalano, G., Gruenert, D. and Prince, A. (1992) "Comparison of adherence of *Pseudomonas aeruginosa* to respiratory epithelial cells from cystic fibrosis patients and healthy subjects." *Infect. Immun.* **60** (7): 2808-2814.

Sanchez, R. and Sali, A. (1998), "Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome." *PNAS.* **95**: 13597-13602.

Schriemer, D.C., Bundle, D.R., Li, L. and Hindsgaul, O. (1999) "Micro-scale frontal affinity chromatography with mass spectrometric dectection: A new method for the screening of compound libraries." *Angew. Chem. Int. Ed.Engl.* **37** (24): 3383-3387.

Schuck, P. (1997) "Use of surface plasmon resonance to probe the equilibrium and dynamic aspects of interactions between biological macromolecules." *Annu. Rev. Biophys. Biomol. Struct.* **26**: 541-566.

Schwarz, F.P., Puri, K.D., Bhat, R.G. and Surolia, A. (1993) "Thermodynamics of monosaccharide binding to concanavalin A, pea (*Pisum sativum*) lectin, and lentil (*Lens culinaris*) lectin." *J. Biol. Chem.* **268** (11): 7668-7677.

Schwede, T., Ko, J., Guex, N. and Peitsch, M.C. (2003) "SWISS-MODEL: an automated protein homology-modeling server." *Nucleic Acids Res.* **31** (13): 3381-3385.

Sharon, N. and Lis, H. (1989), "Lectins as Cell Recognition Molecules." *Science,* **246** (4927): 227-234.

Shi, R., Pan, Q., Guan, Y., Hua, Z., Huang, Y., Zhao, M. & Li, Y. (2007) "Imidazole as a catalyst for in vitro refolding of enhanced green fluorescent protein", *Archives of Biochemistry and Biophysics,* **459** (1): 122-128.

Shinohara, Y., Hasegawa, Y., Kaku, H. and Shibuya, N. (1997) "Elucidation of the mechanism enhancing the avidity of lectin with oligosaccharides on the solid phase surface." *Glycobiology,* **7** (8): 1201-1208.

Sinclair, A.M. and Elliott, S. (2005) "Glycoengineering: The effect of glycosylation on the properties of therapeutic proteins." *J. Pharm. Sci.* **94** (8): 1626-1635.

Sirinavin, S., Techasaensiri, C., Benjaponpitak, S., Pornkul, R. and Vorachit, M. (2005) "Invasive *Chromobacterium violaceum* infection in children: Case report and review." *Pediatr. infect. Dis. J.* **24** (6): 559-561.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J. & Klenk, D.C. (1985) "Measurement of protein using bicinchoninic acid", *Analytical Biochemistry,* **150** (1): 76-85.

Sonawane, A., Jyot, J. and Ramphal, R. (2006) "*Pseudomonas aeruginosa* LecB is involved in pilus biogenesis and protease IV activity but not in adhesion to respiratory mucins." *Infect. Immun.* **74** (12): 7035-7039.

Steinitz, M. (2000) "Quantitation of the blocking effect of Tween 20 and bovine serum albumin in ELISA microwells." *Anal. Biochem.* **282**: 232-238.

Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warrener, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Garber, R.L., Goltry, L., Tolentino, E., Westbrock-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N.,

Folger, K.R., Kas, A., Larbig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Saier, M.H., Hancock, R.E., Lory, S. & Olson, M.V. (2000) "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen", *Nature,* **406** (6799): 959-964.

Sudakevitz, D., Kostlanova, N., Blatman-Jan, G., Mitchell, E.P., Lerrer, B., Wimmerova, M., Katcoff, D.J., Imberty, A. and Gilboa-Garber, N. (2004) "A new *Ralstonia solanacearum* high-affinity mannose-binding lectin RS-IIL structurally resembling the *Pseudomonas aeruginosa* fucose-specific lectin PA-IIL." *Mol. Microbiol.* **52** (3): 691-700.

Surolia, A., Sharon, N. and Schwarz, F.P. (1996) "Thermodynamics of monosaccharide and disaccharide binding to *Erythrina corallodendron* lectin." *J. Biol. Chem.* **271** (30): 17697-17703.

-T-

Tielker, D., Hacker, S., Loris, R., Strathmann, M., Wingender, J., Wilhelm, S., Rosenau, F. and Jaeger, K.E. (2005), "*Pseudomonas aeruginosa* lectin LecB is located in the outer membrane and is involved in biofilm formation." *Microbiology,* **151** (5): 1313-1323.

Tielker, D. (2006) "Lectin-based affinity tag for one-step protein purification." *Biotechniques,* **41** (Sep): 327-332.

Tomana, M., Schrohenloher, R.E., Bennett, P.H., del Puente, A. and Koopman, W.J. (1994) "Occurrence of deficient galactosylation of serum IgG prior to the onset of rheumatoid arthritis." *Rheumatol. Int.* **13** (6): 217-220.

Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D. and Arakawa, T. (2003) "Solubilization of active green fluorescent protein from insoluble particles by guanidine and arginine." *Biochem. Biophys. Res. Commun.* **312** (4): 1383-1386.

Turbadar, T. (1959) "Complete absorption of light by thin metal films." *Proc. Phys. Soc.* **73** (1): 40-44.

-V-

van der Merwe, P.A. and Barclayt, A.N. (1997) "Analysis of cell-adhesion molecule interactions using surface **C**plasmon resonance." *Curr. Opin. Immun.* **8** (2): 257-261.

von Bismarck, P., Schne:enheim, R. and Schumacher, U. (2001) "Successful treatment of *Pseudomonas aeruginosa* respiratory tract infection with a sugar solution-a case report on a lectin based therapeutic principle." *Klin. Padiatr.* **213** (5): 285-287.

Vornholt, W., Hartmann, M. and Keusgen, M. (2007) "SPR studies of carbohydrate–lectin interactions as useful tool for screening on lectin sources." *Biosens. Bioelectron.* **22** (12): 2983-2988.

-W-

Wada, K., Wada, Y., Ishibashi, F., Gojobori, T. and Ikemura, T. (1992) "Codon usage tabulated from the GenBank genetic sequence data." *Nucleic Acids Res.* **20**: 2111-2118.

Waldo, G.G., Standish, B.M., Berendzen, J. and Terwilliger, T.C. (2004) "Rapid protein-folding assay using green fluorescent protein." *Nat. Biotechnol.* **34** (3): 354-363.

Wang, W. (2005) "Protein aggregation and its inhibition in biopharmaceutics." *Int. J. Pharm.* **289** (1-2): 1-30.

Wear, M.A. and Walkinshaw, M.D. (2006) "Thermodynamics of the cyclophilin-A/cyclosporin-A interaction: A direct comparison of parameters determined by surface plasmon resonance using Biacore T100 and isothermal titration calorimetry." *Anal. Biochem.* **359** (2): 285-287.

Weissman, S.J., Moseley, S.L., Dykhuizen, D.E. and Sokurenko, E.V. (2003) "Enterobacterial adhesins and the case for studying SNPs in bacteria." *Trends Microbiol*. **11** (3): 115-117.

Whitaker, J.R. (1963) "Determination of Molecular Weights of Proteins by Gel Filtration on Sephadex." *Anal. Chem.* **35** (12): 1950-1953.

Whitchurch, C.B., Leech, A.J., Young, M.D., Kennedy, D., Sargent, J.L., Bertrand, J.J., Semmler, A.B.T., Mellick, A.S., Martin, P.R., Alm, R.A., Hobbs, M., Beatson, S.A., Huang, B., Nguyen, L., Commolli, J.C., Engel, J.N., Darzins, A. and Mattick, S. (2004) "Characterization of a complex chemosensory signal transduction system which controls twitching motility in *Pseudomonas aeruginosa*." *Mol. Microbio.* **52** (3): 873-893.

Whiteley, M., Lee, K.M. and Greenberg, E.P. (1999) "Identification of genes controlled by quorum sensing in *Pseudomonas aeruginosa.*" *PNAS-Biological Sciences,* **96** (24): 13904-13909.

Williams, P., Baintona, N.J., Swift, S., Chhabraa, S.R., Winsonb, M.K., Stewartb, G.S.A.B., Salmond, G.P.C. and Bycrofta, B.W. (1992), "Small molecule-mediated density-dependent control of gene expression in prokaryotes: Bioluminescence and the biosynthesis of carbapenem antibiotics." *FEMS Microbiol. Lett.* **100** (1-3): 161-167.

Wimmerová, M., Mishra, N.K., Pokorná, M. and Koča, J. (2009) "Importance of oligomerisation on *Pseudomonas aeruginosa* Lectin-II binding affinity. *In silico* and *in vitro* mutagensis." J. *Mol. Mod.* **54** (6): 673-679.

Winzer, K., Falconer, C., Garber, N.C., Diggle, S.P., Camara, M. and Williams, P. (2000) "The *Pseudomonas aeruginosa* lectins PA-IL and PA-IIL are controlled by Quorum Sensing and by RpoS." *J. Bacteriol.* **182** (22): 6401-6411.

Wu, A.M., Wu, J.H., Singh, T., Liu, J.H., Tsai, M.S. and Gilboa-Garber, N. (2006) "Interactions of the fucose-specific *Pseudomonas aeruginosa* lectin, PA-IIL, with mammalian glycoconjugates bearing polyvalent Lewisa and ABH blood group glycotopes." *Biochimie,* **88** (10): 1479-1492.

Wu, A.M., Lisowska, E., Duk, M. and Yang, Z. (2008) "Lectins as tools in glycoconjugate research." *Glycoconj. J.* Epub ahead of print.

-X-

Xia, B., Royall, J.A., Damera, G., Sachdev, G.P. and Cummings, R.D. (2005) "Altered O-glycosylation and sulfation of airway mucins associated with cystic fibrosis." *Glycobiology,* **15** (8): 747-775.

-Y-

Yahr, T.L., Goranson, J. and Frank, D.W. (1997) "Exoenzyme S of *Pseudomonas aeruginosa* is secreated by a type III pathway." *Mol. Microbiol.* **22** (5): 991-1003.

Yang, Z. and Hancock, W.S. (2005) "Monitoring glycosylation pattern changes of glycoproteins using multi-lectin affinity chromatography." *J. Chromatogr. A,* **1070** (1-2): 57-64.

-Z-

Zinger-Yosovich, K., Sudakevitz, D., Imberty, A., Garber, N.C. and Gilboa-Garber, N. (2006) "Production and properties of the native *Chromobacterium violaceum* fucose-binding lectin (CV-IIL) compared to homologous lectins of *Pseudomonas aeruginosa* (PA-IIL) and *Ralstonia solanacearum* (RS-IIL)." *Microbiology,* **152** (2): 457-463.

Zuber, C., Li, W. and Roth, J. (1998) "Blot analysis with Lectins for the Evaluation of Glycoproteins in Cultured Cells and Tissues." in *Use of Lectins for Structural Analysis of Oligosaccharide Chains*, eds. J.M. Rhodes and J.D. Milton, : 159-166.