

**The Characterisation of the Peanut Agglutinin: an Evolved  
plant lectin, with improved specificity to the Thompson**

**Freidenriech antigen.**

**by**

ZAIDA LAGARDIEN

lgrzai001

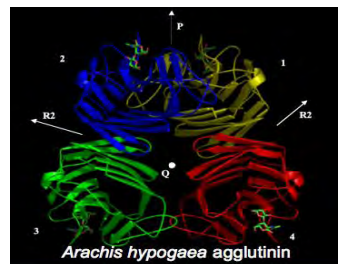
SUBMITTED TO THE UNIVERSITY OF CAPE TOWN

In fulfillment of the requirements for the degree

Msc. Medical Biochemistry

**Faculty of Health Sciences**

**UNIVERSITY OF CAPE TOWN**



**April 2013**

**Professor J.M Blackburn**

**Department of Medical Biochemistry: University of Cape Town.**

# Contents

<b>Declaration.....</b>	<b>I</b>
<b>Acknowledgements.....</b>	<b>II</b>
<b>Abbreviations.....</b>	<b>III</b>
<b>Abstract.....</b>	<b>V</b>
<b>Chapter 1: Introduction.....</b>	<b>1</b>
<b>1.1 The significance of protein carbohydrate interactions.....</b>	<b>1</b>
<b>1.1.1 History of lectins as diagnostic tools.....</b>	<b>2</b>
<b>1.1.2 The study of carbohydrates and proteins and their interactions.....</b>	<b>4</b>
<b>1.2 Types of lectins.....</b>	<b>5</b>
<b>1.2.1 Legume Lectin structure and function.....</b>	<b>7</b>
<b>1.2.2 Common recognition folds of the lectins.....</b>	<b>10</b>
<b>1.3 The Peanut Agglutinin (PNA) and T-antigen.....</b>	<b>12</b>
<b>1.3.1 Structure and molecular recognition properties of the PNA.....</b>	<b>12</b>
<b>1.3.2 Thomsen Friedenreich antigen (Gal<math>\beta</math>1-3 GalNAc<math>\alpha</math><sub>1</sub>-O-Ser/Thr).....</b>	<b>15</b>
<b>1.4 Engineering of plant lectins to improve specificity.....</b>	<b>21</b>
<b>1.5 Aim of thesis.....</b>	<b>23</b>
<b>Chapter 2: Materials and Materials.....</b>	<b>24</b>
<b>2.1 Materials.....</b>	<b>24</b>
<b>2.2 Methods.....</b>	<b>34</b>
<b>2.2.1 DNA characterization and analysis.....</b>	<b>34</b>
<b>2.2.2 Expression and purification of recombinant proteins.....</b>	<b>36</b>
<b>2.2.3 Native gel electrophoresis.....</b>	<b>40</b>
<b>2.2.4 Examination of quaternary structure of proteins.....</b>	<b>40</b>
<b>2.2.5 High Performance Liquid Chromatography (HPLC).....</b>	<b>41</b>

2.2.6 Affinity binding assay.....	43
2.2.7 Microarray assay –Protein-Carbohydrate binding assay.....	47
2.2.8 ELISA Enzyme-linked Immunosorbent assay.....	52
<b>Chapter 3: Results and discussion.....</b>	<b>55</b>
3.1 DNA preparation.....	55
3.1.1 DNA extraction and quantification.....	55
3.1.2 Plasmid sequences from DNA extractions for protein expression and purification.....	55
3.1.3 Restriction digest.....	58
3.1.4 Discussion: restriction digests.....	59
3.1.5 Sequence analysis.....	59
3.2 Protein expression and purification.....	64
3.2.1 Protein expression and purification analysis using SDS page.....	64
3.2.2 Expression of recombinant PNA variants.....	65
3.3 Purification of recombinant PNA variants.....	66
3.3.1 Discussion of Eluted samples of purification.....	66
3.4 Western blot analysis.....	68
<b>Chapter 4: Investigation of quaternary structure of Peanut Agglutinin.....</b>	<b>70</b>
4.1 Native gel electrophoresis/mobility shift assay.....	70
4.1.1 Native vs denaturing gel electrophoresis.....	70
4.2 Investigation of multimeric states and quaternary structure using Mass Spectrometry and High Performance Liquid Chromatography.....	72
4.2.1 MALDI Spectrometry.....	72
4.3 High Performance Liquid Chromatography (HPLC) analysis of PNA quaternary structure.....	77
<b>Chapter 5: Affinity binding assays.....</b>	<b>90</b>
5.1 Streptavidin coating of slides.....	90
5.2 Antibody labeling.....	91

<b>5.3 Microarray printing of NHS-activated, PEG-coated glass.....</b>	<b>91</b>
<b>5.4 ELISA enzyme linked immunosorbent assay.....</b>	<b>92</b>
<b>5.5 Discussion.....</b>	<b>97</b>
<b>5.6 Conclusion.....</b>	<b>98</b>
<b>References.....</b>	<b>100</b>
<b>Appendix.....</b>	<b>107</b>

## **Declaration**

I, Zaida Lagardien, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature .....

Date.....

## **Acknowledgements**

I would like to thank my supervisor Prof. Jonathan Blackburn for his continuous support, patience and motivation throughout the production of this thesis and my Masters project. I would also like to thank Dr. Natasha Beeton Kempen who taught me a great deal.

I would like to extend a thank you to my husband for the continuous support and understanding during this time. I am ever grateful to my family in particular my mother, for the constant motivation and inspiration throughout my studies. To my friends, for helping maintain my sanity, laughter and good spirit through my postgraduate career.

Last but not least, I would like to thank the University of Cape Town for affording me this opportunity to study at this prestigious University, and the National Research foundation for the funding provided through the course of my studies.

## Abbreviations

AB	Antibody
BSA	Bovine Serum Albumin
bp	base pairs
Con A	Concanavalin A
D/P	Dye to protein ratio
ELISA	Enzyme linked immunosorbant assay
EtBr	Ethidium Bromide
His	Histidine
HPLC	High performance liquid chromatography
HRP	Horse radish peroxidase
kDa	Kilo Dalton
Lac	Lactose
LacNAc	N-Acetyl-D-lactosamine
MALDI	Matrix-assisted laser desorption/ionization
MALDI-Tof	Matrix-assisted laser desorption/ionisation- <i>time of flight</i> mass spectrometry
NHS	N-hydroxysuccinimide
Ni-NTA	Nickel Nitrilotriacetic acid
OD	Optical density
PAGE	polyacrylamide gel electrophoresis
PNA	Peanut agglutinin ( <i>Arachis Hypogaea</i> )
PBS	Phosphate buffer saline

PBST	Phosphate buffer saline Tween
PEG	Polyethylene glycol
SPR	Surface Plasmon Resonance
SEC	Size Exclusion Chromatography
SDS	sodium dodecyl sulphate
T-antigen	Thomsen Freidenriech antigen (Gal $\beta$ 1-3 GalNAc $\alpha$ <sub>1</sub> -O-Ser/Thr)

## Abstract

The study of Lectins-specifically *Arachis Hypogaea*, the Peanut Agglutinin- aims to develop the use and application of lectins as diagnostic tools. Peanut agglutinin (PNA), a carbohydrate binding protein, is able to recognise and bind a number of distinct carbohydrate structures that have been implicated in a number of disease pathologies in humans. *In vitro* studies of PNA have previously been shown to have some specificity for the Thomson Freidenriech antigen (T-antigen), found on malignant human cells, and this specificity has made PNA an important target for protein engineering experiments aimed at improving its specificity and affinity. A number of tumour cells are characterised by altered states and patterns of glycosylation on cell surfaces and suitably engineered lectins may be able to recognise tumour specific carbohydrate structures.

This study was aimed at carrying out the biophysical characterisation of a set of PNA mutants which showed apparent improvement in specificity for the T-Antigen. Previous studies have aimed to engineer this lectin in order to direct its recognition properties towards the T-antigen and away from lactose, the preliminary binding affinities of these mutants being determined using Surface Plasmon Resonance (SPR). Here a set of PNA mutants were characterised, proteins expressed and purified to determine binding activities to the T-antigen, N-Acetyl-D-lactosamine (LacNAc) and lactose through the use of Protein Micro Array technology as well as Enzyme linked immunosorbant assays (ELISA).

In addition, the current study aimed to determine the multimeric state of PNA in solution. Previous crystallography data suggests evidence of a tetrameric species, but the present studies demonstrate the predominance of dimeric species instead of tetrameric species. Finally, the present study also demonstrates a pH dependence of the multimeric state of PNA, a result that is rationalised in terms of crystallographic data.

## Chapter 1: Introduction

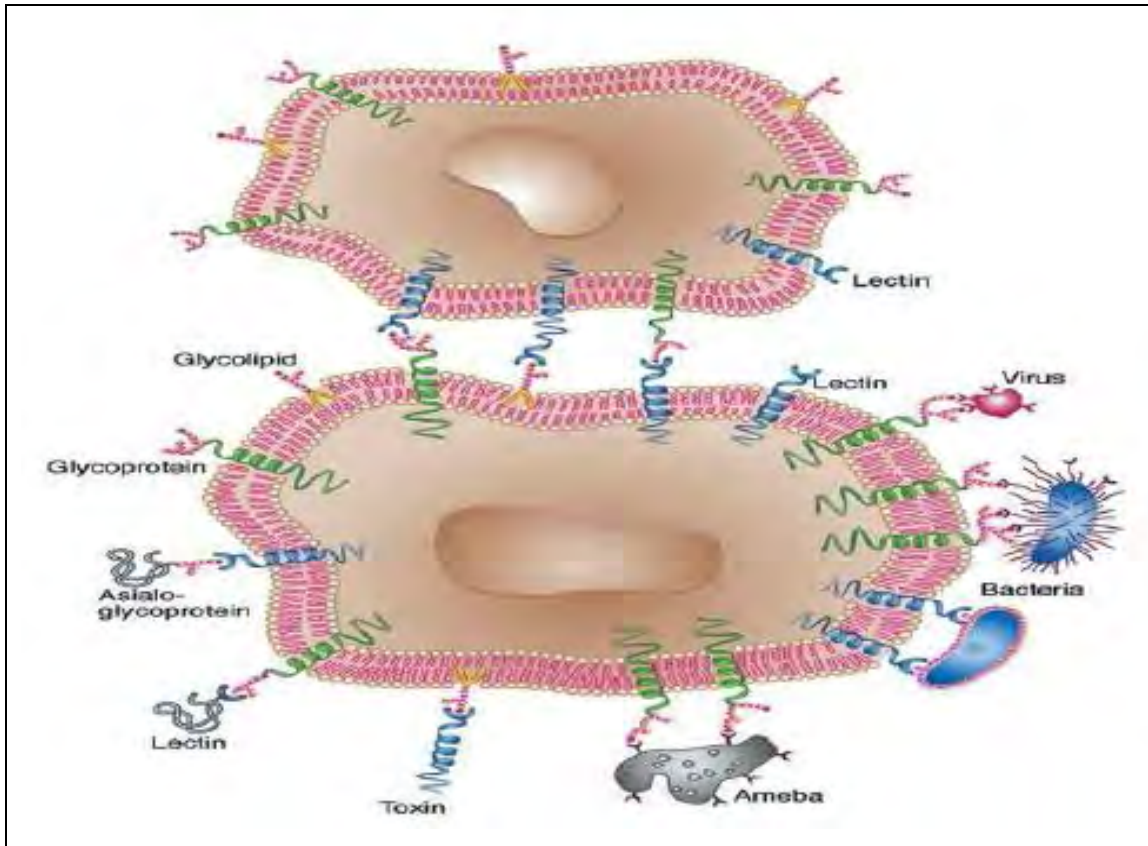
### 1.1 The significance of protein carbohydrate interactions.

Proteins are among the primary constituents of living organisms and form an important class of molecules studied in biochemistry/biological sciences. Proteins mediate many interactions on a chemical and biochemical level and are able to play structural, mechanical and catalytic roles thus providing most of the molecular machinery of cells (Liang *et al.*, 2007) within a cellular environment, many proteins, interact with other proteins or with other molecules to form complexes. Of particular interest in the context of this thesis is the ability of certain proteins to bind to carbohydrate structures. These carbohydrate binding proteins have been shown to be excellent tools for monitoring changes on cell surfaces during physiological, pathological conditions and cell surface glycosylation patterns (Liang *et al.*, 2007).

Carbohydrates are a family of molecules of great structural and functional diversity. They constitute a significant fraction of earth's organic matter. Proteins that interact with carbohydrate structures non-covalently are widespread in nature: these include carbohydrate specific enzymes as those found in metabolism, carbohydrate binding proteins, as well as anti-carbohydrate antibodies (Varki *et al.*, 1993). In human biology, carbohydrate and protein carbohydrate interactions are typically essential in bacterial and viral infection, differentiation and development, immune response and for the progression of metastasis of tumor cells (Liang *et al.*, 2007).

One of the most well studied interactions between proteins and carbohydrates are between the lectin group of proteins and carbohydrates, as they bind to and recognize cell surfaces (Figure 1.1). A study by Sharon 1989 showed that lectins were able to agglutinate red blood cells giving rise to increased interest in this class of proteins in

Latin referred to as ‘*legere*’ meaning to select. What makes lectins of increased interest is because they are found ubiquitously in animals, plants, bacteria etc., making them easy and cheap to access (Slifkin & Doyle, 1990).



**Figure 1.1** Cell surface lectin–carbohydrate interactions. Lectins serve as means of attachment of different kinds of cell as well as viruses to other cells via the surface carbohydrates of the latter (Sharon & Lis, 2004).

### 1.1.1 History of lectins as diagnostic tools

Lectin research dates back to the late 1900’s and throughout the 20<sup>th</sup> century interest in the study of this class of proteins and their function has grown continually. The discovery of an agglutinin named ricin (*Ricin communis*) in the seeds from the castor

tree in 1888 by Peter Hermann Stillmark, and also subsequently the discovery of another agglutinin named abrin extracted from the Jequirity bean, by H.Hellin, is where the interest in the use of lectins as agglutinating agents could be said to originate from (Sharon & Lis, 2004).

These toxic agglutinins were employed by scientist Paul Ehrlich from the Royal Institute of Experimental Therapy as model antigens in immunological research in mice to determine immunity and specificity of the immune response (Sharon & Lis, 2004). Due to their toxicity, ricin and abrin were later used for biological warfare and intended for use as weapons for mass destruction during World War I and in World War II, a ricin bomb was developed and tested but never employed as a weapon of mass destruction (Sharon & Lis, 2004).

One particularly intriguing lectin is Concanavalin A (Con A), from the Jack Bean; a pure form of this hemagglutinin was obtained by James Sumner in 1919 at Cornell University and in 1936, it was discovered that Con A was not only able to agglutinate erythrocytes and yeasts but could also precipitate glycogen from solution (Sharon & Lis, 2004), both activities being inhibited by sucrose. It was at this point that the idea of lectin sugar specificity came into play and it was suggested that Con A-induced hemagglutination was as a consequence of binding between the lectin and carbohydrates found on the surface of red blood cells (Sharon & Lis, 2004).

It was during the 1940s that the scientists William C. Boyd and Karl O. Renkonen from Boston University and University of Helsinki respectively, made the discovery for the specificity of hemagglutinins for human blood groups/types (Sharon & Lis, 2004). It was found that crude extracts of the *Phaseolus limensis*, Limabean, and the tufted vetch, *Vicia (V) cracca*, agglutinated blood type A erythrocytes but not blood type B or O cells, but an extract of the asparagus pea, *Lotus tetragonolobus*, was able to specifically agglutinate blood type O erythrocytes (Sharon & Lis, 2004). Additional hemagglutinins specific for blood types A and O (but not B) have also been discovered, as well as several for other blood types, such as N (*V. graminea* lectin), T

(peanut agglutinin, PNA) and Tn (the lectins of *V. villosa* and *Moluccella laevis*). (Sharon & Lis, 2004).

These blood type-specific hemagglutinins played a pivotal role in early investigations on the structural basis of the specificity of the antigens involved with the ABO blood group system. During the 1950s, Winifred M. Watkins and Walter J. T. Morgan at the Lister Institute, London, discovered that the agglutination of type 'A' red cells by lima bean lectin was best inhibited by  $\alpha$ -linked N-acetyl-D-galactosamine, and type O cells by the lectin of *L. tetragonolobus*, was best inhibited by  $\alpha$ -linked L-fucose. It was concluded that  $\alpha$ -N-acetyl-D-galactosamine and  $\alpha$ -L-fucose, were the sugar determinants conferring A and H (O) blood group specificity, respectively (Morgan & Watkins, 2000).

The work of Watkins and Morgan provided early evidence for the existence of carbohydrates on cell surfaces where they could play a role as identity markers-this is now an accepted theme in modern glycobiology. The potential for plant agglutinins to distinguish between erythrocytes of different blood types allowed Boyd and Shapleigh (1954) to propose the name lectins, meaning 'to pick out' or 'choose'. This name was later generalized to include all sugar-specific agglutinins of none immune origin, irrespective of source and blood type specificity (Sharon & Lis, 1972).

### **1.1.2 The study of carbohydrates and proteins and their interactions**

Protein-carbohydrate interactions have generated much interest over the past few years due to their great structural and functional diversity and the roles they play as energy stores/sources and as protein recognition molecules that mediate cellular adhesion and targeting (Varki *et al.*, 1993). Glycoconjugates, glycoproteins and glycolipids all mediate interesting and valuable interactions such of which only exhibit functionality or activity once bound by lectins.

Lectins are typically classified according to their carbohydrate specificity and are often highly specific, multivalent carbohydrate binding proteins (Salunke *et al.*, 1985). While some lectins are specific to only one sugar, others are able to bind to more than one sugar. Over the last 120 years lectin research has come a long way. Much more could be elaborated on the history of these proteins, but for this study, lectins as diagnostic tools and their structural characteristics will now be addressed.

## 1.2 Types of lectins

Three interesting lectin types have been found with the following classifications: Merolectins which possess only one carbohydrate binding site or domain. These are unable to agglutinate cells, as it is required that a lectin possess at least two domains to enable it to stably agglutinate a cell; hololectins, which possess at least two carbohydrate binding domains, which can either be identical or highly homologous and thereby are able to agglutinate red blood cells; and chimerolectins, possessing a carbohydrate binding domain and another unrelated domain that is unable to bind carbohydrates but which exhibits defined catalytic function, independent of the carbohydrate binding domain.

These chimero lectins can act either as homolectins or merolectins, depending on the number of binding sites they possess (Van Damme *et al.*, 1998). Lectins are typically classified into four groups, depending on their ability to bind to Glucose, Galactose/ N-acetylamine acetylgalactose, L-fucose, or the sialic acids (Table 1.1). The experimental classification by which this specificity was observed was by means of firstly, observing the protein's ability to agglutinate red blood cells and then the monosaccharide specificity by observing which monosaccharide's blocked this agglutination. The individual lectins show much greater specificity at the disaccharide and oligosaccharide levels.

**Table 1.1** Legume lectins. Examples of legume lectins with structural data, sugar specificity and oligomeric states (Sharpe, 2003).

<b>Legume</b>	<b>Common name</b>	<b>Abbreviation</b>	<b>Oligomeric state</b>	<b>Sugar specificity</b>	<b>PDB Code</b>
<i>Canavalia brasiliensis</i>	Brazilian jack bean	ConBr	Tetramer	Man/Glc	1AZD
<i>Canavalia ensiformis</i>	Jack bean	ConA	Tetramer	Man/Glc	2CNA
<i>Dioclea grandiflora</i>	Mucuna	DGL	Tetramer	Man/Glc	1DGL
<i>Dolichos biflorus</i>	Horse gram	DBL	Tetramer	GalNAc	1LUI
<i>Dolichos biflorus</i>	Horse gram	DB58	Tetramer	GalNAc	IG7Y
<i>Dolichos lablab</i>	Hyacinth pea	FRIL	Dimer	Man/Glc	1QMO
<i>Erythrina corollodendron</i>	West Indian coral tree	EcoL	Dimer	Gal/GalNAc	1AXO
<i>Erythrina crystallogali</i>	Cocks comb, coral tree	ECL	Dimer	GalNAc	1GZ9
<i>Vicia fabia</i>	Broad bean	Favin	Tetramer	Man/Glc	ND6
<i>Griffonia simplicifolia</i>	Griffonia	GS-1	Dimer	Gal/GalNAc	1HQL
<i>Griffonia simplicifolia</i>	Griffonia	GS-4	Dimer	Gal/GalNAc	1GSL
<i>Lathyrus ochrus</i>	Yellow-flowered pea	LOL-1	Dimer	Man/Glc	1LOE
<i>Maackia amurensis</i>	Amur maackia	MAL	Tetramer	Complex	1DBN
<i>Pisum sativum</i>	Pea	PSL	Dimer	Man/Glc	2LTN
<i>Arachis hypogaea</i>	Peanut	PNA	Tetramer	Gal	2TEP
<i>Phaseolus vulgaris</i>	Kidney bean	PHA-L	Tetramer	Complex	1FAT
<i>Robinia pseudoacacia</i>	Bark lectin I	1FNY	Tetramer	Complex	1FNY
<i>Glycine max</i>	Soybean	SBA	Tetramer	Gal/GalNAc	2SBA
<i>Ulex europaeus</i>	Common gorse	UEA-1	Tetramer	Fucose	1FXS
<i>Vicia villosa</i>	Hairy winter vetch	VVL-B4	Tetramer	Gal/GalNAc	1N47
<i>Psophocarpus tetragonolobus</i>	Winged bean agglutinin	WBA1	Dimer	Gal/GalNAc	1WBL
<i>Psophocarpus tetragonolobus</i>	Winged bean agglutinin	WBA2	Dimer	Gal/GalNAc	1FAY

However, to determine the biophysical function, the interactions between lectins and carbohydrates required more extensive study than just defining their ability to agglutinate erythrocytes or act as immune response agents. Crystallographic structural studies have thus provided answers to some insight into this class of protein (Kumar *et al.*, 2012).

### **1.2.1 Legume lectin structure and function.**

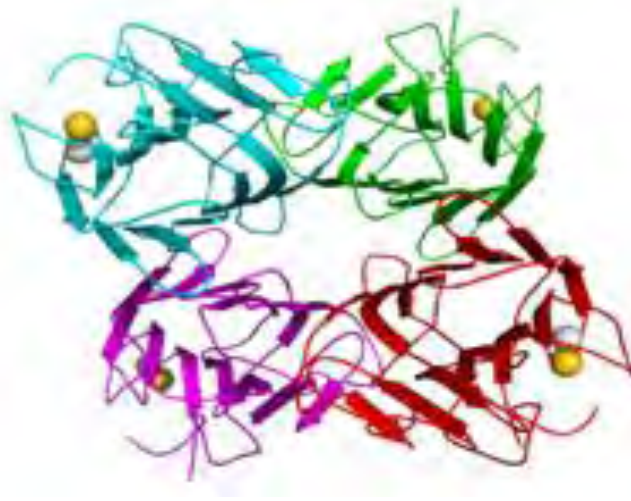
Lectins largely consist of carbohydrate binding proteins or glycoproteins of non immune origin. They are able to bind cells or precipitates and on occasion both (Kumar *et al.*, 2012). Carbohydrate associated enzymes with multiple combining sites will allow precipitation or agglutination of the glycoconjugates and may therefore also be labelled as lectins (Kumar *et al.*, 2012).

The specific binding capacity of the lectins is typically defined by the monosaccharides, which when present are able to inhibit the lectins associated reactions. Lectins are infact usually able to accommodate large sugars and their multivalency allows for stronger interactions with branched oligosaccharides (Kumar *et al.*, 2012). Lectins do not typically exhibit strong cross-reactivity with protein elements of glycoprotein ligands so they are able to afford greater specificity for the carbohydrate structures in glycoproteins than antibodies are able to (Kumar *et al.*, 2012). Lectin histochemistry is based on the ability of the plant lectins, and in particular those from leguminosae, to discriminate between large numbers of structurally diverse cell surface sugars (Kumar *et al.*, 2012).

The legume family of lectins are the most widely studied group of lectins and have been purified from more than 70 species of legumes (Sharon and Lis 1990, Van Damme *et al.*, 1998b). Legume lectins are predominantly specific for Gal/GalNAc or Man/Glc monosaccharides but there exists the few that are specific for fucose (Sharon & Lis 1990). Legume lectins are typically ~ 250 amino acid residues in length and



Figure 1.2 displays the completely conserved and the partially conserved sequences across the legume lectin family. The amino acids in bold black are those that are completely conserved and the grey highlighted amino acids, indicate partially conserved. Legume lectins are synthesised in the ER as pre-proteins, and after removal of the N-terminal signal sequence, further post-translational processing may occur. This may be through cleavage of C-terminal residues, removal of covalently bound carbohydrate, proteolytic cleavage (for example leading to the formation of two chain lectins of the *Viciae* tribe) or cleavage followed by re-ligation of the N- and C-termini (for example producing the circularly permuted ConA) (Loris *et al.*, 1998). Legume lectin monomers have a conserved tertiary structure, the prototypical example of which is the Jack Bean *Canavalia Ensiformis* lectin, Con A (Figure 1.3).



**Figure 1.3** Concanavalin A tetrameric structure. Each colour differentiates a monomeric form of the tetramer of Concanavalin A. These monomers are identical to each other and are an indication of the conserved tertiary structure (Swiss PDB viewer).

This is similar to the fold observed in viral coated proteins (Chelvanayagam *et al.*, 1992), where it is essentially a scaffold of 3  $\beta$ -pleated sheets where the carbohydrate binding site is grafted onto.

The classic legume lectin fold consists of a flat 6-stranded ‘back’ sheet, a curved 7-stranded ‘front’ sheet, and 5-stranded ‘top’ sheet which is vital in keeping the two larger sheets together. Similar to the jellyroll fold seen in viral coat proteins, this fold is not in fact restricted to the legume lectins but is found in 14 other families despite low sequence homology (Chandra *et al.*, 2001). More than half of the lectin structure consists of loops and two hydrophobic cores, with the first formed between the back and front sheets and the second formed between the front sheet and loops curling over the front sheet (Loris *et al.*, 1998).

The legume lectins show much variation in quaternary structure despite displaying great homology in tertiary structure. It is at the dimer interface that the main differences are evident with lectins, as they exist as dimers or dimers of dimers. (Covell *et al.*, 1994). The extensive variation in quaternary structure appears to be due to small differences in tertiary structure and this phenomenon is evident in other protein families, such as the galectins,  $\alpha$  and  $\beta$  chemokines, and cysteine knot growth factors (Covell *et al.*, 1994).

### **1.2.2 Common recognition folds of the lectins**

Legume lectins display a highly conserved monosaccharide binding site consisting of a narrow cleft in the protein surface containing four loops termed A, B, C, and D which contain certain residues including an Asp (Loop A) an Asn (Loop C) and the main chain NH of a Gly or Arg (Loop B) (see Figure 1.4). The primary specificity to the monosaccharide is formed by hydrogen bonding interactions formed by a triad of the abovementioned invariant residues.

Figure 1.4 presents a legume lectin, peanut agglutinin (PNA) with Loops A, B, C, and D. The bound sugar (lactose) is shown as a “ball-and-stick” image. Important calcium and manganese ions are required for ligand binding. The size of binding-site loop d

and monosaccharide specificity show an explicit correlation. Key residues are highlighted in blue and highly conserved residues and been indicated with an asterisk (Vijayan *et al.*, 1999).



Lectin	Loop A	Loop B	Loop C	Loop D	Specificity
EcorL	GPPYT-RPLPADGLVF	AQ-GYGYLG	VEFDTFSN----PWDP	GLSGATG----AQRDAAETHDVYSW	GalNAc
DBL	APSK---ASFADGIAP	RR-NCGYLG	VEFDTLSNS---GWDP	GFSATTGLSDG----YIETHDVLSW	GalNAc
E-PHA	VPNN---EGPADGLAF	KD-KGGLLG	VEFDTLYNV---HWDP	GFTATTGITKG----NVETNDILSW	Complex
L-PHA	VPNN---AGPADGLAF	KD-KGGFLC	VEFDLYNK---DWDP	GFSATTGINKG----NVETNDVLSW	Complex
SBA	APDT---KRLADGLAF	QT-HAGYLG	VEFDTFRN----SWDP	GFSAAATGLDIP-----GESHDVLSW	GalNAc
PNA	KD--IKDYDPADGIIF	GSIGGGTLG	VEFDTYSNS--EYNDP	GFSASGSL-----GGRQIHLIRSW	Gal
LTA	IR--ELKYTPTDGLVF	GS-TGGFLG	VEFDSYHN----IWDP	GFSATTGN-----PEREKHDIYSW	Fuc
UEA-I	SANP---KAATDGLTF	RRA-GGYFG	VEFDTY-GSPVNFDDP	GFSGGTYI-----GRQATHEVLNWSW	Fuc
UEA-II	EPDE--KIDGVDGLAF	GS-SAGMFG	VEFDSYFGKAYNPWDP	GFSGGVGN-----AAKFDHDVLSW	GlcNAc
LAAl	PPIQSRKADGVDGLAF	GS-SAGMFG	VEFDTYFGKAYNPWDP	GFSAGVGN-----AAKFNHDILSW	GlcNAc
LSL	RPNSDS-QVVADGFTF	RG-DGGLLG	VEFDTFHNQ---PWDP	GLSASTATYY-----SAHEVYSW	Man/Glc
Con A	SPDS---HPADGIAP	GS-TCRLLG	VELDTYPNT--DIGDP	GLSASTGL-----YKETNTILSW	Man/Glc
DiocL	SPDH---EPADGITF	GS-GGRLLG	VELDSYPNT--DIGDP	GLSATTGL-----YKETNTILSW	Man/Glc
LenL	SPNG---YNVADGFTF	QT-GCGYLG	VEFDTFYNA---AWDP	GFSATTGAEP-----AAQEVHSW	Man/Glc
PSL	APNS---YNVADGFTF	QT-GCGYLG	VEFDTFYNA---AWDP	GFSATTGAEY-----AAHEVLSW	Man/Glc
FAVIN	APNG---YNVADGFTF	QT-GCGYLG	VEFDTFYNA---AWDP	GFSATTGAEY-----ATHEVLSW	Man/Glc

**Figure 1.4** Loops involved in sugar binding. This shows the three-dimensional structure of a legume lectin, peanut agglutinin (PNA) monomer showing the four loops involved in sugar binding (Sharpe, 2003).

A *cis* peptide bond is located between the conserved asparagine in loop A and the previous residue which is typically an alanine, and serves to orientate the asparagine correctly. Stabilisation of the bound sugar is generated by stacking interactions with a Leu, Tyr, Trp, or Phe residues. Most legume lectin specificity towards Man/Glc or Gal/GalNAc, may be explained largely by the orientation of the 3' and 4' hydroxyl groups in the ligand. Ligands are only bound if their orientation is such that an equatorial OH (3'OH in Gal/GalNAc or 4'OH in Man/Glc) (Sharpe, 2003), is directed towards the centre of the triad.

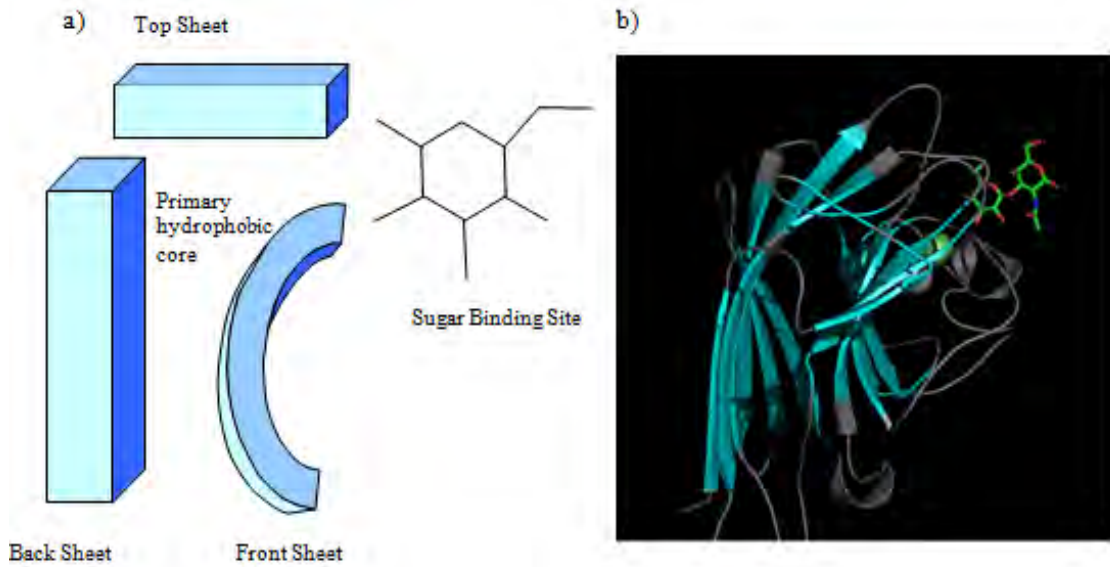
An axial OH is required to establish contacts with the Asp from loop A. The binding site is incapable of accepting two equatorial hydroxyl groups facing the triad, leading to a difference in the orientation of Man/Glc and Gal/GalNAc in order to make the required binding contacts (Sharpe, 2003). The specificity observed between mannose and galactose binding lectins is thought to be generated by Loop D, which is highly variable with regards to length and sequence. Only some residues of this loop interact with the sugar and these differ between lectins within the same monosaccharide specificity group. It seems, the length of Loop D in Gal/GalNAc specific lectins is longer than in the Man/Glc specific group and the length rather than its sequence has been suggested to be the main determinant for discrimination between these two specificity groups (Sharma & Surolia, 1997).

## **1.3 The Peanut Agglutinin (PNA) and T-antigen**

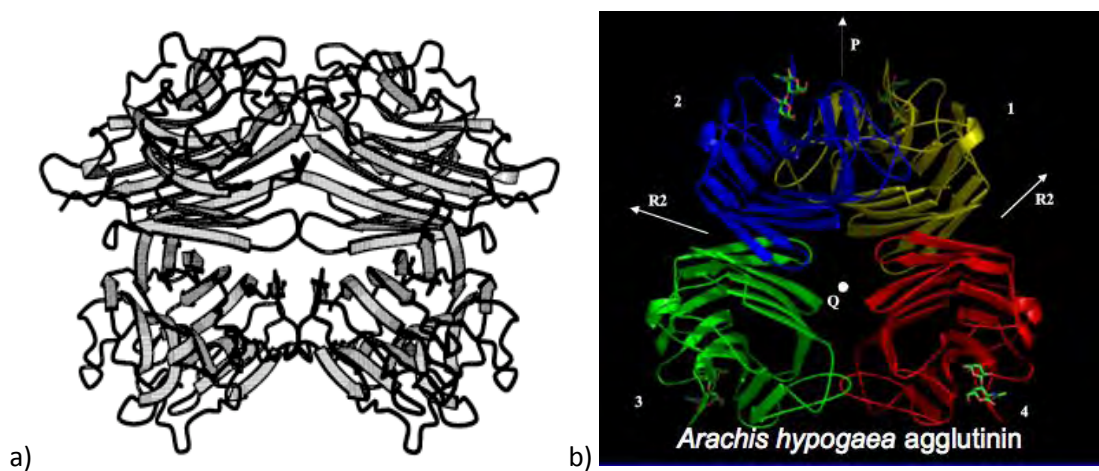
### **1.3.1 Structure and molecular recognition properties of the PNA**

Literature suggests that the peanut agglutinin, PNA, is a non-glycosylated homotetramer of roughly 110 kDa (Figure 1.5), with each monomer consisting of 236 amino acids made up of 4 subunits of around 27.5 kDa (Young *et al*, 1991.). It has one

carbohydrate binding site per monomer/subunit (Figure 1.6) along with a single calcium and magnesium ion.



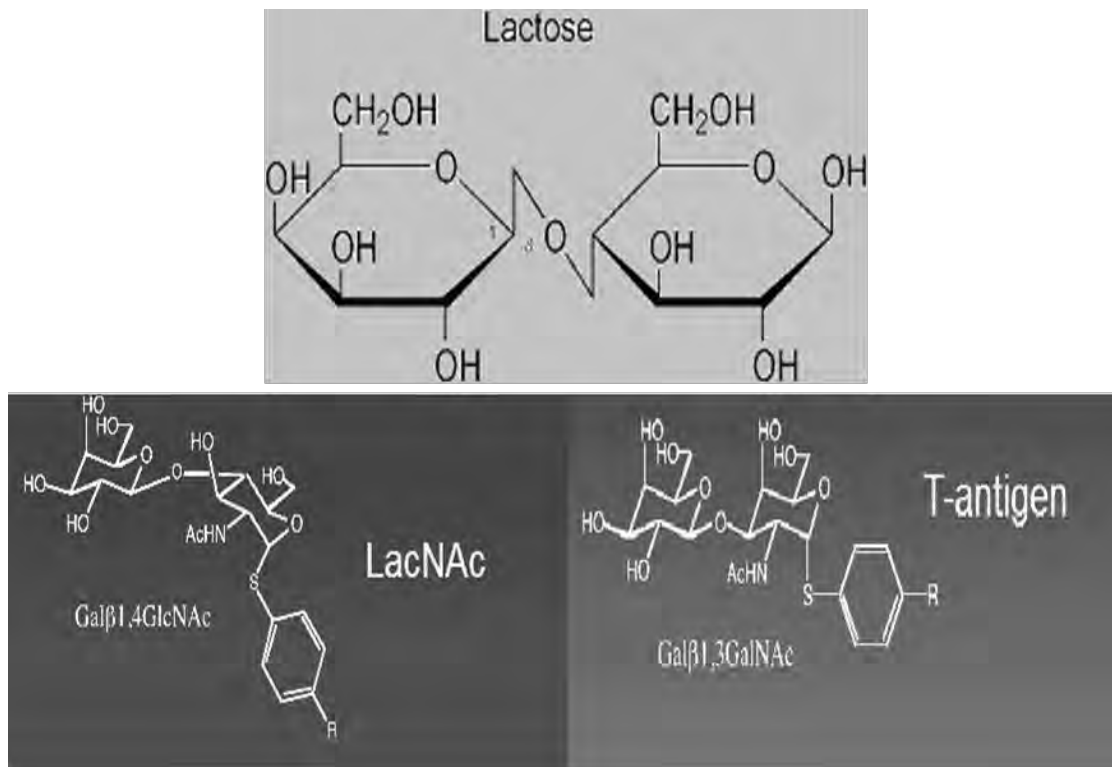
**Figure 1.5** The legume lectin fold. The lectin fold shown as a) Diagrammatic view and b) represented by one subunit of peanut agglutinin (PNA) (Sharpe, 2003).



**Figure 1.6** Tetrameric fold of PNA. Displaying the four subunits of PNA (Sharpe, 2003).

It is interesting that PNA differs from many other lectins through the use of the stabilizing water bridges between the monomers (Dev & Surolia, 2006). What makes the PNA unique in the tetrameric lectin group is the dimeric interface between monomer 3 and 4, (see Figure 1.6) (Dev & Surolia, 2006). Ligand specificity on the other hand, is generated by both direct and water mediated hydrogen bonding to key hydroxyls on the sugars. Divalent cations such as  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  are critical to ligand binding by PNA since they help to control the shape of the binding site and as previously mentioned, lectins differ in their specificity for ligand binding as well as quaternary structure, (Dev & Surolia, 2006).

The crystal structure of PNA reveals an apparent homotetrameric, non-glycosylated nature, whilst a wealth of literature has shown that PNA is able to recognize the sugars lactose, N-acetyllactosamine (LacNAc), glucose and the Thompson Freidenreich antigen (T-antigen) – a cell marker with a proven link to malignancy which is therefore an important cancer biomarker (Figure 1.7).



**Figure 1.7** Structures of Lactose, LacNAc and the T-antigen. Chemical structures of the sugars that recognise PNA (Swiss PDB viewer).

Although aberrant glycosylation is typically shown to be present in all human cancers, studies suggest that it is a result of oncogenic transformation, rather than the cause (Hakomori *et al.*, 2002), and the altered sugar epitopes produced are thus considered as tumour associated antigens. PNA has been tested as a diagnostic and prognostic indicator in many cancers (Sotozono *et al.*, 1994; Campbell *et al.*, 1995) because it allows for the monitoring of differential expression of the T-antigen as well as allowing for early detection of ‘T-polyagglutinability’, and the separation of mature and immature thermocytes for use in bone marrow transplant (Reisner *et al.*, 1979).

### 1.3.2 Thomsen Freidenreich antigen (Gal $\beta$ 1-3 GalNAc $\alpha$ <sub>1</sub>-O-Ser/Thr)

T- and Tn (GalNAc $\alpha$ <sub>1</sub>-O-Ser/Thr) blood group precursors, occur in normal tissues as the more abundant cryptic T- and Tn that are masked by sialylation, however in most

cancerous tissues, these T- and Tn antigens are not sialylated (Springer *et al.*, 1997). PNA has an advantage over other anti-T antigen lectin probes, such as Amaranthin and Jacalin, since PNA is able to distinguish between sialylated and non sialylated T-antigens (Lotan *et al.*, 1975, Rinderle *et al.*, 1989, Swamy *et al.*, 1991).

The limiting factor in use of PNA as a cancer diagnostic tool though is its cross recognition of the healthy cell epitope LacNAc and although the wildtype PNA binds to the T-antigen with  $\sim 10$  fold higher affinity than to LacNAc, this cross reactivity remains limiting. Other methods of selective recognition of the T-antigen have also been explored. For example, monoclonal antibodies were produced to be specific for the T-antigen, but these still displayed some cross reactivity towards normal cells and gave conflicting results in their binding to carcinoma cells (Balduş *et al.*, 2000). Small peptides have also been selected from libraries for binding to T-antigen (Peletskaya *et al.*, 1997) and the initially low affinity of these peptides has recently been improved to around  $K_D \cong 60$  nM (Landon *et al.*, 2003) but the use of these peptides, particularly with regard to their inability to agglutinate, remains to be seen.

Lectin crystal structures of 26 legume lectins in complexed and uncomplexed states are currently known and accessible, (3D Lectin Databank, <http://www.cermav.cnrs.fr/lectines/> accessed in March 2013) and classified as tetramers, comprised of dimers of dimers or strictly dimers. The first X-ray study on PNA was in the early 1980's (Salunke *et al.*, 1981) but the crystal structure was not solved until 1994 (Banerjee *et al.*, 1994).

PNA has been described by crystallography as being a dimer of dimers with each dimer being formed by back to back association of monomers 2&3 and 1&4 through flat back beta sheets. The monomers in each dimer are related by a two-fold axis, and the two dimers are related by another two fold axis (Salunke *et al.*, 1985). The interface of each dimer between each monomer is made up of side by side alignments of two flat  $\beta$  sheets which is a common form of dimerization as seen in many other

legume lectins (Dev & Surolia, 2006). However, there is some evidence that PNA might in fact exist in solution as a dimer, not a tetramer (Sharpe, 2003) which warrants further investigation.

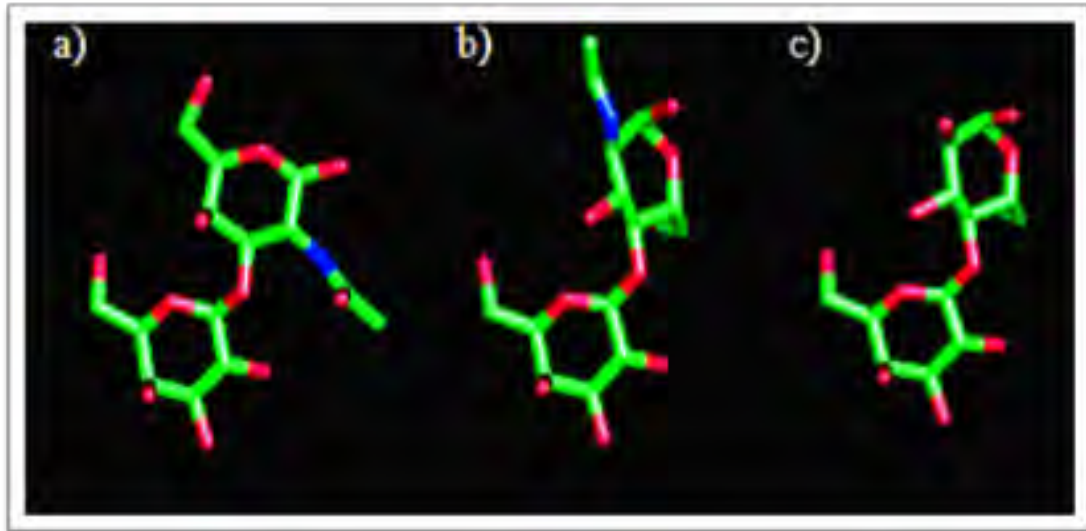
The affinity of lectins binding for simple monosaccharides is usually weak, with dissociation constants, ( $K_D$ ) typically in the range 0.1-10 mM reflecting the relatively shallow binding site in legume lectins which contrast the bacterial periplasmic sugar binding proteins that envelope the ligand and binds with much higher affinity (Vyas *et al.*, 1991). However, increased selectivity and binding affinity in lectins is achieved by expansion of the primary binding site to include a secondary binding site, thereby enabling selective higher affinity binding of disaccharide.

The quaternary structure of legume lectins gives rise to a clustering of binding sites, enabling the high affinity recognition of multivalent sugar ligands through the binding of one ligand to each of the subunit binding sites. The relative orientation of the various binding sites in the quaternary structure also determines specificity; in some lectins, such as ricin (a galactose-binding protein), and cholera toxin, the binding sites all face the same direction, enabling binding to multivalent carbohydrates on a cell surface; in others, such as Con A and PNA, the binding sites are at opposite ends of the molecule suggesting a bridging function, that is consistent with the known function of some galectins in cell-cell interactions and the ability of certain legume lectins to agglutinate red blood cells (Drickamer *et al.*, 1995).

PNA is also able to recognise sugars including LacNAc, lactose and galactose (see fig 1.8), with around 10, 30 and 50 fold lower affinities than towards T-antigen respectively (Sharma and Surolia 1996, Adhikari *et al.*, 2001).

The PNA is able to bind with greatest affinity to the T-antigen with a reported dissociation constant ( $K_D$ ) between 3-22 $\mu$ M at 4 to 25 °C (Neurohr *et al.*, 1982, Adhikari *et al.*, 2001). The structural basis for the discrimination of PNA between

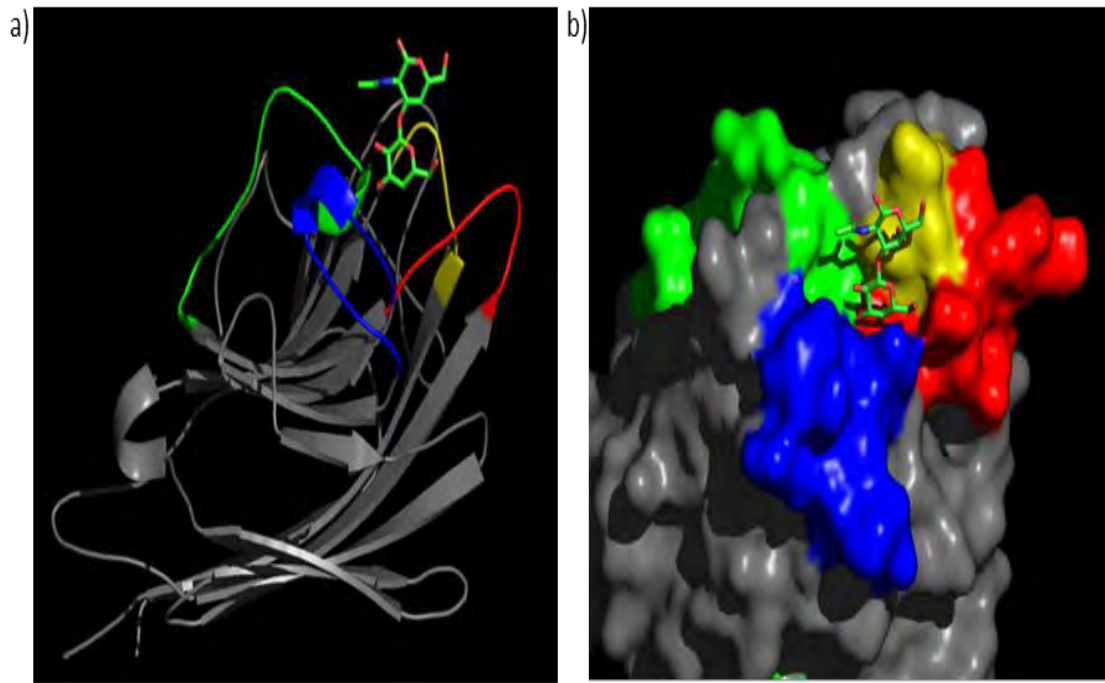
these specific sugars has been revealed by crystal structures of PNA in complex with different sugars



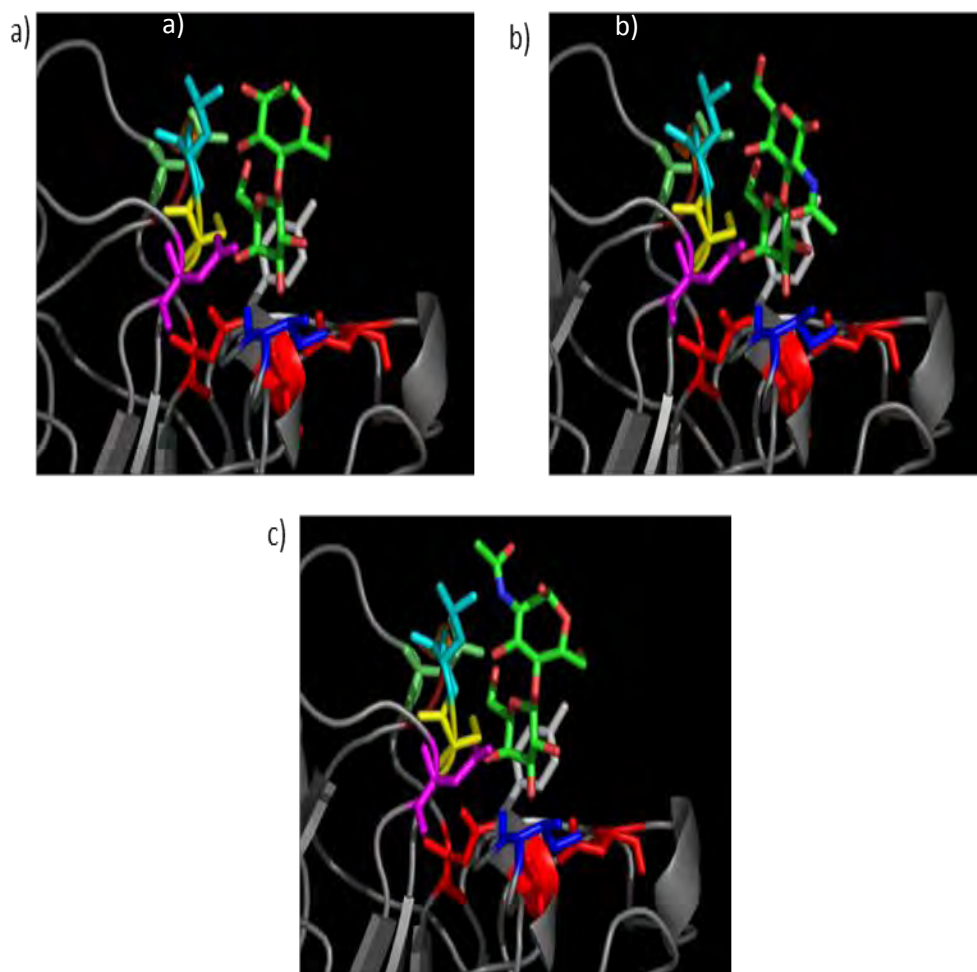
**Figure 1.8** Structures of a) the T-antigen, b) LacNAc c) lactose. Sugars as they appear in the binding site of PNA, showing the similar orientation of the primary galactose ring (Sharpe, 2003).

The PNA crystal structure has been solved in complex with the T-antigen (Ravishankar *et al.*, 1999a), lactose (Banerjee *et al.*, 1996) and LacNAc (Ravishankar *et al.*, 1999b). It is interesting to note that, in all crystal structures only the first 232 residues are seen clearly in the electron density map, with the last 4 residues appearing to be disorganised.

Typically, as is seen in all other legume lectins the highly conserved carbohydrate binding site is established from residues from the four loops A (75-83) B (91-106) C (125-135) and D (211-216) (Young & Oomen, 1992) (Fig. 1.9).

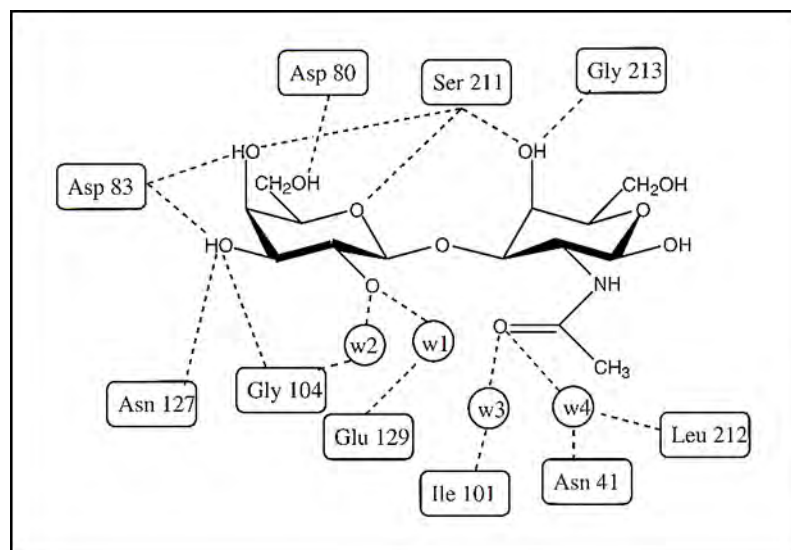


**Figure 1.9** PNA monomer in complex with the T-antigen. Showing a) the position of the four carbohydrate binding loops and b) Surface model of the carbohydrate binding pocket. For each, loop A= red, loop B=green, loop C=blue, loop D=yellow (Sharpe, 2003).



**Figure 1.10** PNA in complexes with a) T-antigen, b) Lactose and c) LacNAc. The sugar is coloured green in all structures and the invariant triad of Asp83, Asn127, and Gly104 are shown in red. Other residues highlighted include: Ser211= yellow, Leu212= light blue, Gly213= dark red, Gly214= green (Sharpe, 2003).

The water-mediated interactions increase the affinity for Gal at the primary binding site in PNA and are mediated by two water molecules, w1 and w2 as can be seen below (Figure 1.11). The Gal O2 interacts with Glu 129 OE1 through w1 and with Gly104 N through w2 in the T-antigen, lactose, and LacNAc complexes respectively. The galactose moiety of each sugar is anchored in the primary binding site by the same contacts. The difference in the contacts made by the second ring of each sugar is therefore the origin of the specificity of PNA (Sharpe, 2003).



**Figure 1.11** The carbohydrate binding site of PNA in complex with the T-antigen representation. Here the water molecules that aid in the mediation between sugar and protein can be seen (Sharpe, 2003).

The specificity of PNA for the T-antigen over lactose and LacNAc is directly related to the presence of these additional water bridges (w3 and w4). PNA does not bind to GalNAc like other lectins similar to it (Adhikari *et al.*, 2001).

#### 1.4 Engineering of plant lectins to improve specificity

The fields of medical diagnostics and biotechnology have explored the use of lectins due to their potential value in the identification of the presence of specific carbohydrates, glycolipids and glycoproteins on cell surfaces (Cook, 1986). Their use as reagents for blood typing and mitogenic stimulation of lymphocytes are still employed, but today one of their major uses is in the detection of change in cell surface glycosylation patterns, which is evident in a number of conditions such as

malignant transformation, cell maturation and aging. Additional uses of lectins include typing of bacteria, identification and separation of cells, an example of the latter being the removal of mature T-cells from bone marrow by soybean agglutinin prior to transplantation treatment of children with severe combined immunodeficiency. Lectins also find application in the purification of recombinant glycoproteins for therapeutics, (Sharon & Lis, 1986, Van Damme *et al.*, 1998b).

The desire and ability to modify lectins for use in medicine is driven today by the growing understanding of the evolution of binding proteins. Common recognition principles are evident in, for example the binding of carbohydrates to proteins through the combined use of hydrogen bonding and Van der Waals interactions. Specificity of the carbohydrate recognition on the other hand is generated by the direct water mediated hydrogen bonding to important hydroxyls and by the presence of  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$  cations which interact with the sugar indirectly as seen in legume lectins or directly. This will impact on the shape or orientation of the binding site therefore altering the positioning of the sugar (Sharpe, 2003)

It is therefore tempting to speculate that the binding specificity of legume lectins might be engineered through site directed mutagenesis that alters the pattern of hydrogen- binding between lectin and bound carbohydrate. One lectin that has been the subject of several mutagenic studies is PNA.

It has been demonstrated that PNA also shows some non-specificity for the T-antigen, as it binds lactose and N-acetyl-lactosamine (LacNAc), commonly found at the termini of cell surface sugars. Because of this, studies, redesigning the substrate specificity of this lectin have been undertaken (Sharma *et al.*, 1996, Adhikari *et al.*, 2001). Sharma *et al.* (1996) described the creation of two mutants within the carbohydrate binding loop, at positions, L212A and L212N. The L212N mutation apparently leads to an increased affinity towards lactose, unchanged binding to the T-antigen but loss of binding to LacNAc, while L212A leads to elimination of binding affinity towards the T-antigen, but an increased affinity for LacNAc. In a later study Adhikari *et al.* (2001)

targeted residue N41 which is important in making critical water mediated contacts with the T-antigen. All three mutants analysed (N41A, N41D, and N41Q) reportedly retained an unchanged affinity for lactose, but exhibited a marked decrease, slight decrease and 4-fold increase in binding affinity for the T-antigen respectively. The results of these interesting studies indicate the potential of engineering PNA for the creation of improved specificity towards the T-antigen.

In more recent work, Sharpe and colleagues (2003) used a novel *in-vitro* phenotype-genotype linkage plasmid display system to select PNA variants with altered binding properties from a random mutation library. Starting from a library in which a number of residues had been randomised, a number of PNA variants were identified based on their binding to immobilised T-antigen and preliminary characterisation by Surface Plasmon Resonance (SPR) suggested that these mutants had improved recognition properties towards the T-antigen- these mutant PNA therefore warrant further investigation.

## **1.5 Aim of thesis**

The aims of this Master's thesis are therefore:

- 1) To provide further experimental evidence for altered specificity of three variant forms of PNA that originated from the work of Sharpe and colleagues.
- 2) To determine the multimeric state of PNA under varying experimental conditions.

## Chapter 2: Materials and Methods

Unless otherwise stated, the fine chemicals and reagents were supplied by Fluka, Saarchem Chemicals, Merck Chemicals and Laboratory supplies OR Sigma Aldrich chemical company. DNA size markers, protein ladders and restriction endonucleases/enzymes were supplied by Fermentas. The *pceseq* and *6xHisfor* oligonucleotide primers were made and purchased from Stellenbosch Sequencing Unit University of Stellenbosch, South Africa. Carbohydrates were purchased from (Lectinity, Moscow). Substrate for ELISA was purchased from R&D systems. Microarray slides were from Schott/ Nexterion. All solutions were prepared using sterile milli-Q water.

### 2.1 Materials

#### LB media

Tryptone 10g

Yeast 5g

NaCl 10g

LB agar – all above including Agar 15g/l

Above chemicals were prepared into 1L of distilled water.

Autoclaved prior to use after which appropriate antibiotic added.

#### Ethanolamine (200 mM, pH 8.0)

Ethanolamine (98%)

NaOH

Add 2.55 mL of ethanolamine to ~150 mL of H<sub>2</sub>O. Adjust the pH to 8.0 with NaOH and bring to 200 mL final volume with H<sub>2</sub>O.

**1 M Na<sub>2</sub>HPO<sub>4</sub> (100 ml)**

Na<sub>2</sub>HPO<sub>4</sub> 14.2 g

Add distilled H<sub>2</sub>O to final volume of 100 ml

**1 M NaH<sub>2</sub>PO<sub>4</sub> (100 ml)**

NaH<sub>2</sub>PO<sub>4</sub> 12 g

Add distilled H<sub>2</sub>O to final volume of 100 ml

**1M Sodium Phosphate buffer pH 8.0**

1 M Na<sub>2</sub>HPO<sub>4</sub> 93.2 ml

1 M NaH<sub>2</sub>PO<sub>4</sub> 6.8 ml

100ml

Adjust pH with NaH<sub>2</sub>PO<sub>4</sub> and Na<sub>2</sub>HPO<sub>4</sub>.

Autoclave and store at room temperature.

**Guanidine (6M pH6.4)**

Guanidine-HCl 573. 18 gram (g)

1M Tris-HCl (pH7.6) 50ml

0.5 M EDTA 40ml

dH<sub>2</sub>O to 1l

Total volume 1l

**1M KCl**

KCl 7.5 gram (g)

dH<sub>2</sub>O to 100ml

Total volume 100ml

**1 M H<sub>2</sub>SO<sub>4</sub>**

Sulfuric Acid (H<sub>2</sub>SO<sub>4</sub> 95-97%) 5.6 ml

dH<sub>2</sub>O to 100 ml

Total volume 100 ml

Note: Dilute by adding slowly sulfuric acid to 80 ml water and mix, adjust final volume to 100 ml.

### **TFBI**

30 mM potassium acetate 1.47 g

50 mM manganese chloride 4.95 g

100 mM rubidium chloride 6.05 g

10 mM calcium chloride 0.56 g

15% (w/v) glycerol 75 ml

Bring the volume up to 500 ml using distilled water.

Adjust the pH with 0.2N acetic acid to pH 5.8. Throw the solution out if the pH is less than 5.8. (i.e. 5.82 or so is useful)

Sterile filter the solution and keep at 4 °C

### **TFBII**

10 mM mops pH7.0 0.42 g

75 mM calcium chloride 1.66 g

10 mM rubidium chloride 0.24 g

15% (w/v) glycerol 30 ml

Bring the volume up to 200 ml by adding distilled water.

Adjust the pH with 0.2N acetic acid to pH 5.8. Throw the solution out if the pH is less than 5.8. (i.e. 5.82 or so is useful)

### **Chlororamphenicol**

34 mg/ml in EtOH: -20°C

Chloroamphenicol stock 34mg/ml

Required at 30 ug/ml

### **1 M HEPES, pH = 7.0**

HEPES (free acid) 119.15 g

Distilled water to 400 ml

Add solid NaOH a few pellets at a time while mixing until the pH is ~6.8

Add concentrated NaOH dropwise to achieve pH = 7.0

Distilled water to 500 ml

Sterile filter and store at 4°C

### **1 x Spotting Buffer pH 8.3**

Final concentrations of 100 mM KCl

Triton X-100 0.01%

HEPES pH 8.5 25 mM

Made up to 100ml final volume in distilled water.

### **10 M Ammonium acetate**

Dissolve Ammonium acetate 770 g

H<sub>2</sub>O 800 mL

Adjust volume with H<sub>2</sub>O. 1 L

Sterilize by filtration.

Store the solution in tightly sealed bottles at 4°C or at room temperature. Ammonium acetate decomposes in hot H<sub>2</sub>O and solutions containing it should not be autoclaved.

### **SDS page separating gel 10% gel**

Water 4.1ml

1.5M Tris-HCl pH 8.8 2.5ml

20% (w/v) SDS 0.05ml

Acrylamide/Bis concentrate  
(30%/ 0.8 % w/v) 3.3ml

10% Ammonium persulphate 0.07ml

TEMED 0.01ml

Ammonium persulphate and TEMED added immediately prior to casting

**SDS page stacking gel**

Water	3.07ml
0.5M Tris-HCl pH 8.8	1.25ml
20% (w/v) SDS	0.025ml
Acrylamide/Bis concentrate (30%/ 0.8 % w/v)	0.67ml
10% Ammonium persulphate	0.025ml
TEMED	0.01ml

Ammonium persulphate and TEMED added immediately prior to casting

Coomasie staining and destaining solutions prepared by technical staff.

**1M Sulphuric acid**

Add 5.33ml of 95% H<sub>2</sub>SO<sub>4</sub> to 100ml of water (always add acid to water)

**50mM Biotin**

Dissolve biotin in	0.12215g
dH <sub>2</sub> O	5ml
Add 1M NaOH to dissolve solids with shaking	
Make up to	10ml
Filter sterilize	

**Blocking solution (Streptavidin 1mM)**

Streptavidin	1ml
1x PBS triton solution	49ml

**0.5 M EDTA pH 8.0 1 liter**

Na <sub>2</sub> EDTA.2H <sub>2</sub> O	186.1 g
Milli-Q H <sub>2</sub> O	800 ml

Bring pH to 8 with NaOH pellets (about 20 g); Bring volume to 1 liter with dH<sub>2</sub>O and autoclave.

**1 M Tris (pH as desired)**

1 M Tris pH 8.5 1 liter

Tris base                      121.1 g

H<sub>2</sub>O                                ~750 ml

Adjust pH to 8.5 with ~20 ml conc. HCl

Bring to 1 liter with distilled H<sub>2</sub>O; don't autoclave.

Store at RT

**2 M Tris pH 7.5**

To make 1 liter:

242.2 g Tris base

600 ml Milli-Q H<sub>2</sub>O

Dissolve Tris in H<sub>2</sub>O and adjust pH with ~ 125 ml conc. HCl

Bring volume to 1 liter with Milli-Q H<sub>2</sub>O; don't autoclave

**3 M NaCl**

To make up 2 liters

350.6 g NaCl

Make up to 2 liters with dH<sub>2</sub>O

Store at RT

**7.5 M Ammonium acetate**

To make 1 liter:

578.1 g NH<sub>4</sub>Ac

Dissolve in dH<sub>2</sub>O to final volume of 1 liter.

Check the pH; it should be about 7.4

Bring volume to 1 liter with dH<sub>2</sub>O and autoclave.

### **10 N NaOH**

400 g NaOH per liter dH<sub>2</sub>O

### **10X TBE**

1 L

Tris 1 08.8 g

Boric Acid 55.0 g

0.5M EDTA 40 mL

Make up to 1L with distilled water.

### **Lysis-Equilibration-Wash Buffer (1 x LEW Buffer, 1 liter Protino Kit)**

50 mM NaH<sub>2</sub>PO<sub>4</sub> 7.8 g NaH<sub>2</sub>PO<sub>4</sub> x 2 H<sub>2</sub>O (MW = 156.01 g/mol)

300 mM NaCl 17.5 g NaCl (MW = 58.44 g/mol)

Adjust pH to 8.0 using NaOH

Prepare in 1L of distilled water

### **1 x Elution buffer (Protino Kit)**

50 mM NaH<sub>2</sub>PO<sub>4</sub> 7.8 g NaH<sub>2</sub>PO<sub>4</sub> x 2 H<sub>2</sub>O (MW = 156.01 g/mol)

300 mM NaCl 17.5 g NaCl (MW = 58.44 g/mol)

8 M urea 480.5 g (MW = 60.06 g/mol)

250 mM imidazole 17.0 g imidazole (MW = 68.08 g/mol)

Adjust pH to 8.0 using NaOH

Prepare in 1L of distilled water

### **5X TBS**

1 L

Tris 24.2 g

NaCl 292.4 g

Make up to 1L in distilled water.

pH to 7.5

**Add Ethidium bromide (EtBr) to the Gels NOT the running buffer.**

Add 5ul (10mg/ml) EtBr per 100ml TAEI (caution: mutagenic)

**5x loading dye for agarose gels**

50% glycerol            50 ml

0.5 M EDTA pH 8      10 ml

1 M Tris pH 7.5        5 ml

Add a tiny amount (a few mg) of the dye xylenol orange--just enough to color the solution. Add 35 ml dH<sub>2</sub>O to total volume 100 ml.

**6X Loading Dye (Fermentas):**

0.09% bromophenol blue (just add enough until you get a color you like)

0.09% xylene cyanol FF

60% glycerol

60mM EDTA

**10X PBS**

Phosphate buffered saline, recipe from Harlow and Lane.

To make up 1 liter

NaCl                            80 g

KCl                              2 g

Na<sub>2</sub>HPO<sub>4</sub>                        14.4 g

KH<sub>2</sub>PO<sub>4</sub>                        2.4 g

Dissolve in 800 ml dH<sub>2</sub>O

Adjust pH to 7.4 with HCl

Bring volume to 1 liter, autoclave

To make up 1x PBS, dilute 100ml of a 10x solution in 900ml of dH<sub>2</sub>O.

**PBS-Tween:**

1.5 mM KH<sub>2</sub>PO<sub>4</sub>,

4.3 mM Na<sub>2</sub>HPO<sub>4</sub>

137 mM NaCl  
3 mM KCl  
0.1% (v/v) Tween20  
Adjust to pH 7.4 with HCl.

**1 M CaCl<sub>2</sub>**

To make 1 liter  
110.9 g CaCl<sub>2</sub>  
dH<sub>2</sub>O to 1 liter  
Aliquot 100 ml/bottle, autoclave

**1 M MgCl<sub>2</sub>**

To make up 100 ml  
20.3 g MgCl<sub>2</sub>·6H<sub>2</sub>O  
Make up to 100ml with dH<sub>2</sub>O  
Autoclave  
Store at RT  
Glycerol (100%) dilute in dH<sub>2</sub>O. Sterilize by autoclaving.

**Destain**

	<u>1 L</u>
Methanol	50 mL
Acetic Acid	42 mL

Make up to 1L with distilled water.

**Stain**

	<u>1 L</u>
Methanol	150 mL
Acetic Acid	100 mL
Coomassie Blue	2 g

Make up to 1L with distilled water.

**10X Transfer Buffer**

	<u>1 L</u>
Tris	24.7 g
Glycine	112.6 g

Do not pH the transfer buffer

Make up to 1L with distilled water.

**Blot Transfer Buffer**

	<u>1 L</u>
Methanol	200 mL
Glycine	14.4 g
Tris	3.03 g

Make up to 1L with distilled water.

**10% (w/v) Ammonium Persulfate (APS)**

	<u>500 mL</u>
Ammonium persulfate	50 g

Make 50 mL aliquots in distilled water. Store at -20°C.

**Lysozyme, 50 mg/mL**

	<u>10 mL</u>
Lysozyme	0.5 g

Make up 10ml volume in distilled water.

Store at -20°C.

## **2.2 Methods**

### **2.2.1 DNA characterization and analysis**

#### **Preparation and growth of competent cells for transformation of plasmid DNA.**

Competent *E.coli* JM 109 cells were prepared using the Rubidium Chloride method (Promega Protocols and Applications Guide (3<sup>rd</sup> edition), p. 45-46). The cells were then transformed to Chloramphenicol antibiotic resistance with plasmid DNA of interest using the Inoue method of transformation (Inoue et al. 1990 Gene 96, 23-28). Single colonies containing transformed cells were plated onto a LB Cam agar and incubated overnight at 37°C.

pCESN20 (6xHis), pCES N41R (6xHis) or pCES I101R (6xHis), (see Figure 3.1 and 3.2) clones were picked and inoculated into 5ml LB Cam media and incubated overnight at 37°C with shaking at ~ 180 rpm and glycerol stocks of samples were prepared for long term storage at -80 °C. This was done by adding 100µl of sample into a 1ml tube with 400µl of an 80% glycerol stock.

#### **2.2.1.1 Analysis and manipulation of DNA**

Plasmid DNA was purified from resulting 5ml cultures using the QIAGEN mini-prep kit according to manufacturer's protocol. The purified plasmid was eluted in 30-50ul TE buffer/water and the concentration was determined through quantification using a Nano Drop® ND -100 spectrophotometer (Nano Drop Technologies®). The use of zymogen clean and concentrator kits facilitates the removal of DNA polymerases, modifying enzymes, RNA polymerases, ligases, kinases, nucleases, phosphatases and restriction

endonucleases etc. to enable cleaner more concentrated DNA.

Plasmid DNA size of the three clones, was determined through 0.8% gel agarose gel electrophoresis using a 1x TBE buffer.

### **Restriction endonuclease digestion of DNA**

Samples were quantified using a Nano Drop ® (Nano Drop Technologies ®) ND-100 spectrophotometer and prepared for sequencing.

Using the appropriate buffers for each digest, restriction digest of 1µg of plasmid DNA was carried out using 1 unit of enzyme each: *NdeI* and *Xho I* were used for the double digest and *NcoI* was used for the single digest. To identify plasmids based on sizes of fragments, digests were separated on a 1% agarose gel (supplemented with EtBr) in TBE running buffer at 80 Volts for 45 minutes. The uncut plasmid size was 5139bp.

### **Sequencing of plasmid DNA**

*Pcseq* and *6xHisfor* primers (see appendix page 107) were synthesized by Dr. Di James at the Sequencing Unit, University of Cape Town Molecular Cell Biology or Stellenbosch sequencing Unit, University of Stellenbosch. Plasmid identity was confirmed by Sanger sequencing using the primers *pcseq For* and 6x His (Stellenbosch sequencing unit, Department of Genetics at the University of Stellenbosch). Concentration requirements for sequencing were ~600- 800 ng of DNA per plasmid sample.

## **2.2.2 Expression and purification of recombinant proteins**

### **2.2.2.1 Protein expression**

Protein was expressed from glycerol stock containing *E.coli* JM 109 transformed with pCES N20 (6xHis), pCES N41R (6xHis) or pCES I101R (6xHis) clones. Each transformation mix was streaked out on LB Cam (Chloramphenicol at 34ug/ml) plates and incubated overnight at 37°C. Single colonies were aseptically inoculated into 5ml LB supplemented with Chloramphenicol (30µg/ml) and grown at 37°C overnight in an orbital shaking incubator at 200rpm. Cultures were scaled up by adding the 5ml into 500ml in LB Cam, and grown overnight at 30 °C, with shaking at 230 rpm.

### **2.2.2.2 Cell harvesting and fractionation**

Cells were harvested by centrifugation at 12 000 rpm (Beckman J2-21M) for 30 minutes at 4°C. The supernatant was removed and the pellets stored at –20°C or resuspended in 10ml Lysis buffer (as prepared from a 4x concentrate, provided with Protino Purification kit) along with the addition of 1 mg/ml lysozyme, 500 U DNase I (New England Biolabs), 0.5mM CaCl<sub>2</sub>, 2.5mM MgCl<sub>2</sub>, giving a final volume of 11ml for each culture. The cell suspensions were incubated on ice for 30 min with gentle agitation prior to sonication.

Samples were subjected to sonication (Virsonic 100) for 3x 10 second on and 10 seconds off bursts on ice at 70% duty cycle. The sonicated cell lysate was then clarified by centrifugation at 10 000 rpm for 20 minutes at 4 °C and soluble lysate was removed and filtered through a 0.45µM filter. The insoluble pellet was discarded. The cell suspensions were then centrifuged at 10,000 rpm for 30 min at 4°C and the resulting supernatants were the crude soluble lysate preparations. 100 µl of each lysate was removed and stored at -20°C. All soluble proteins were then filter sterilized through a 45 micron filter prior to use in Protino columns.

### **2.2.2.3 Protein purification**

Protein was purified using either Protino His tag columns or Ni-NTA agarose beads (Qiagen purification system) for 6 x His tagged proteins. Protino columns were pre equilibrated with 1x lysis/wash buffer (see Materials and Methods), after which 5-6 ml of the crude lysate was added to the column, and washed twice with 2ml lysis solution. The HIS tagged protein was eluted from the column twice with a 2ml of 1x elution buffer as per the Protino kit. The Bradford assay was used to quantify yield of purified protein.

To prepare the Ni-NTA agarose beads, 500 $\mu$ l slurry of resin was centrifuged in a 15ml tube at 4000rpm for 10 minutes and storage buffer solution was removed by aspiration. The resulting resin pellet was resuspended in 2ml of the crude lysate and was then centrifuged again at 4000rpm for 10 minutes. The supernatant was removed and beads were washed twice in 2ml of wash buffer and recovered by centrifugation as previously described. Beads were then resuspended in 0.5ml of elution buffer and then pelleted by centrifugation as before, this was repeated 4 times, the supernatant in each being kept for SDS-PAGE analysis. Samples were analysed on 10 % SDS gel or stored at 4 °C, for further analysis.

### **2.2.2.4 Determination of protein concentration by Bradford assay and spectrophotometer readings of pure lysate**

Samples were prepared at a dilution of 1:10 with Bradford (undiluted) reagent. The concentration of the protein was determined at 595nm for both crude and purified protein. A 60 $\mu$ l of pure protein sample was added to 540 $\mu$ l elution buffer and vortexed. 500 $\mu$ l of this was mixed with 500 $\mu$ l Bradford's reagent and allowed to equilibrate for 20 minutes. The OD<sub>595</sub> was measured on spectrometer. Protein concentration was

determined by comparison to standard curve treated under same conditions as sample was prepared, by known concentrations of BSA as per reagent suggestions and whether standard/micro-assay.

#### **2.2.2.5 Polyacrylamide gel electrophoresis (PAGE).**

10% polyacrylamide gels were prepared for crude proteins and for purified proteins as follows: two glass plates were immobilized onto a gel tank stand (BIORAD) and the preferred percentage separating gel was prepared and allowed to polymerize with a covering layer of isopropanol. After 20 minutes, the isopropanol layer was decanted off and the top of the separating gel was washed off with water. The stacking gel was then poured above the separating gel and a comb inserted; the stacking gel was allowed to set for 20 minutes. After polymerization the comb was removed and the glass sheets removed from the stand and immobilized onto the gel tank (BIORAD) and immersed in a 1x SDS PAGE running buffer. Pre-stained protein ladder was loaded directly into the first lane without boiling. 10-15 $\mu$ g of each sample was loaded mixed with 5 $\mu$ l of a 4x SDS loading buffer and the soluble fractions boiled at  $\sim$ 90 $^{\circ}$ C for 5 minutes, debris removed by centrifugation and 20 $\mu$ l of the supernatant was directly loaded into each well. The gel was run at 35mA constant current for 2 hours until the Bromophenol blue of the loading buffer reached the bottom of the gel. The gel equipment was then disassembled and the gel removed and stained with Coomassie staining solution for 4 hours followed by destaining the gel in destaining solution overnight to enable visualization of protein bands.

#### **2.2.2.6 Western blot analysis**

After proteins were subjected to SDS page gels, protein gels were placed between sheets of filter paper and nitrocellulose membrane (Amersham, Hybond ECL) and soaked in transfer buffer. The physical set up was as follows:

Positive electrode → Filter paper  
Nitrocellulose membrane  
Gel

Negative electrode → Filter paper completely immersed and soaked.

The blotting apparatus was set up using a BIORAD Miniprotean® tank, with an ice pack, and transfer of proteins onto the membrane was achieved by electrophoresis at 100 volts for 1 hour in transfer buffer. The blot was removed, the membrane stained in Ponceau's stain for a minute to check for the presence of any band's. The membrane was then removed from the Ponceau stain and incubated in 5% skim milk powder to block any unbound protein an hour on a shaker at room temperature. Milk (skim) contains a generic mix of a relatively high concentration of known proteins; it is used to block the parts of the membrane that don't already have protein on so that the antibodies won't bind there.

The membrane was then washed 3 x 5min in 1x PBS. Anti His HRP conjugate was diluted to 1:1000 in 10 ml of a 5 % milk powder and incubated on the membrane used to probe the blot for 1 hour. The blot was then washed for 3x 5min in 10 ml PBS + 0.1 % Tween, followed by a final wash in distilled water for 5 min. The blot was then air dried and probed with an HRP chemiluminescent substrate (Supersignal West Pico) (Sigma Aldrich) according to manufacturer's protocol; chemiluminescent signal was recorded on autoradiography film and the film was developed according to standard protocols. Alternatively, the membrane was probed with *Anti-Arachis Hypogaea*, agglutinin (Sigma Aldrich, #L7759) as the primary antibody and the secondary Goat anti Rabbit IgG antibody (Santa Cruz, #sc2030), HRP conjugated, dilution at 1:1000 in 10 ml 1xPBS buffer and was then incubated for 1 hour.

## **2.2.3 Native gel electrophoresis**

### **2.2.3.1 Mobility shift assay**

Native mobility shift assay was carried out according to standard protocols (Ausubel *et al.*, 1994) A 4%Tris Glycine gel was prepared, and re-run for 30-60 min at 100 volts. The gel was loaded with a prestained protein ladder, unstained protein ladder, and varying concentrations of wildtype PNA treated with different concentrations of Guanidine Hydrochloride, as well as untreated wildtype PNA and Concanavalin A. A stock solution of 6M guanidine was prepared and diluted down to 4M, 2M, 1M and 0.5M concentrations in dH<sub>2</sub>O prior to mixing with purified PNA and analysis by native gel.

## **2.2.4 Examination of quaternary structure of proteins**

### **2.2.4.1 MALDI Matrix Assisted Desorption Ionization - Time of Flight mass spectrometry Confirmation of quaternary structure of PNA**

MALDI mass spectrometry sample preparation.

MALDI mass spectrometry of purified samples was carried out by Dr. Putuma Gqamana in the Blackburn lab U.C.T, the aim being to establish whether the protein structure and size changes were due to pH, as well as whether there was any N-terminal proteolysis occurring.

Typically, minimum concentrations of 0.1 mg/mL were required in order to observe high-quality spectra for the intact protein by matrix-assisted laser desorption/ ionization

mass spectrometry (MALDI MS). Hence, the protein extract was further purified and concentrated to the desired volume by buffer exchanging with HPLC grade water, or an appropriate pH buffer in the case of pH studies, using the 3 kDa molecular weight cutoff (MWCO) Nanosep<sup>®</sup> ultracentrifugation devices (Pall Corporation).

Purified protein samples were then buffer exchanged into 0.1 M Ammonium acetate buffers; pH 3, 7 and 9 using Vivaspin desalt spin columns (GE Healthcare Life Sciences, # 28-9322-18). Proteins were used from concentrated form or at 1:10 or 1:100 dilutions. Columns were washed with 500µl water. The sample was then added to the column and centrifuged for 10 min at 14000rpm, leaving a residue of 100µl of 400µl ammonium acetate at the concentration of required pH range was then added and the tube was spun at 14000rpm for ten minutes leaving a residue of 100µl; this was repeated once, after which the samples were diluted to a final volume of 500µl, by addition of 400µl 0.1M ammonium acetate.

Thereafter, 0.5 µL of the purified, buffer-exchanged PNA sample was spotted on a 384 Opti-TOF 123mm x 81mm plate (MDS Sciex, Ontario, Canada), overlaid with equal volume of saturated  $\alpha$ -cyano-4-hydroxycinnamic acid (CHCA) matrix solution (on the order 5 – 15 mg/mL in 30:70 (v/v) 0.1 % aqueous trifluoroacetic acid (TFA) /acetonitrile (ACN) solution). The spots were vacuum-dried at room temperature and analyzed by a 4800 MALDI TOF/TOF<sup>™</sup> Analyzer in the high mass linear mode with external calibration using BSA (66, 431 *m/z* units)

### **2.2.5 High Performance Liquid Chromatography (HPLC)**

High-performance liquid chromatography (HPLC) is a chromatographic technique used to separate mixtures of compounds in chemistry/biochemistry/biotechnology with the purpose of identifying, quantifying and purifying the components of the mix individually.

It is also useful in establishing quaternary structure of purified proteins. Size exclusion chromatography, or SEC, is used primarily for the analysis of large molecules such as proteins or polymers and SEC, works by trapping these smaller molecules in the pores of a particle; The bigger molecules will simply pass by the pores, as they are too large to enter the pores. Larger molecules therefore flow through the column quicker than smaller molecules.

The wildtype (wt) PNA purified proteins, were prepared at a concentration of ~ 1mg/ml. Purified protein samples were buffer exchanged into either citric acid (pH 3), sodium phosphate (pH 7) or -Tris (pH 9) buffers respectively. For HPLC analysis by SEC the mobile phase used was sodium phosphate at 0.1M in pH 8 buffer. A sample volume of 50µl was loaded at roughly 1mg/ml onto the column (Waters Biosuite 250 7.5 diameters / internal 3mm length instrument) where a maximum load volume of 200µl can be used. Prior to analyzing the N20 (wild type) sample, the protein was centrifuged to remove any residue/precipitants and filter sterilized to avoid clogging the HPLC column. Eluted protein was monitored at 215 and 280 nm where, 215 corresponds to peptide bonds and 280 to aromatic residues absorption, respectively (tyrosine, tryptophan and phenylalanine).As the control samples BSA and Concanavalin A were prepared at 1mg/ml, and from this, 50µl of each were injected separately onto the column.

## **2.2.6 Affinity binding assay**

### **Antibody labeling and streptavidin coating of glass slides and antibody labeling Cy5 his and Cy5 PNA**

#### **2.2.6.1 Preparation of Cy3-BSA and Cy5-biotin-BSA or for Anti HIS/Anti PNA antibody Cy5 labeling**

To establish the labelling of BSA with Cy5-labelled antibody, 500  $\mu$ l of 10 mg/ml BSA in PBS pH 7.4 was aliquoted into two separate 1.5ml plastic tubes. 0.5 mg of Biotin-NHS ester was added to one tube at a 1 mg/ml final concentration. The tubes were incubated for 4 hrs at room temperature, on a platform shaker to enable the biotin-NHS to react with amine groups on the BSA in the one aliquot. A total of 2 mg each of the biotin-BSA and unbiotinylated BSA from each of these tubes (in the original buffer and concentration) were added to two separate foil-wrapped tubes of Cy5 dye powder (#PA23001, GE Healthcare). The vials were vortexed to ensure complete dissolving of the content of the vials, and the vials were then incubated for an hour at RT with vortexing after every 10 minute interval. After incubation, unbound Cy5 dye was removed by applying the samples to Sigma Spin Post-Reaction Purification Columns (#S5059, Sigma Aldrich) as per manufacturer's instructions for Cy-labelled DNA. Aliquots of Cy5-BSA and Cy5-biotin-BSA were stored at -20°C for later use.

Anti His and anti PNA were also labelled in the same manner, but a total of at least 1 mg/ml of antibody is added to each foil wrapped vial due to light sensitivity of the dye as exposure to light will compromise the activity of the antibody. Vials were incubated for an hour with 10 min vortex bursts. These were applied to Sigma spin columns and the protocol followed as described above.

Spectrophotometry assays were carried out to determine the dye-to-protein (D/P) ratio,

to ensure that adequate Cy5-labelling had been obtained, which, as per the manufacturer's information, should optimally be  $\geq 2$ . This is especially important in the case of the Cy5-biotin-BSA, which was initially reacted with biotin prior to dye labelling, as both labels will react with amine groups on the protein surface. This ratio can be calculated by measuring the protein's absorbance at  $A_{280}$  and the dye absorbance at  $A_{552}$  (8% of the dye's  $A_{552}$  reading is then subtracted from the  $A_{280}$  reading to correct for the dye's absorbance at  $A_{280}$ ). However, the  $A_{280}$  measurement requires large quantities of the protein sample unless micro volume cuvettes are accessible. Therefore, a Bradford standard assay was carried out at  $A_{595}$  instead to determine the BSA concentration. No absorbance of the dye at  $A_{595}$  was observed.

The concentrations of the BSA and the dye were determined and the D/P ratio was calculated (molar dye concentration/molar BSA concentration) using a MW of 66.43 kDa for BSA and an extinction coefficient of  $0.25 \mu\text{M}^{-1}.\text{cm}^{-1}$  for Cy5. A D/P of 21 was observed for Cy5-BSA, while a D/P of 29 was found for Cy5-biotin-BSA. The reason for the higher D/P in the case of biotin-BSA was not determined. A concentration of 2mg of both Anti His and Anti PNA was also prepared as described above. The Anti His antibody D/P was calculated to be around 19.9 D/P. whilst Anti PNA was shown to have a D/P of around 2.12.

### **2.2.6.2 Streptavidin coating of glass slides and antibody labeling Cy5 Anti His and Cy5 Anti PNA**

#### **Preparation and testing of home-made streptavidin-coated array surfaces**

A vial of streptavidin lyophilisate (#S0677, Sigma Aldrich) was equilibrated to RT. The lyophilisate was then dissolved in water to  $\sim 10 \text{ mg/ml}$  and the reconstitution was allowed to proceed for 30 min. The solution was divided into 1 mg aliquots and stored at  $-20^\circ\text{C}$ . A 1 mg/ml working solution was prepared in 1 x Spotting Buffer (final

concentrations of 100 mM KCl, 0.01% Triton X-100, 25 mM HEPES pH 8.5) as needed. Care was taken that the Spotting Buffer had a pH > 8.3, as this was necessary for the protein amine groups to remain mostly un-protonated and reactive towards amine-reactive groups on the slide surface

### **Preparation of streptavidin-coated slides for microarray printing**

A Nexterion Slide P (or H) microarray slide was equilibrated to room temperature for an hour and removed from the foil packaging. A 1mg/ml streptavidin solution in 150mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8.5) was made up and glass microarray ‘lifterslip’ was placed over the microarray surface. 60µl of the streptavidin solution was pipetted along the edge of the lifterslip in order for the solution to be drawn under the cover slip uniformly by capillary action. Slides were then left for 1hr at room temperature in a humidified chamber. The lifterslip was removed and the slide washed for 1hr at room temperature in 10ml 150mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8.5) containing 50mM ethanolamine to deactivate any remaining amine-reactive groups. The slide was washed for 3 x 5 min in 10ml wash buffer and then for 5 min in 10ml water. The slide was placed in a 50 ml Falcon tube and centrifuged at 1000rpm for 1 min at 20°C to spin dry. The newly streptavidin-coated slides were placed into slide boxes and sealed in Ziploc bags and stored at -20°C.

### **Testing of binding specificity to streptavidin-coated slides**

The slides were removed from -20°C storage and equilibrated to RT for 1 hr. Cy5-BSA and Cy5-biotin BSA were diluted 1:1250 (~6 ng/ml) in 1 x Spotting Buffer. Two slides each were incubated in Cy5-BSA and Cy5-biotin-BSA for 30 min at RT on a platform shaker. The slides were then washed 3 x 5 min in PBST (0.1% Tween-20 in 1 x PBS pH 7.4) and 1 x 5 min in dH<sub>2</sub>O with shaking. Slides were spun dry at 1000 rpm for 5 min and scanned at 120 PMT at 10 µm resolution.

### **Testing of coating homogeneity across slide surface**

After each fresh batch of streptavidin-coated slides was prepared, a batch quality control test was performed to determine whether the homogeneity across the slide surface was adequate. One to two slides from each batch (depending on batch size) were incubated with Cy5-biotin-BSA and then washed, spun dry and scanned (as described for the binding specificity test above). Using GenePix Pro measuring tools, a rectangular region of 22 x 58 mM (the active area of the slide) was drawn. The average and median pixel fluorescence and the standard deviation across this area were then calculated. The same measurements were also performed for a rectangular region of 7 x 7 mM to determine homogeneity across a smaller area.

### **Establishing uniformity of the streptavidin derivitisation.**

A single streptavidin-coated slide was equilibrated to room temperature and washed for 1 min in 150mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8.5) then centrifuged briefly (1000rpm; 1 min) to remove excess surface liquid. 1mg/ml solution of Cy5-labelled, biotinylated BSA in 150mM Na<sub>2</sub>HPO<sub>4</sub> buffer (pH 8.5) was prepared.

A glass microarray 'lifterslip' was placed over the microarray surface and 60µl of the Cy5-labelled; biotinylated BSA solution was pipetted along the edge of the lifterslip such that the solution was drawn under the cover slip uniformly by capillary action.

The slide was incubated for 1hr at room temperature in a humidified chamber after which the lifterslip was removed and the slide washed for 3 x 5 min in 10ml PBST, then centrifuged briefly (1000rpm; 1 min) to remove excess liquid. The slide was then scanned in a fluorescence microarray scanner (excitation 550nm; emission 570nm; set at 10 uM resolution and automatic gain control) and the uniformity of fluorescence across the slide determined.

An alternative method was also employed whereby all steps requiring lifter slips covering the slides were instead replaced by immersing the slides in around 10 ml of each solution required at incubation or wash steps. These slides were shaken on an orbital shaker for different time incubations and wash steps required.

## **2.2.7 Microarray assay –Protein-Carbohydrate binding assay**

### **2.2.7.1 Hand spotting streptavidin coated slides**

Prior to the preparation of the slides for hand spotting the microarrays, the lyophilised Biotinylated Polysaccharide Sugars were resuspended at 1mg/ml in 50mM phosphate buffer each sugar, i.e Lactose, N-acetyllactosamine, and the Thompson-Fredenreich antigen (T-antigen) were diluted into 50mM Sodium Phosphate buffer with 20 % glycerol to 1mM concentrations. These sugars were then further diluted to 200uM as a working solution.

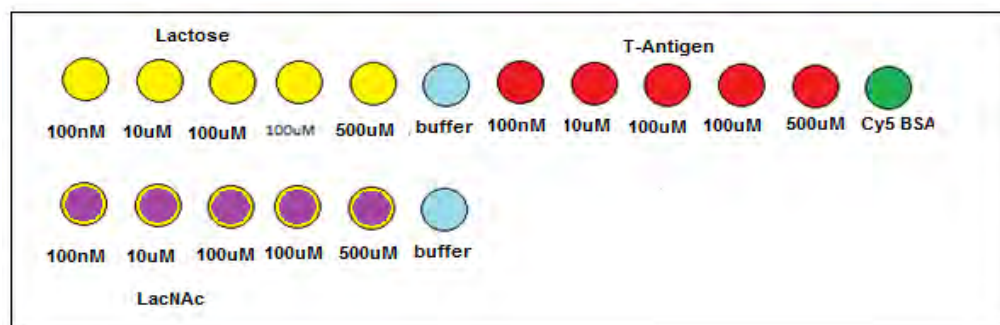
Streptavidin coated slides were removed from -20°C storage and allowed to equilibrate to room temperature. Using a hand spotting template, 1ul of sample of biotinylated sugar was spotted onto the slide surface in increasing concentrations across each slide with a range of 1mM to 100nM, (1mM, 1uM, 100uM, 100nM) including spotting 1ul of a buffer only control. Biotinylated Cy5-labelled BSA was also spotted as a positive control (Figure 2.1).

These hand spotted slides were incubated for 30 min at room temperature. Slides were then washed 3x5 min in 1x PBS supplemented with 100 µM Biotin and 0.1mg/ml BSA in order to remove excess unbound sugar on the surface and to block any residual biotin binding sites. Varying concentrations of protein; wildtype (wt) PNA, PNA (I101R) or PNA (N41R) ranging from 100µM to 100nM were placed underneath a lifter slip above the hand printed slide, or alternatively allowed to wash over the slide continuously with slow movement.

If incubated under a lifterslip, about 50µl liquid was needed to cover the slide. If proteins were spotted in areas allowing a 16 well gasket to fit over the spots, then each resultant 16 plex well needed to be incubated with 50µl of protein at the desired concentration. Each protein was diluted in the 1 x Elution Buffer (see Protino elution buffer in Materials and Methods section). The slides were incubated with PNA for one hour and then washed in PBST for 2 x 5 minutes. In this case, we employed the use of a lifterslip to allow the protein to wash across the slide.

Fluorescently labeled Primary antibody was prepared at the required dilution of 1:1000 in a 5% milk powder solution in PBS. Enough antibody was prepared so that, the slides could be incubated with the antibody in a container able to hold about 15ml of antibody dilution; alternatively using the 16 plex gasket, 50ul of Ab solution was added to each well and incubated in the dark overnight at 4°C.

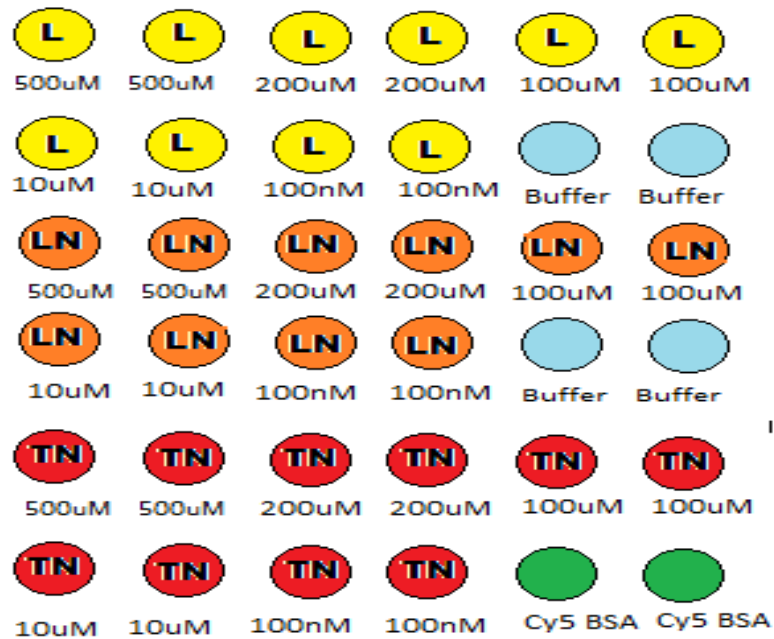
The slides were then washed in 3 x 5 min PBST (0.05% Tween20 in 1 x PBS)- followed by a single 5min wash in dH<sub>2</sub>O and slides were then spun dry in a 50ml Falcon tube at 500-1000rpm for 1 minute at 20 °C to dry. Scanning was performed at the desired 120 PMT voltage at the CPGR on a TECAN reloaded microarray scanner.



**Figure 2.1.** Hand spotted array. Slides that were hand spotted with sugars were plotted in the following manner incorporating both positive and negative control spots and varying dilutions of sugar. Lactose, Nacetyllactoseamine or T-Antigen and washing with either pCES N20, pCES I101R or pCES N41R. BSA spotted with a Cy5 fluorescent label. Sugars were diluted from a 1mM stock solution of each sample.

### 2.2.7.2 Microarray printing

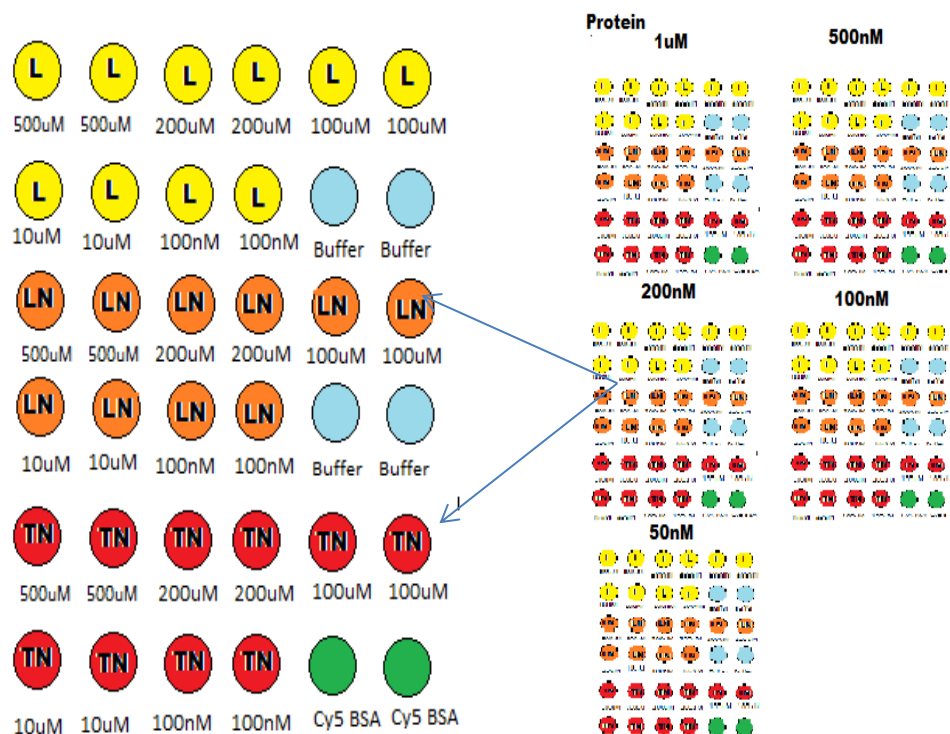
Streptavidin coated slides were removed from 20 °C freezer and equilibrated at room temperature for an hour before washing in 10mM sodium phosphate solution for 10 minutes. Slides were spotted by hand and spotting 1µl of each sugar on the slides at varying concentrations, or the spotting by microarray, printed abstractly on a Q Array 2 printer (Genetix) at the following conditions (fig.2.2) and dilutions of the sample to be printed was prepared according to the pattern in Figure 2.2. The slide is left in the humidifier after printing, for 30 minutes Q Array 2 (Genetix). Each sugar was prepared from the 1mM stock solution, where dilutions ranging from 500uM to 100nM were spotted.



**Figure 2.2.** Arrangement of print of sugars onto slide. Layout of microarray printed format.

One printed slide was initially scanned under automatic gain control after printing of the sugars, to determine whether there was any auto fluorescence of the biotinylated sugars.

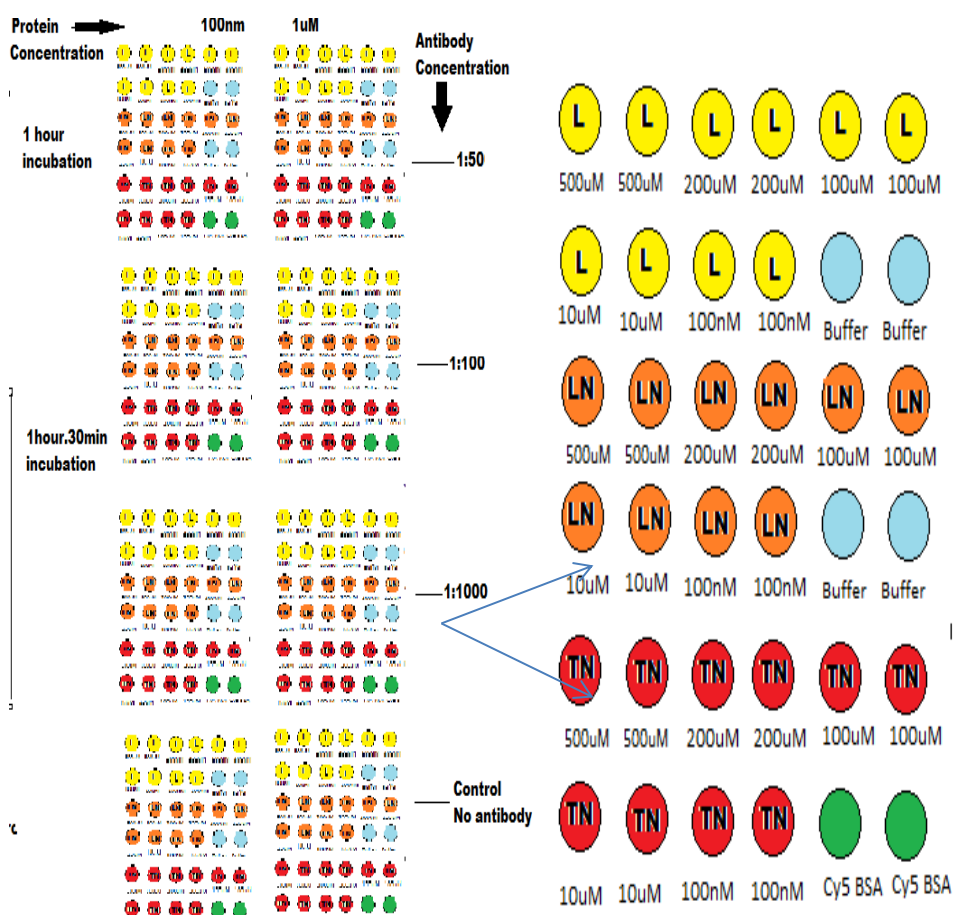
The slide pattern was arranged in a 16-plex format, with each area consisting of nano liter volume spots of each sugar, allowing for a larger number of sugar dilutions to be printed reproducibly across each 16 plex area. The slide was washed and then blocked in 5 % skim milk powder in PBS supplemented with 100 $\mu$ M Biotin at room temperature with gentle agitation. 100  $\mu$ M of each PNA variant was prepared, 50 $\mu$ l of each protein, was added to individual wells and the slide incubated for 1 hour at room temperature. This procedure was repeated across varying protein concentrations with each unique-protein/ concentration combination being added to a discrete replica 16 plex microarray. Slides were then blocked for 30 min in 5% milk powder in PBS at RT and then washed in 1xPBST and probed with an appropriate Cy5 labeled antibody (either Cy5 Anti His or Cy5 Anti PNA) at optimized concentration of 1:1000 in 5% milk powder PBS for one hour in the dark at RT. Slides were then washed: 3x5min PBST followed by one wash in distilled water, the slide was spin dried in a 50ml Falcon tube at 500rpm for 2 minutes at 20 °C to dry and scanned as before.



**Figure 2.3.** Print layout for each replica 16 plex array.

The slide orientation for each well (16wells) was prepared in this orientation to allow broad range of sugar concentrations to be evaluated in each well with all sugars.

In Figure 2.3 each protein was prepared from 1uM through to a lowest concentration of 50nM. One 16-plex array was used for a buffer only control. In the first instance varying concentrations of protein and sugar were employed in order to determine whether binding was taking place, initially only one concentration dilution of antibody was used to determine whether there was binding and then later on the antibody concentration was varied to determine the optimal antibody dilution to maximize specific signal.



**Figure 2.4.** Illustration of the assay. This format was used to determine the optimal antibody concentration and incubation time.

Figure 2.4 Above illustrates one of many attempts to establish the best assay conditions. In this instance we evaluated various primary antibody concentrations at two different

protein concentrations. This was aimed at identifying which antibody dilution gave the least amount of background and best foreground fluorescent signal for subsequent binding assays. Three slides were prepared in this manner in order to assay all three proteins. This assay made use of 2 protein concentrations down each side of the slide. Each biotinylated sugar was prepared at 5 different concentrations (500 $\mu$ M, 200  $\mu$ M, 100  $\mu$ M, 10  $\mu$ M and 100nM) along with a control sample (buffer only) and a Cy5 BSA sample (provided and prepared by Dr. Aubrey Shoko as a gift to us). Each spot was printed in duplicate at nL ranges across each individual well. Every printed assay was prepared in this manner, with the only changing parameters being varying the protein concentrations that were washed across the slide as well as the antibody concentration. Figure 2.4 displays the layout of each well with the variation of the incubation times of each antibody concentration used and the use of the two different protein concentrations used here.

### **2.2.8 ELISA Enzyme-linked Immunosorbent assay**

This experiment required pure protein and biotinylated carbohydrates i.e. T-antigen Lactose and LacNAc. NHS activated PEG (Poly-ethylene glycol) coated 96 well plates, were derivatised with 1mg/ml streptavidin for an hour at room temperature.

Streptavidin was then removed and the wells washed in 100ul of 1xPBS buffer. 100uM of each carbohydrate (50ul) was added to individual wells and incubated for an hour at room temperature. Wells were then washed in 100ul of PBS and blocked in 1mg/ml BSA for an hour. All liquids were aspirated with a pipette to remove any remaining liquid in the well. Sugar derivatised wells were then incubated with appropriate PNA variants at varying concentrations (3  $\mu$ M, 1.5  $\mu$ M, 0.75  $\mu$ M, 0.3  $\mu$ M, 0.18  $\mu$ M, 0.09  $\mu$ M, 0.04  $\mu$ M, 0.02  $\mu$ M) as per table 2.2 for an hour at room temperature, wells were washed with 100ul 1x PBS and incubated with 50 $\mu$ l of 1:1000 dilution of primary Anti His antibody HRP conjugate for an hour. Wells were then washed (3 x 1 minute) with 100ul PBST, followed by one water wash. Substrate (Substrate Reagent Pack, R&D systems catalogue

# DY999) was added and incubated for 20 minutes in the dark. Reaction was promptly stopped following the addition of 1M H<sub>2</sub>SO<sub>4</sub> to stop the reaction here. The microwell plate was then read on a multiplate /well reader at 450nm. Data was exported to excel for analysis.

**Table 2.1** .The steps of the ELISA assay are described.

Microwell derivitisation (step 1); addition of biotinylated sugar (step 2) ; addition of PNA variant (step 3); addition of Primary antibody (step 4); the combinations of sample set up in each well and of these varying concentration of each sample was analysed on PEG and high binding plates to determine which gave best signal when read on microplate reader at450 nm. Blocking attempts were made with milk powder, 1mg/ml BSA with 1mg/ml Biotin, 5% milk and 1mg/ml BSA

<b>STEP 1</b>	<b>STEP 2</b>	<b>STEP 3</b>	<b>STEP 4</b>
-----	sugar	protein	antibody
Streptavidin	-----	protein	antibody
Streptavidin-	sugar	-----	antibody
Streptavidin-	sugar	protein	-----
Streptavidin-	sugar	-----	antibody
Streptavidin	sugar	-----	-----
-----	sugar	-----	-----
-----	-----	-----	-----

The initial steps taken, prior to the first proper ELISA, performed, involved a process of elimination to determine whether our sugars, proteins, antibody or the streptavidin was causing any auto fluorescence. In table 2.1 above, the inclusion or absence of a specific parameter was examined to test for any auto fluorescence.

**Table 2.2.** The steps of the ELISA assay are described. Arrangement of sugar concentrations, and protein concentrations used.

	A	B	C	D	E	F	G	H	I
1	LAC	LAC	LAC NAC	LAC NAC	T - ANT	T - ANT	NO SUGAR	NO SUGAR	PNA conc
2	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	3uM
3	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	1.5uM
4	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	0.75uM
5	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	0.3uM
6	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	0.018uM
7	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	0.09uM
8	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	0.04uM
9	100uM	100uM	100uM	100uM	100uM	100uM	0uM	0uM	0.02uM

The table (2.2) above as previously referred to, displays the arrangement of the sugars, their concentrations and the varying PNA concentrations used in this experiment.

## **Chapter 3: Results and discussion**

### **3.1 DNA preparation**

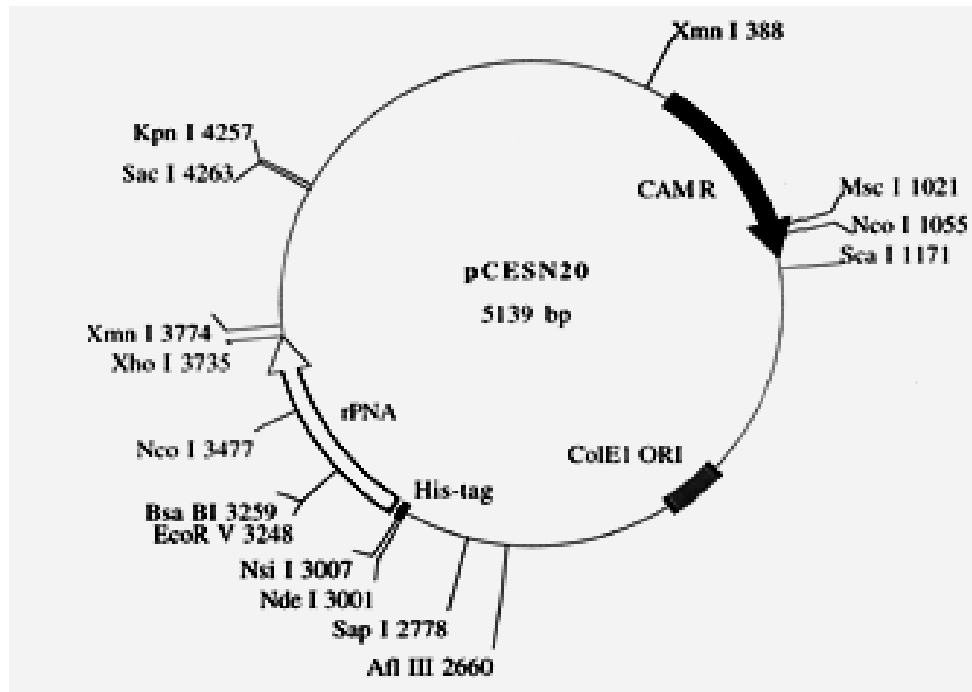
#### **3.1.1 DNA extraction and quantification**

Plasmid preparations using Qiagen purification kits yielded substantial amounts of DNA with ranges of between 100- 200ng/ul for each 5ml start culture. To improve yields, Zymogen® kits were used to concentrate DNA further, to ensure sufficient sample for separation on agarose gels and for further characterization. Upon concentrating, the samples yielded between 300- 400 ng/ul of DNA.

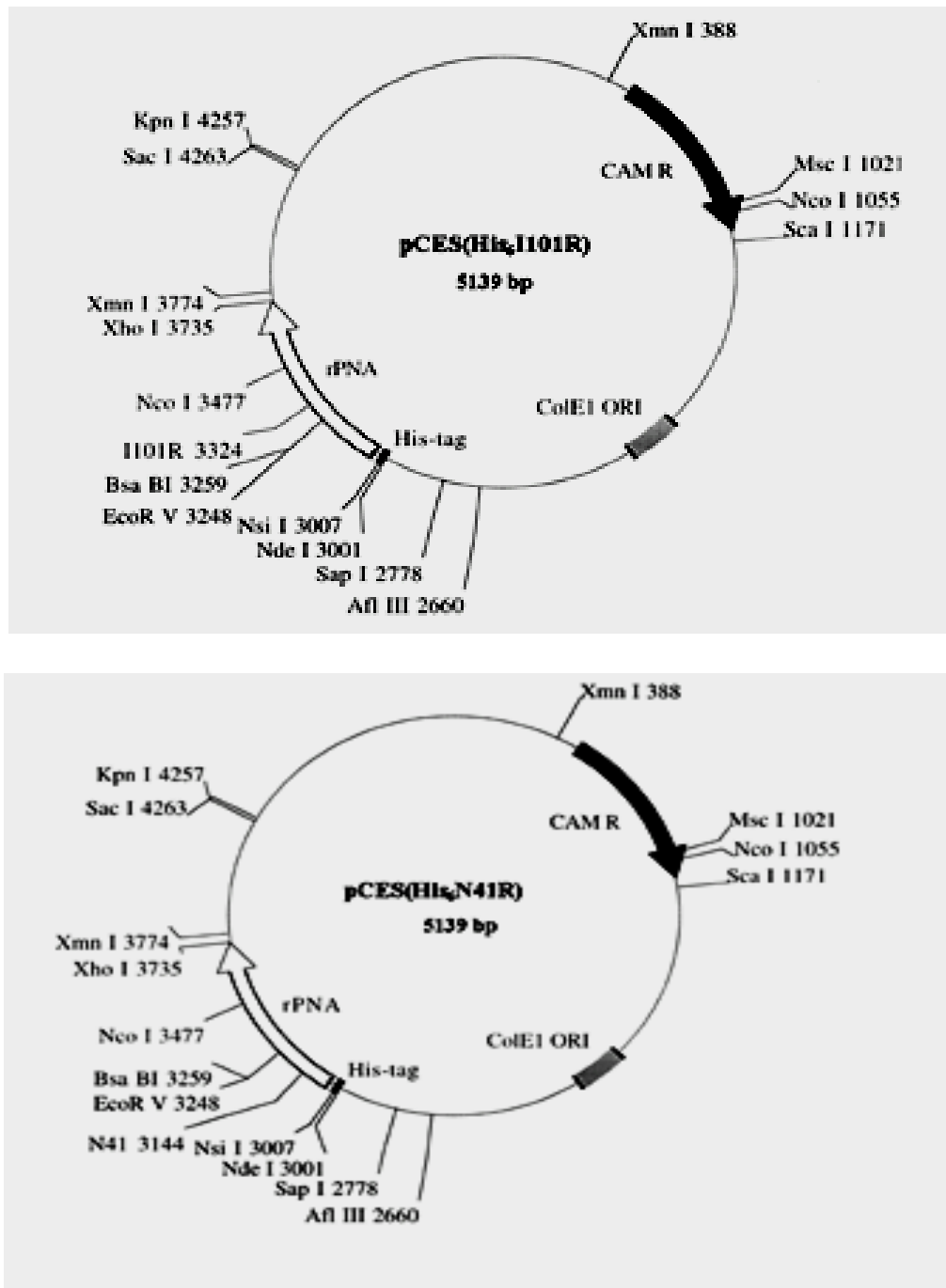
#### **3.1.2 Plasmid sequences from DNA extractions for protein expression and purification**

Historical plasmid constructs pCESN20 (6xHis), pCES N41R (6xHis) and pCES I101R (6xHis) (Sharpe, 2003) were used to transform fresh *E.coli* JM109 cells and new plasmid DNA was prepared. These constructs contained a Chloramphenicol resistance gene thereby making use of Chloramphenicol resistant antibiotics. Restriction analysis and sequencing confirmed clone identities (see section 3.1.5 and appendix for chromatograph sequences). These sequence–confirmed clones were used in all subsequent protein expression work

**Plasmid constructs encoding wildtype and variant PNA's.**

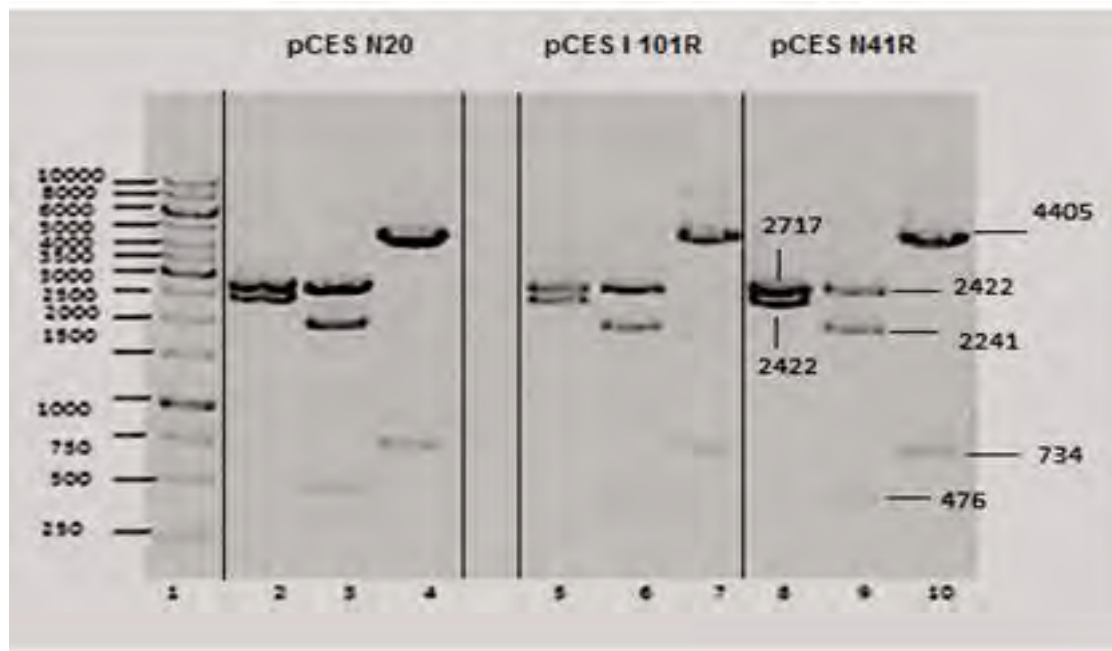


**Figure 3.1.** Plasmid map for pCES N20 (Sharpe, 2003) encoding wt PNA. Plasmid contains a chloramphenicol resistance gene and ColE1 origin of replication.



**Figure 3.2** Plasmid maps for pCES N411R and pCES I101R (Sharpe, 2003). These maps encode the PNA variants.

### 3.1.3 Restriction digest



**Figure 3.3.** Restriction digests of samples. pCES N20 (6XHis) , pCES PCES I101R, pCES PCES N411R. Lane 1 DNA ladder (Fermentas) Lane 4,7,10 is the NdeI/XhoI digest. Lane, 3, 6, 9 NcoI/NdeI double, Lane 2, 5, 8 NcoI single digest.

**Table 3.1** The expected fragment sizes in base pairs of each restriction digest.

	<b>pCES N20</b>	<b>pCES I101R</b>	<b>pCES N41R</b>
<i>NdeI/NcoI</i>	2422,2241,476 (bp)	2422,2241,476 (bp)	2422,2241,476 (bp)
<i>NcoI</i>	2717, 2422 (bp)	2717, 2422 (bp)	2717, 2422 (bp)
<i>NdeI/XhoI</i>	4405, 734 (bp)	4405, 734 (bp)	4405, 734 (bp)

### 3.1.4 Discussion: restriction digests

In Figure 3.3, restriction digestion products were separated on varying percentage agarose gels depending on sizes that needed confirmation. Gels were prepared in 1x TBE running buffer run at 80 Volts for ~ 2 hours. Restriction digests of plasmid encoding wild type PNA and the two mutants displayed the same pattern of digestion for single and the two double digestions. Digest sizes for both single and double digest were as follows: *NcoI* (2717bp, 2422bp) *NcoI/NdeI* (2422bp, 2241bp, 476bp). *NdeI/XhoI*(4405bp, 734bp) ( Table 3.1). Restriction digests suggested that the plasmids encoding both the mutants and wild type PNA's were correct, as judged by the observed expected band sizes. In some instances incomplete digestion was apparent, presumably due to insufficient or old enzyme. It could also be due to impure DNA, phenol or detergent contaminations as well as improper dilution or addition of enzyme.

### 3.1.5 Sequence analysis

Sequencing data (Figure 3.4 and see appendix chromatograph data) confirmed that the plasmids were correct as compared to original sequence (Uniprot, accession#P02872). Sequence results also confirmed the expected site directed mutations. Figure 3.4 the translated sequence of the encoded recombinant PNA's with amino acid changes at positions 41 and 101 in the two variants, are shown. The DNA sequence for the PNA was extracted and translated into amino acid sequences. Each protein sequence was aligned to the original PNA using SIM - Alignment Tool for protein sequences, Expasy tool. Chromatograph data for pCES N20 was not of the best quality, the chromatograph was unfortunately not available at the time to include in this thesis.

## Original protein PNA sequence as per Uniprot entry accession number P02872

### Original PNA precursor

MKPFCVFLTFLLLAASKKVDS....

```
.
AETVSFNFNS FSEGNPAINF QGDVTVLSNG NIQLTNLNKV NSVGRVLYAM PVRIWSSATG 60
NVAFLTSFS FEMKDIKDYD PADGIIFFFIA PEDTQIPAGS IGGGTLGVSD TKGAGHFVGV 120
EFDTYSNSEY NDPPTDHVGI DVNSVDSVKT VPWNSVSGAV VKVTVIYDSS TKTLVAVTN 180
DNGDITTIAQ VVDLKAKLPE RVKFGFSASG SLGGRQIHLI RWSFTSTLI TTTRSIDNN 240
EKKIMMASA 249
```

Original precursor aligned against wt PNA construct used here.

```
AETVSFNFNS FSEGNPAINF QGDVTVLSNG NIQLTNLNKV NSVGRVLYAM PVRIWSSATG 60 N20
AETVSFNFNS FSEGNPAINF QGDVTVLSNG NIQLTNLNKV NSVGRVLYAM PVRIWSSATG 60 ORIG

NVAFLTSFS FEMKDIKDYD PADGIIFFFIA PEDTQIPAGS IGGGTLGVSD TKGAGHFVGV 120 N20
NVAFLTSFS FEMKDIKDYD PADGIIFFFIA PEDTQIPAGS IGGGTLGVSD TKGAGHFVGV 120 ORIG

EFDTYSNSEY NDPPTDHVGI DVNSVDSVKT VPWNSVSGAV VKVTVIYDSS TKTLVAVTN 180 N20
EFDTYSNSEY NDPPTDHVGI DVNSVDSVKT VPWNSVSGAV VKVTVIYDSS TKTLVAVTN 180 ORIG

DNGDITTIAQ VVDLKAKLPE RVKFGFSASG SLGGRQIHLI RWSFTSTLI TTTRS 236 N20
DNGDITTIAQ VVDLKAKLPE RVKFGFSASG SLGGRQIHLI RWSFTSTLI 236
```

### Sample identity “wild type” pCES N20(6XHIS)

MHHHHHH...

```
1 AETVSFNFNS FSEGNPAINF QGDVTVLSNG NIQLTNLNKV NSVGRVLYAM PVRIWSSATG
61 NVAFLTSFS FEMKDIKDYD PADGIIFFFIA PEDTQIPAGS IGGGTLGVSD TKGAGHFVGV
121 EFDTYSNSEY NDPPTDHVGI DVNSVDSVKT VPWNSVSGAV VKVTVIYDSS TKTLVAVTN
181 DNGDITTIAQ VVDLKAKLPE RVKFGFSASG SLGGRQIHLI RWSFTSTLI TTTRS
```

100.0% identity in 236 residues overlap; Score: 1206.0; Gap frequency: 0.0%

```
PNA 1 AETVSFNFNSFSEGNPAINFQGDVTVLSNGNIQLTNLNKVNSVGRVLYAMPVRIWSSATG
N20, 1 AETVSFNFNSFSEGNPAINFQGDVTVLSNGNIQLTNLNKVNSVGRVLYAMPVRIWSSATG
*****
```

PNA 61 NVASFLTSFSFEMKDIKDYPADGIIFFIAPEDTQIPAGSIGGGTLGVSDTKGAGHFVGV  
N20, 61 NVASFLTSFSFEMKDIKDYPADGIIFFIAPEDTQIPAGSIGGGTLGVSDTKGAGHFVGV  
\*\*\*\*\*

PNA 121 EFDTYSNSEYNDPPTDHVGIDVNSVDSVKTPWNSVSGAVVKVTVIYDSSTKTLVAVTN  
N20, 121 EFDTYSNSEYNDPPTDHVGIDVNSVDSVKTPWNSVSGAVVKVTVIYDSSTKTLVAVTN  
\*\*\*\*\*

PNA 181 DNGDITITIAQVVDLAKALPERVKFGFSASGSLGGRQIHLIRSWSFSTLITTTTRS  
N20, 181 DNGDITITIAQVVDLAKALPERVKFGFSASGSLGGRQIHLIRSWSFSTLITTTTRS  
\*\*\*\*\*

### Sample identity pCES I101R

MHHHHHH...

1 AETVSFNFNS FSEGNPAINF QGDVTVLSNG NIQLTNLNKV NSVGRVLYAM PVRIWSSATG  
61 NVASFLTSFS FEMKDIKDYPADGIIFFIA PEDTQIPAGS RGGGTLGVSD TKGAGHFVGV  
121 EFDTYSNSEY NDPPTDHVGI DVNSVDSVKT VPWNSVSGAV VKVTVIYDSS TKTLVAVTN  
181 DNGDITITIAQ VVDLAKALPE RVKFGFSASG SLGGRQIHLI RSWSFSTLI TTTTRS

99.6% identity in 236 residues overlap; Score: 1199.0; Gap frequency: 0.0%

PNA 1 AETVSFNFNSFSEGNPAINFQGDVTVLSNGNIQLTNLNKVNSVGRVLYAMPVRIWSSATG  
I101R,1 AETVSFNFNSFSEGNPAINFQGDVTVLSNGNIQLTNLNKVNSVGRVLYAMPVRIWSSATG  
\*\*\*\*\*

PNA 61 NVASFLTSFSFEMKDIKDYPADGIIFFIAPEDTQIPAGS **I**GGGTLGVSDTKGAGHFVGV  
I101R,61 NVASFLTSFSFEMKDIKDYPADGIIFFIAPEDTQIPAGS **R**GGGTLGVSDTKGAGHFVGV  
\*\*\*\*\* **I**\*\*\*\*\*

PNA 121EFDTYSNSEYNDPPTDHVGIDVNSVDSVKTPWNSVSGAVVKVTVIYDSSTKTLVAVTN  
I101R,121EFDTYSNSEYNDPPTDHVGIDVNSVDSVKTPWNSVSGAVVKVTVIYDSSTKTLVAVTN  
\*\*\*\*\*

PNA 181DNGDITITIAQVVDLAKALPERVKFGFSASGSLGGRQIHLIRSWSFSTLITTTTRS  
I101R,181DNGDITITIAQVVDLAKALPERVKFGFSASGSLGGRQIHLIRSWSFSTLITTTTRS  
\*\*\*\*\*

### Sample identity pCES N41R

MHHHHHH....

```
1   AETVSFNFNS FSEGNPAINF QGDVTVLSNG NIQLTNLNKV RSVGRVLYAM PVRIWSSATG
61  NVASFLTSFS FEMKDIKDID PADGIIFFIA PEDTQIPAGS IGGGTLGVSD TKGAGHFVGV
121 EFDTYNSSEY NDPPTDHVGI DVNSVDSVKT VPWNSVSGAV VKVTVIYDSS TKTLSVAVTN
181 DNGDITITIAQ VVDLKAKLPE RVKFGFSASG SLGGRQIHLI RSWSFSTSLI TTTRRS
```

99.6% identity in 236 residues overlap; Score: 1200.0; Gap frequency: 0.0%

```
PNA      1 AETVSFNFNSFSEGNPAINFQGDVTVLSNGNIQLTNLNKVNSVGRVLYAMPVRIWSSATG
N41R,    1 AETVSFNFNSFSEGNPAINFQGDVTVLSNGNIQLTNLNKVRSVGRVLYAMPVRIWSSATG
          *****R*****

PNA     61 NVASFLTSFSFEMKDIKDIDPADGIIFFIAPEDTQIPAGSIGGGTLGVSDTKGAGHFVGV
N41R,   61 NVASFLTSFSFEMKDIKDIDPADGIIFFIAPEDTQIPAGSIGGGTLGVSDTKGAGHFVGV
          *****

PNA    121 EFDTYNSSEYNDPPTDHVGI DVNSVDSVKTVPWNSVSGAVVKVTVIYDSS TKTLSVAVTN
N41R, 121 EFDTYNSSEYNDPPTDHVGI DVNSVDSVKTVPWNSVSGAVVKVTVIYDSS TKTLSVAVTN
          *****

PNA    181 DNGDITITIAQVVDLKAKLPERVKFGFSASGSLGGRQIHLIRSWSFSTSLITTTTRRS
N41R, 181 DNGDITITIAQVVDLKAKLPERVKFGFSASGSLGGRQIHLIRSWSFSTSLITTTTRRS
          *****
```

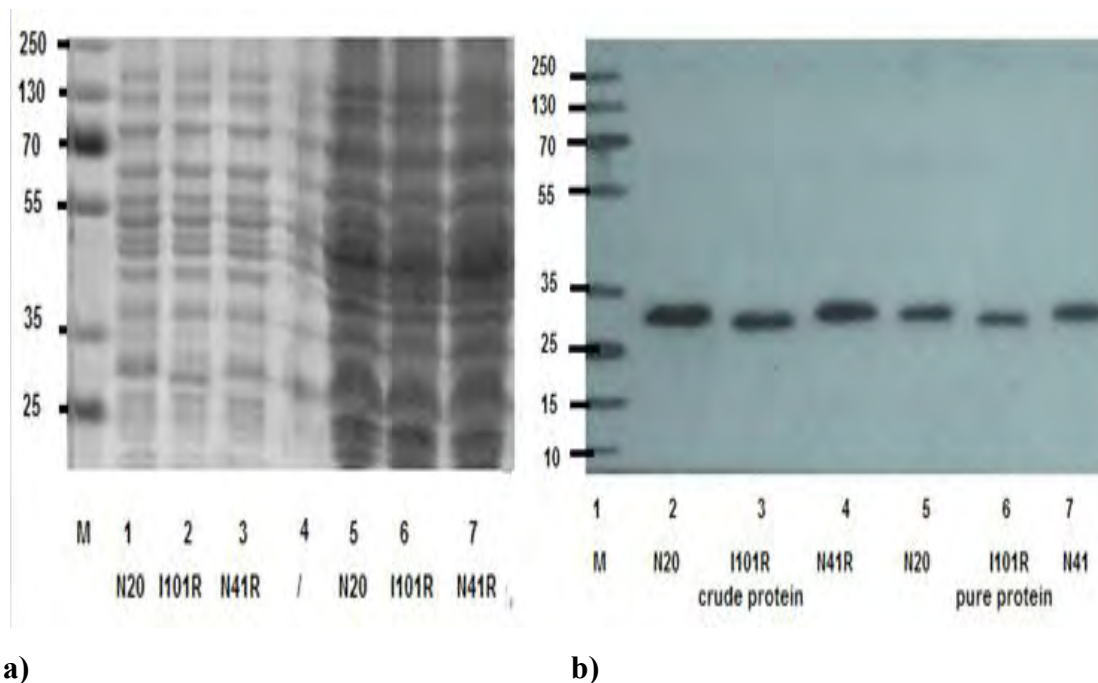
**Figure 3.4.** Amino acid sequences of wt and variant PNA are translated from DNA sequences. Clearly indicated in red are the amino acid changes compared to the wild type. Each protein sequence was aligned to the original PNA using SIM - Alignment Tool for protein sequences, ExPASy tool.

This conserved PNA region has been cloned identically to the precursor however, besides the mutations that are highlighted above, the N terminal of the Original sequence i.e. ‘MKPFCVFLTFFLLLAASKKVDS’ was removed and replaced with MHHHHHH which is the 6xHis tag. The original PNA sequence was compared to our sequence to confirm sequence homology identity. The purpose of this N- terminal tag being replaced

with a HIS tag, was to enable ease of HIS tag affinity purification using the Protino kit.

## 3.2 Protein expression and purification

### 3.2.1 Protein expression and purification analysis using SDS page



**Figure 3.5.** SDS analysis of protein expression and western blot.

Figure A represents proteins expressed at 30°C, for the wildtype N20, lane 1 and 5, I101R lane 2 and 6 and N411R lane 3 and lane 7. The amounts loaded on each gel were ~10-15 µg of protein. Only soluble fractions of protein were loaded. Figure B displays the western blot analysis of the purified proteins, with PNA band's appearing where expected.

The protein expression of PNA was performed at 37°C initially and then again at 30°C. The protein expression image in the figure above was to display how the expression of PNA looked on an SDS gel and will be discussed further below.

### 3.2.2 Expression of recombinant PNA variants

The relevant plasmids were introduced into *E.coli* JM109 cells for expression studies. Protein was expressed at 30°C without induction due to the presence of the lac O promoter mutation present in these vectors. Total protein yield as per the Bradford assay was approximately 1 to 5mg/ml for each sample but this was also dependent on the method of lysis employed. Analysis of total crude lysate on a 10% acrylamide gel revealed a prominent band at 30kDa; (between 25 and 35 kDa) the expected molecular weight of monomeric PNA was ~27.5 kDa (Figure 3.5). Proteins often migrate slightly differently in SDS-PAGE gels compared to their predicted molecular mass. Rate of migration depends upon the net charge, size and shape and it is generally assumed that upon heat denaturation, each protein will adopt an unstructured, globular structure with bound SDS. Any subtle variation in any of these parameters will (for example the unfolded shape or amount of bound SDS) make them migrate at an apparently different molecular weight (Sharpe, 2003).

The recombinant expression of some proteins is found to be temperature dependent and increased yield of soluble protein can be observed on decrease in expression/incubation temperature from 37° to 30°C (data not shown for 37°C), due to a reduction in the rate of protein synthesis and /or reduction in plasmid copy number thereby leading to the decrease in the concentration of partially folded intermediates prone to aggregation and ultimately to a higher yield of protein in the soluble fraction (Sharpe, 2003).

Here recombinant PNA variants were expressed in 500 ml cultures at 30°C in order to maximize the yield of expression of proteins in the soluble fractions. The soluble fractions were analysed by SDS -PAGE and, in comparison to the expression at 37°C, a much more prominent band was evident at 30kDa, presumed to be recombinant PNA.

The gel was also loaded with LB and a non-transformed *E.coli* cell lysate, as the controls in this experiment, which showed a difference in band migration/distribution in comparison to the crude PNA lysates (data not shown). Bradford assay was used to determine the concentration of protein for each sample and crude protein range was found to be between 1 and 5mg/mls; 10ml crude lysate obtained per 500ml culture.

Protein expression was initially very low due to a number of possible factors including old glycerol stocks having been used to prepare inoculum and/or inefficient cell lysis. The preferred method of lysis involved cell fractionation and sonication. Fresh reagents, newly transformed *E.coli* cells, fresh solutions were therefore prepared in order to ensure expression yields were not compromised by external factors. This resulted in the improved expression yield of the crude lysate proteins.

### **3.3 Purification of recombinant PNA variants.**

#### **3.3.1 Discussion of eluted samples of purification.**

Since the recombinant PNA variants all carried an N-terminal His<sub>6</sub>-Tag, proteins were purified for crude lysate using the Qiagen, 'QIA Expressionist' kit that is specific for binding His-tagged proteins with high specificity whilst reducing nonspecific binding to the column. This kit allowed for the purification of ~5-10mg of crude lysate proteins of about 5mls, thus making it suitable for our samples.



**Figure 3.6.** SDS gel of protein purification

Two sets of purification are shown. The first elution is seen in lanes 2, 5 and 8 and the second in lanes 3, 6 and 9. Prominent bands are evident just below 30 kDa indicative of the recombinant PNA that has been purified. This was one batch of purifications with two rounds of elution's.

Proteins were purified using this kit and figure 3.6 displays the first set of elution's (2, 5, and 8) of each protein, and (3, 6, 9) as the second round of elution of the three proteins. Nonspecific band's' were apparently removed in these steps (second round of elution), as the band's intensity and presence starts to decrease with each wash step.

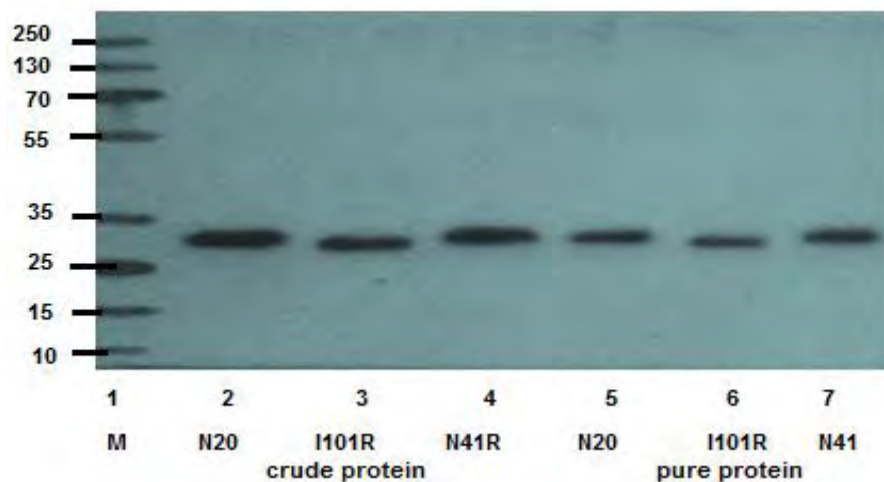
The proteins were subjected to two sets of elutions as mentioned, using the elution buffer as per kit requirements, and although the protocol suggests 4x 0.5ml elution's, the gel above indicates very strong band's after two elution's, with further elution's (not shown above), showing a reduction in nonspecific band's. The imidazole containing elution allows for the purified, His-tagged proteins to be released from the beads. Presence of a dominant band at 30 kDa in each eluted fraction apparently confirmed the previous assumption that PNA migrates aberrantly heavily on SDS-PAGE.

Eluted proteins all ranged from concentrations of 0.5mg/ml to about 2mg/ml and these

could be further concentrated using Vivaspin / Zebaspin desalt or Centriprep columns. These samples were then stored at 4°C for later use.

The elution profile from the column showed minor contamination with other proteins, perhaps due to the column being over-loaded or incorrect ratio of bead to proteins resulting in overflowing/overcrowding.

### 3.4 Western blot analysis.



**Figure 3.7** Lane 1- prestained plus protein ladder. Lane2-4 crude samples wild type, PCES I101R, PCES N41R respectively. Lane 5- 7 elution of purified protein of wild type, PCES I101R, PCES N41R respectively

Western blot analysis of crude and purified protein, using an Anti-His antibody, revealed strong band's at 30 kDa, consistent with the Coomassie Blue stained gels, where a band of ~30 kDa was observed in both, the crude and purified samples, and further confirming the identity of the ~30kDa protein as the recombinant, His –tagged PNA seen in Figure 3.7.

This conclusion was further verified by repeating the western blot using an Anti-PNA

primary antibody followed by a secondary goat anti-rabbit-HRP conjugated antibody, which revealed the same 30kDa band's (data not shown). These purified preparations of wt PNA, PNA (I101R) and PNA (N41R) were deemed of sufficient purity and quantity for further studies on the quaternary structure and specificity of the PNA variants.

## **Chapter 4: Investigation of quaternary structure of Peanut agglutinin**

### **4.1 Native gel electrophoresis/mobility shift assay**

#### **4.1.1 Native vs denaturing gel electrophoresis**

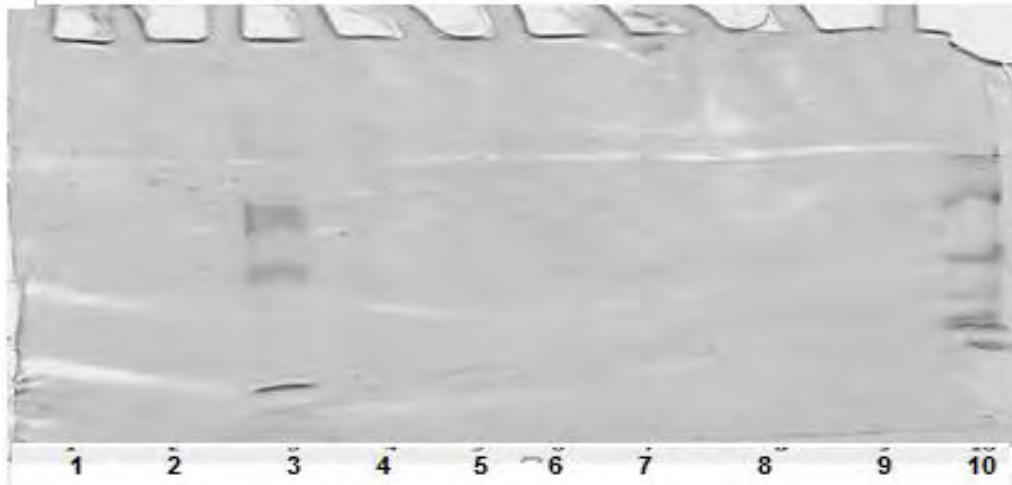
Denaturing gel electrophoresis is a type of electrophoresis in which the native structure of macromolecules and in particular, the quaternary structure is typically not maintained. Thus, quaternary structures of proteins cannot be investigated using this means.

Native PAGE separations however are run in non-denaturing conditions with which secondary, tertiary or quaternary structure is typically maintained. However, native proteins may not separate predictably, since the assumption made in SDS-PAGE regarding the size –based mobility of proteins break down because different proteins of the same mass can have very different shapes and volumes. Unlike denaturing methods, such as SDS-PAGE, native gel electrophoresis does not employ the use of a charged denaturing agent. The molecules being separated therefore differ not only in molecular mass but in intrinsic charge, and experience different electrophoretic forces dependent on the shape of the structure.

Here native gel assays are used to assess the quaternary structure of PNA, given preliminary data from Sharpe (2003) suggesting that PNA forms dimers in solution, not the commonly assumed tetramers (see introduction). In particular we wished to determine whether differing detergent concentrations or differing pH ranges had an impact on the quaternary structure of PNA. The samples loaded included that of a prestained protein ladder, unstained Con A, non- denatured wildtype PNA as well as Guanidine treated wildtype at varying concentrations.

Given the problems associated with accurate sizing of proteins on native gels, we decided to use Con A- a known dimer of similar monomer size to PNA- and a Guanidine- denatured PNA as a control for dimer and monomer mobility on native gels.

However, this assay gave rise to smeared bands or no sample band visibility see Figure 4.1. Despite repeated attempts, this approach proved unsuccessful and was not pursued further.



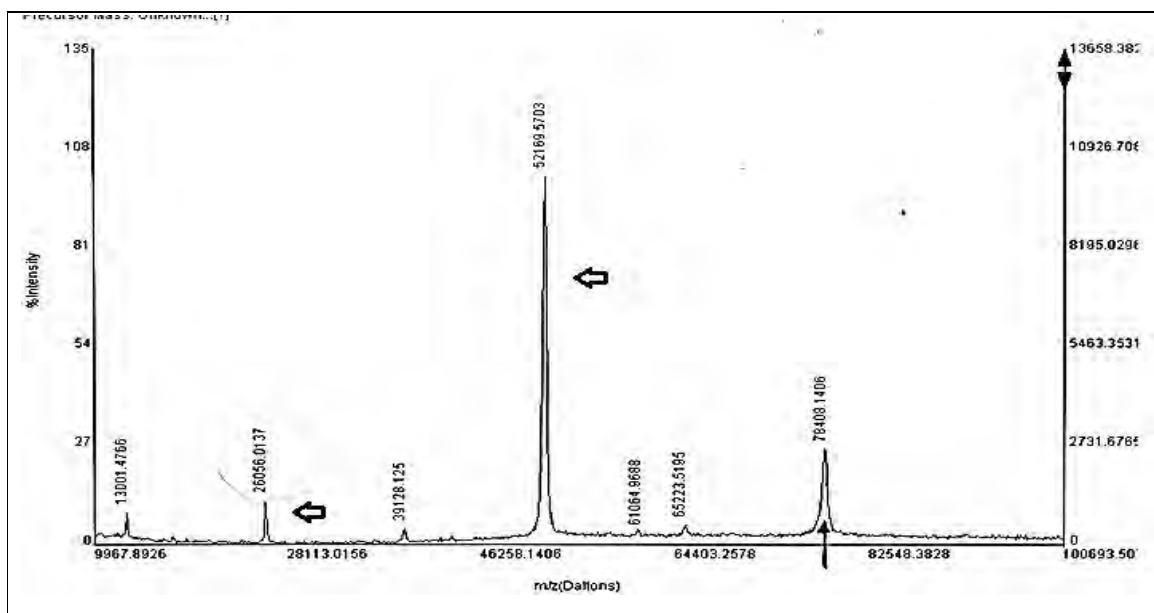
**Figure 4.1.** Native gel electrophoresis. Only visible bands were those of the - unstained and stained ladders. Lane 2-Concanavalin A, Lane 4-untreated wildtype, Lane 5-0.05uM, Lane 6-1uM, Lane 7 -2uM, Lane 8-4 uM, Lane 9-6uM, Lane 9. All prepared at 10ug protein concentrations but no protein bands were evident on the gel. Lane 3 –unstained protein ladder, Lane 1- blank and Lane 10-prestained protein ladder.

## **4.2 Investigation of Multimeric States and quaternary Structure using Mass Spectrometry and High Performance Liquid Chromatography**

### **4.2.1 MALDI Spectrometry**

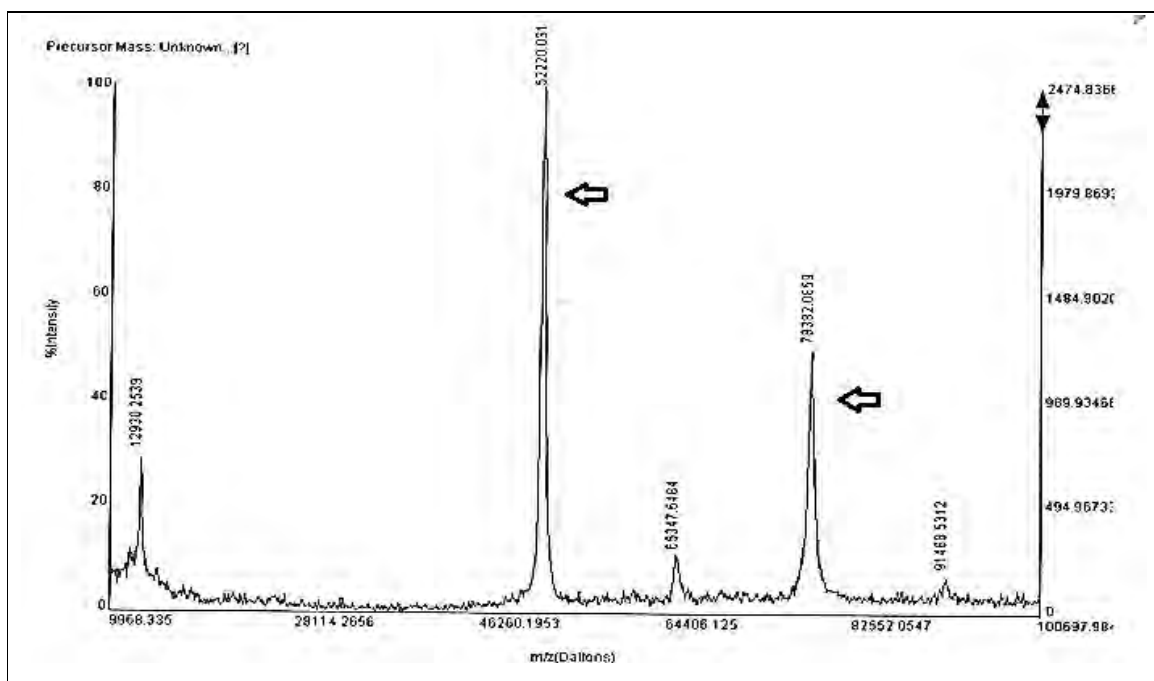
The investigation of the quaternary structure of the PNA was initially performed by means of Matrix Assisted Laser Desorption-Ionisation (MALDI) mass spectroscopy at three different pH levels. Purified wild type PNA was prepared at between 0.001 and 0.01mg/ml in 0.1 M citrate (pH3), 0.1M phosphate or 0.1M Tris (pH9). For each pH, 0.5ul protein solution was mixed with an equal volume of acidified CHCA matrix and spotted onto a MALDI target plot.

Each resultant protein/matrix spot was then analyzed on an AB Sciex 4800 MALDI TOF/TOF mass spectrometer. The resultant Mass spectrometry data revealed the apparent monomers, dimers and trimers at pH 3, dimers and trimers at pH 7, and primarily dimers at pH 9. There was no error with the sample concentrations as there was some variation between concentrations. Future work would make better use of proteins at the same concentrations.



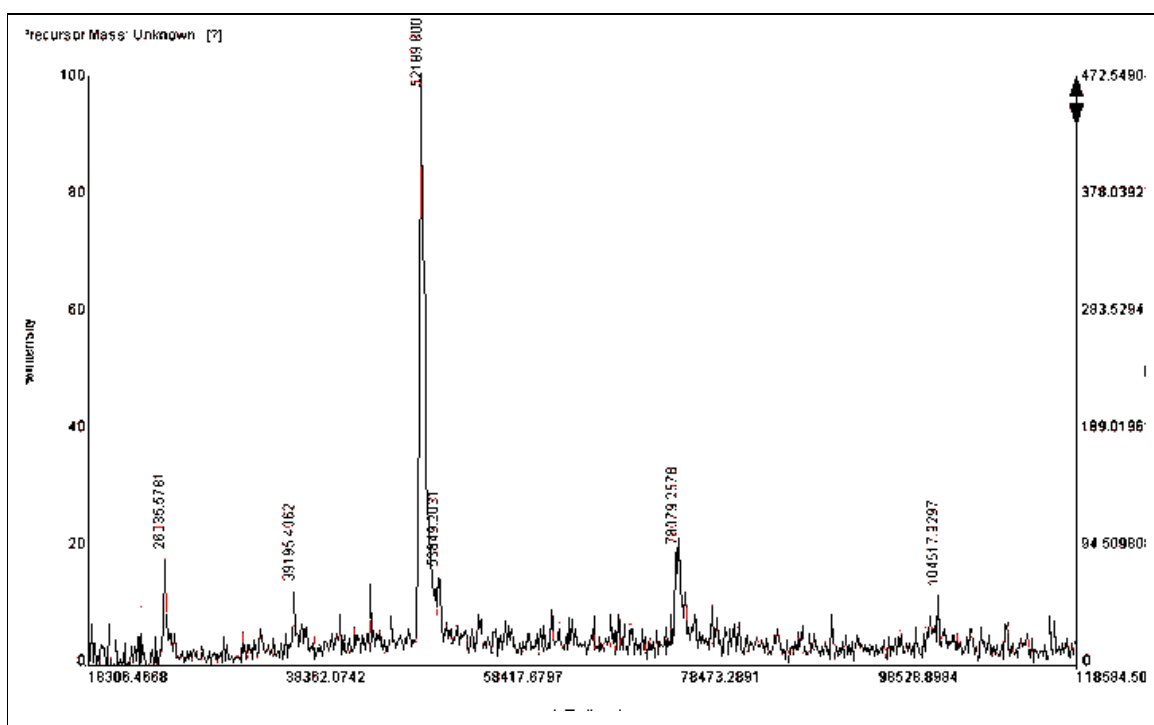
**Figure 4.2** Purified PNA at 0.01mg/ml in 0.1M Citric acid buffer pH 3. Analysed using a saturated CHCA matrix in acidic 30:70 0.1% TFA(pH2): ACN solution.

The predicted exact mass of the wild type His<sub>6</sub>- tagged PNA monomer is 26.143 Da. Following calibrations of the AB 4800 using standards, Figure 4.2 reveals peaks at 26,056.014 Da, 52,180.170 Da, and 78,408.141 Da, which we interpreted as PNA monomer, dimer, and trimer respectively.



**Figure 4.3.** Purified PNA at 0.001mg/ml 0.1M Sodium Phosphate buffer in pH 7. Analysed using a saturated CHCA matrix in acidic 30:70 0.1% TFA (pH2): ACN solution.

Figure 4.3 indicative of predominantly dimeric species compared to pH 3 samples, this pH range shows no evidence of the monomeric species, and strong evidence of an apparent trimer. Figure 4.4 shows a similar trend, although here there is weak evidence for both monomer and trimer species as well as possible evidence for a tetramer at 104,517.930 Da.



**Figure 4.4.** Purified PNA at 0.001mg/ml in pH 9 0.1M Tris HCL buffer. Analysed using a saturated CHCA matrix in acidic 30:70 (0.1% TFA (pH2): ACN solution. Image clarity could not be improved after scan.

**Table 4.1** Illustrating the kDa of each species at different pH ranges

	<b>Monomer</b>	<b>Dimer</b>	<b>Trimer</b>
<b>pH 3</b>	26056.0137	52169.5703	78408.1406
<b>pH 7</b>	N/A	52220.031	78382.0859
<b>pH 9</b>	26035.5781	52169.8008	78079.2578

**Table 4.2.** Monomeric species observed at pH 3, 7 and 9

	<b>pH 3</b>	<b>pH 7</b>	<b>pH 9</b>	<b>Predicted mass</b>	<b>Delta (predict-observed, pH3)</b>
<b>Monomer</b>	26.056.0	N/D	26.035.6	26.143.1	87.1

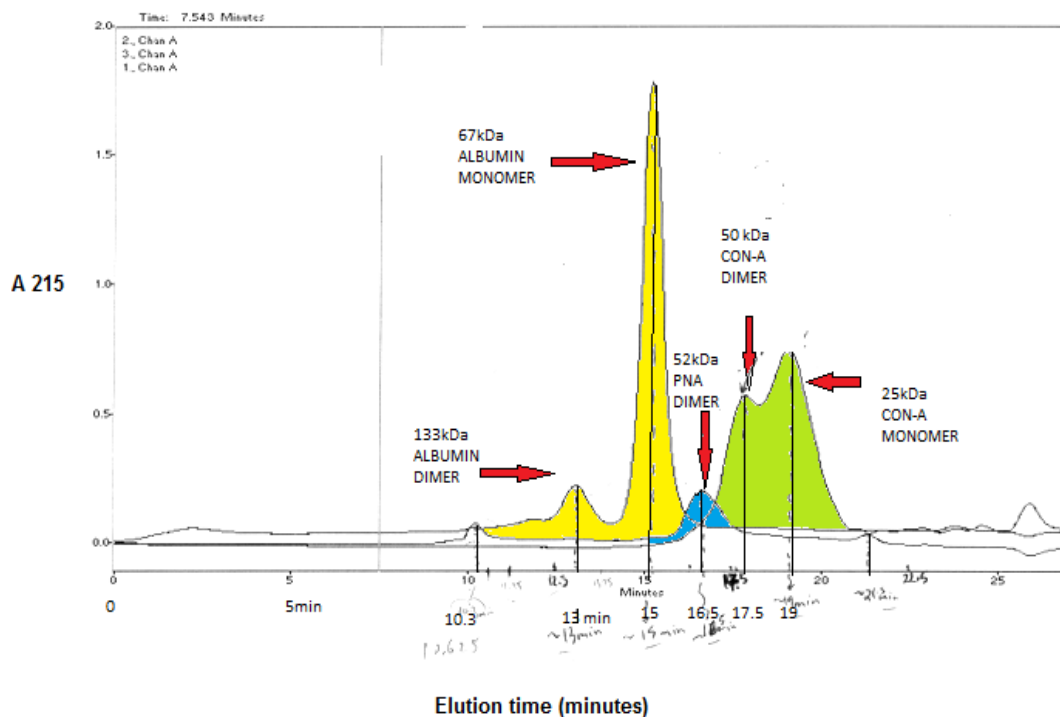
Ordinarily, MALDI mass spectrometry is considered to be a denaturing technique in which protein complexes are not observed. The observation here of a predominantly dimeric PNA species at all pHs is therefore interesting and suggestive of a tightly bound dimer as the dominant solution- phase species of PNA. The observation of an apparent trimeric PNA species by mass spectrometry is perplexing and we do not have a robust explanation for this, but suggest that it might represent an artifact of the co-crystallisation of PNA dimers and monomers with matrix. In the mass spectrometry experiments described here, it is worth noting that, whilst it is difficult to measure pH in a mixed aqueous: organic mixture, theoretical calculations suggest that the addition of an equal volume of 0.1 % TFA (pH 2)/ acetonitrile (30:70 v/v) to a 0.1M buffer will not significantly perturb the pH of the buffer solution. (N.B this equate to mixing equal volumes of 0.1M buffer with ~2.6 mM TFA). (Note, we are ignoring the effect of the organic acid CHCA in the approximations).

Although it is tempting to calculate ratio of monomer vs dimer vs trimer from the areas under the curves under each peak, in reality this is not strictly legitimate because of potentially different ionization affinities for each species. We therefore instead decided to evaluate and quantify the solution-phase quaternary structure of PNA further by analytical size exclusion chromatography.

The most accurate data here exists with the monomeric data. It is clear to see that the mass is relatively consistent in pH 3 and pH 9, but not with the predicted mass of monomer. It is however, consistent with the possible loss of a serine on the C- terminal, where the molecular weight of serine is 87.1 Da

### 4.3 High performance liquid chromatography (HPLC) analysis of PNA quaternary structure.

Size exclusion chromatography was used to further explore the results of the Mass spectrometry experiments, which were apparently in accordance with previous preliminary data of Sharpe (2003). PNA and various controls were independently injected into an analytical size exclusion column at a concentration of ~1mg/ml and the elution profile was monitored at 215nm; the elution time was later converted to an elution volume by reference to the flow rate of 0.5ml/min.

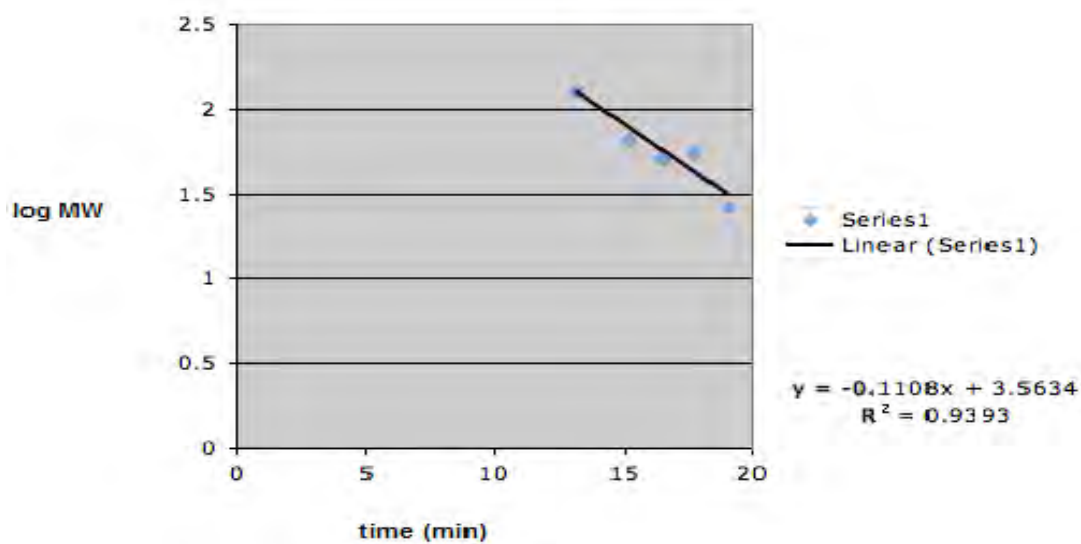


**Figure.4.5** An overlay of chromatography for PNA and two controls, Concanavalin A and BSA. Absorbance at 215nm vs elution time in minutes is seen here.

Figure 4.5 shows the initial chromatogram of wild type PNA overlaid with the chromatogram for Con A and BSA samples. At 15 minutes a peak corresponding to the monomeric form of BSA (~67 kDa) appeared in its dimeric form. At 17.5 minutes a Con A peak was evident and assumed to be dominant Con A dimer (~50 kDa). The PNA (at pH 8) elution peak (~16.5 min) is overlapped with these two control peaks, suggesting a molecular weight of between the 50 and 67 kDa and therefore a dimeric form of PNA. Table 4.3 shows the elution time and estimated molecular weights of PNA and the controls.

**Table 4.3** Elution times of PNA and control samples. Here, their corresponding molecular weights, as well as Log value of these molecular weights can be seen.

<b>Protein standards</b>	<b>Elution time</b>	<b>Molecular weight</b>	<b>Log MW</b>
<b>Void volume</b>	~10.3 min	~170kD-200kDa	2.30
<b>BSA dimer</b>	~13.13	~133	2.12
<b>BSA monomer</b>	~15.25	~67	1.83
<b>PNA dimer</b>	~16.5	~52	1.72
<b>Concanavalin A dimer</b>	~17.25	~50	1.75
<b>Concanavalin A</b>	~19 min	~25	1.425



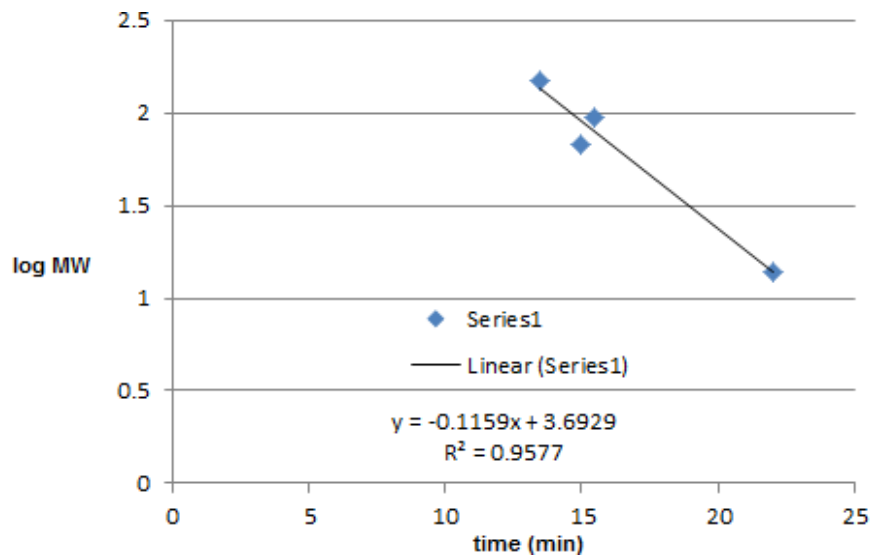
**Figure 4.6** Standard curves for determination of estimated PNA size.

The generation of a standard curve was used to provide an indication of the size of the PNA in the HPLC experiment. This provided an estimated size of the predominant PNA peak is an apparent dimer at pH 8. This standard curve was used for further size approximations of our PNA at different pH ranges in figures/ tables to follow.

**Table 4.4** The standard curve (Con A, BSA and PNA at pH 8). Standard curve equation generated : $y = -0.1108x + 3.5634$

Structure	pH	Time (min)	Elution vol. (ml)	Log (MW)	Mol. Weight (kDa)
<b>Dimer</b>	9	~16.6	~8.3	1.60	~40
<b>Dimer</b>	3	~17	~8.5	1.68	~48
<b>Dimer</b>	7	~17	~8.5	1.68	~48
<b>Monomer</b>	9	~19.68	~9.84	1.38	~24
<b>Monomer</b>	3	~19.68	~9.84	1.38	~24
<b>Monomer</b>	7	~19.68	~9.84	1.38	~24
<b>Tetramer</b>	9	~14.66	~7.33	1.94	~87

Although we expected/predicted to see sizes of ~52 kDa for the dimeric form of PNA, the size estimates here are slightly lower than 52 kDa. Gel filtration does not measure mass of molecule but actually its physical size therefore molecules don't necessarily migrate at precisely exactly expected molecular weights. The fact that we're observing it at a lower mobility of ~40 is consistent with a dimer species. It is also important to bear in mind that although these approximate times are used, because of a lack of tools to export the data and to acquire proper size estimates, the molecular weights seen here are still lower than "tetramers" and higher than "monomers". It can then be assumed that these values are, or rather these estimates are the assumed dimers. Table 4.4 shows the molecular weights calculated by means of using the formula. Using a different calibration curve, after the column was recalibrated, shows a slightly more accurate result.



**Figure 4.7.** Standard curve using different proteins. IgG, Lysozyme Phosphorylase B, and BSA. These known protein sizes were used to generate this standard curve.

**Table 4.5** Standard curve data for re calibrated column.  $y = -0.1159x + 3.6926$ : equation generated after re calibration of column.

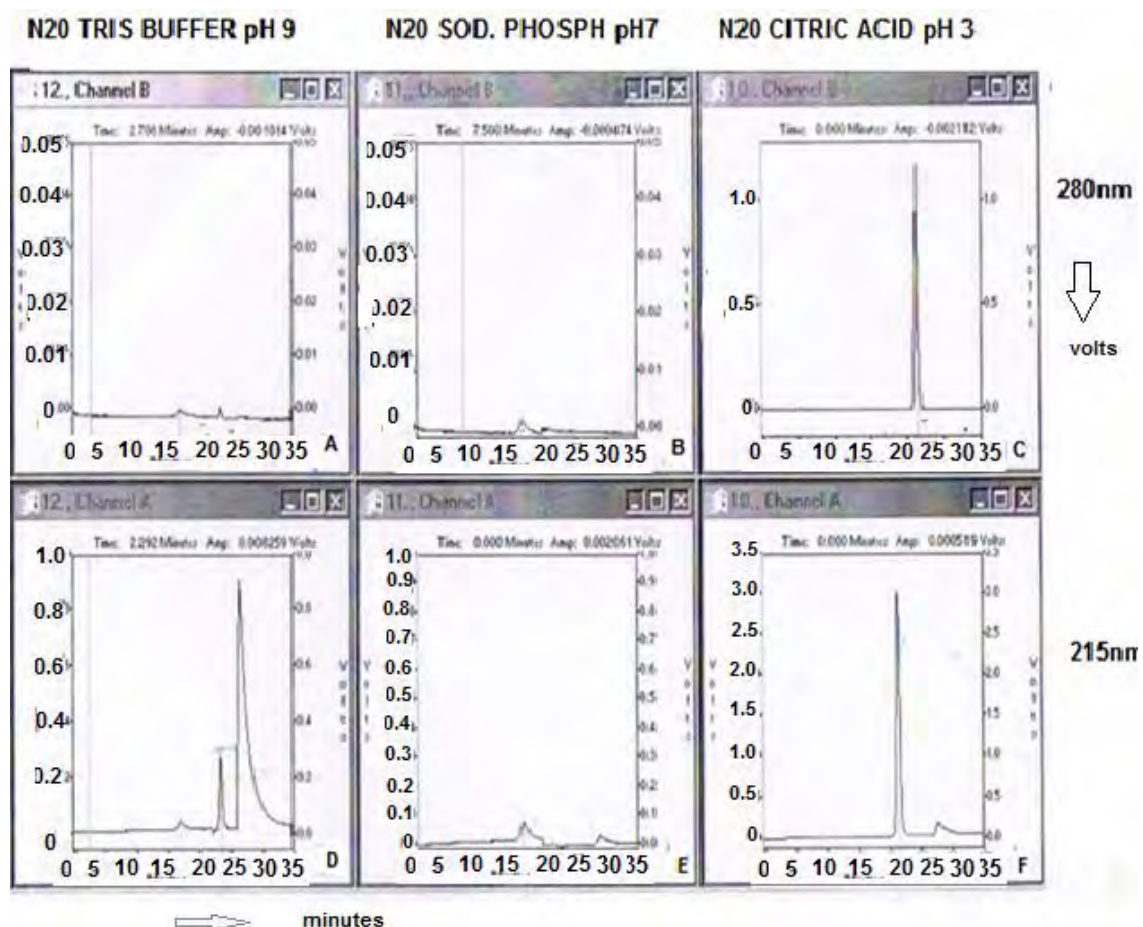
	<b>Time (min)</b>	<b>Elution vol</b>	<b>Log MW</b>	<b>MW</b>
<b>IgG</b>	13.5	6.75	2.17	150
<b>Phosph B</b>	15.5	7.95	1.98	97
<b>BSA</b>	15	7.5	1.82	67
<b>Lysozyme</b>	22	11	1.146	14

**Table 4.6** pH range calculations using new standard curve.  $y = -0.1159x + 3.6926$

<b>Structure</b>	<b>pH</b>	<b>Time (min)</b>	<b>Elution vol.</b>	<b>Log mw</b>	<b>Mol. Weight (kDa)</b>
<b>Dimer</b>	9	~16.6	~8.3	1.76	~57.5
<b>Dimer</b>	3	~17	~8.5	1.72	~52.4
<b>Dimer</b>	7	~17	~8.5	1.72	~52.4
<b>Monomer</b>	9	~19.68	~9.84	1.41	~25.7
<b>Monomer</b>	3	~19.68	~9.84	1.41	~25.7
<b>Monomer</b>	7	~19.68	~9.84	1.41	~25.7
<b>Tetramer</b>	9	~14.66	~7.33	1.99	~97.7

Table 4.5 and 4.6 show the predicted molecular weights of the PNA at varying pH ranges, based on a new equation following the re calibration of the column in order to investigate whether the sizes that were previously observed in table 4.4 was in fact the correct sizes. The curves are based on a different standard curve using proteins IgG, BSA, Lysozyme and Phosphorolase B. This suggested a molecular weight closely related to the expected size of the dimeric protein (Sharpe, 2003). This curve allowed

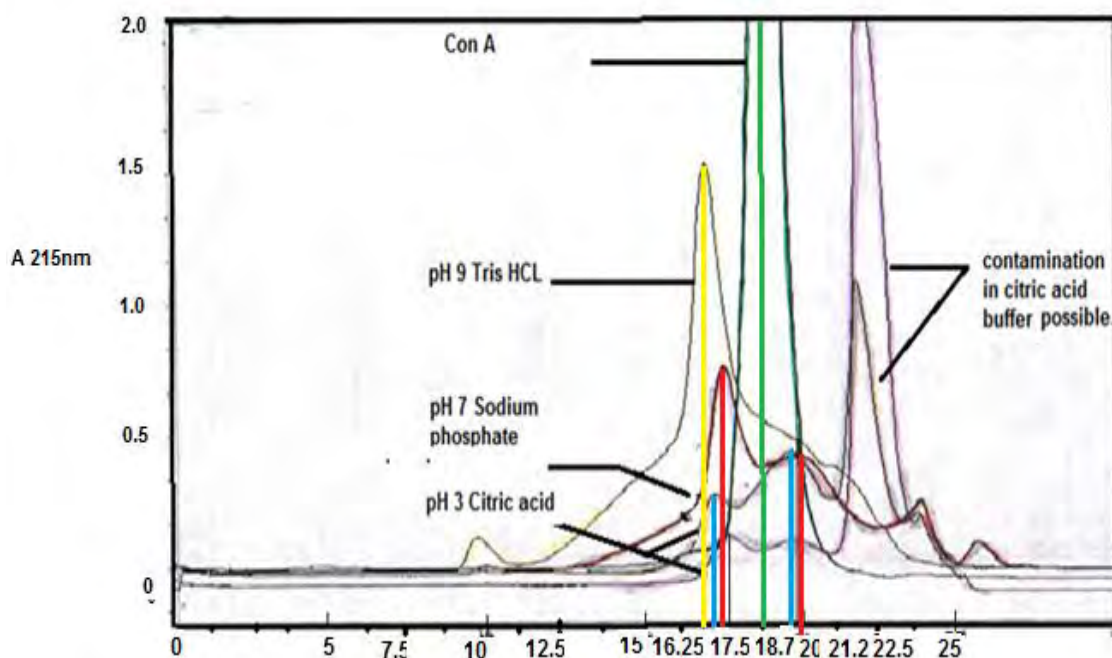
for a much better estimation of size range of all species of protein. Once again it must be emphasized that gel filtration gives an estimation and both curves are within experimental error. The use of these four points was preliminary, as this was intended to be a range finding experiment. Typically four points are not as good as ten points and this could be included in future experiments.



**Figure 4.8.** pH dependency of PNA elution profile from a size exclusion column. Block A and B are the 215/280 traces of pH 3 sample, block C and D sample at pH7 and. Block E and F sample at pH 9.

In order to explore the pH dependency of the PNA quaternary structure, various PNA samples (1mg/ml) were buffer exchanged into 0.1M sodium citrate (pH3), 0.1M sodium phosphate (pH7) or 0.1M Tris (pH9) and then analysed by size exclusion chromatography see Fig 4.8. In each case, the column was pre equilibrated in the relevant pH buffer and that buffer was also used as the mobile phase during elution.

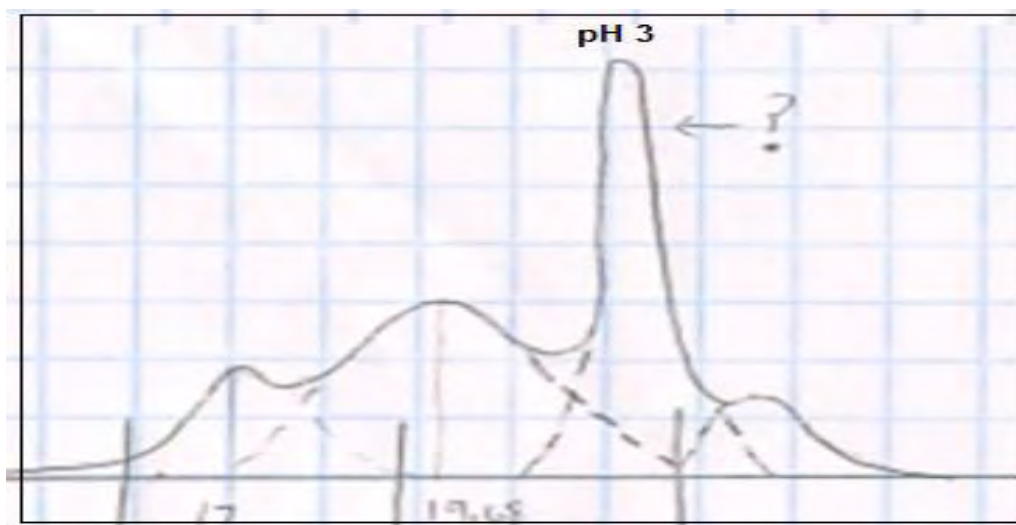
This figure depicts the initial test to determine whether there would be any traces appearing on the HPLC. This data is unclear as the run was not a complete success, thus requiring the experiment to be repeated in future.



**Figure 4.9** Shows an overlay of the PNA elution profiles at the three different pH's. Showing HPLC output data of the wild type sample showing the overlapping traces of pH 3 (blue), pH 7 (red) and pH 9 (yellow). Con A (green) was used as a control sample in this experiment.

Con A sample in solution was a dimeric species. A tetrameric form of Con A would not sit well on the HPLC column. This Con A was prepared at pH 8 in 0.1 M sodium phosphate buffer and used as a control. The data above suggests the likely presence of dimeric as well as monomeric forms of PNA. The presence of largely dimeric PNA is seen at pH 9 whereas in pH 7 and pH 3 there appear to be a large presence of monomeric PNA (see area under curve, Figure 4.10, 4.11 and 4.12). In order to investigate whether the large peak eluting at 22.5 minutes at pH 3 reflected some acid-catalysed autolysis of PNA, equivalent samples of PNA were pre incubated at either pH 9 or pH 3 and then analysed on SDS PAGE. Figure 4.9 and 4.13 is an overlay of the peaks found at the varying pH ranges thus making the only relevant information, the

time overlaps between the samples. Figures 4.10, 4.11 and 4.12 were included to show area under curve estimations. This was only because the program could not export the data into any other format (i.e. excel) and could only be printed. Area under the curve information had to be gathered manually.



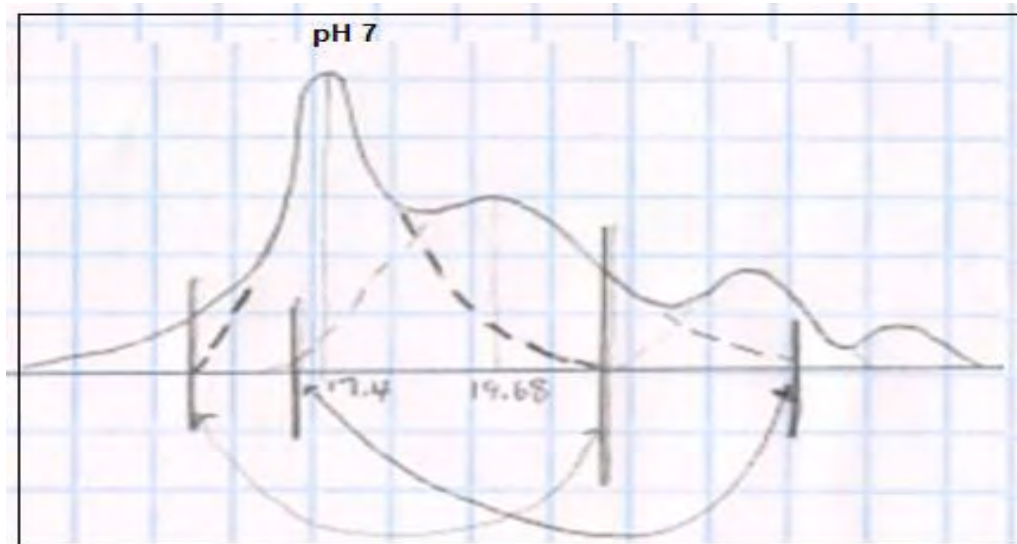
**Figure 4.10** pH 3 citric acid buffer data plotted onto graph paper. The area under curve data was manually calculated from this image.

**Table 4.7** Fraction of each quaternary structure of wt PNA at pH 3.

pH3	# blocks	Fraction %
Tetramer	/	/
Dimer	~ 3	0.315
Monomer	~ 6.5	0.68
Total	~9.5	1.0

Figure 4.10 and Table 4.7 indicate the presence of both dimeric and monomeric species of PNA at pH3. The images presented in figures 4.10, 4.11 and 4.12 could be imported from the HPLC program as there was no digitized format/program to interpret this image. The use of graph paper to plot and calculate the amount/ fraction of each species, was because of the absence of an available tool to measure/ calculate the data efficiently as the HPLC output only allowed the data to be printed and not exported into

a program to calculate. The presence of a peak eluting off the column between ~20 and 22.5 minutes was puzzling and required some investigation in order to determine its identity/cause. This is addressed in figure 4.14.

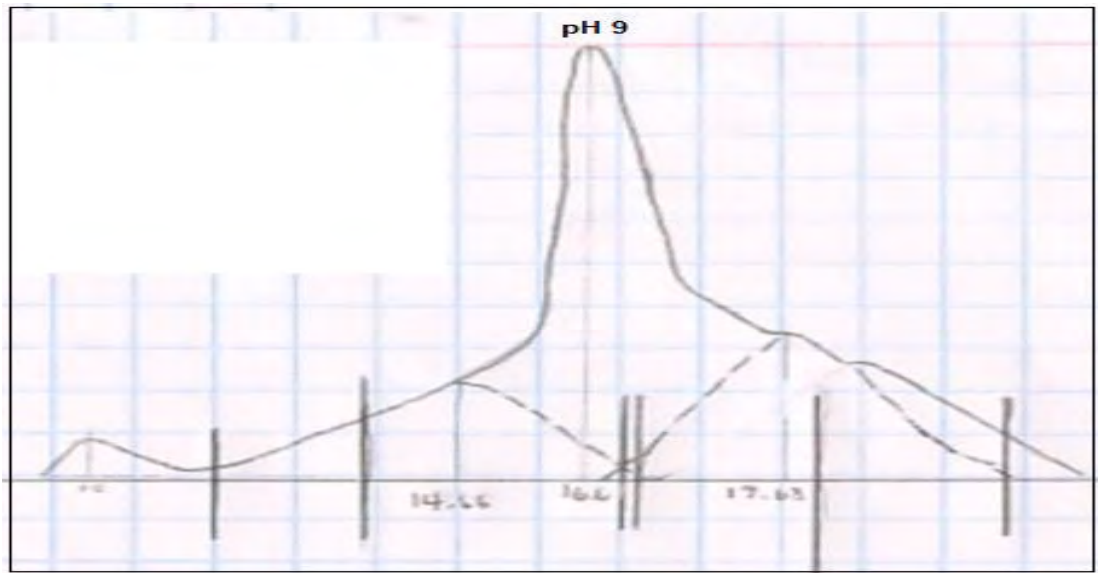


**Figure 4.11** Graphical representation of pH 7 sodium phosphate buffer data. The area under the curve was manually calculated from this curve as is seen in 4.10.

**Table 4.8** Fraction of each quaternary structure of wt PNA at pH 7.

pH 7	# blocks	Fraction %
Tetramer	/	/
Dimer	~9.5	0.55 %
Monomer	~8.5	0.45
Total	18	1.0

Figure 4.11 and Table 4.8 represents the fraction of species upon eluting off the HPLC column in pH 7. The presence of a nearly equal percentage of monomer to dimer exists.

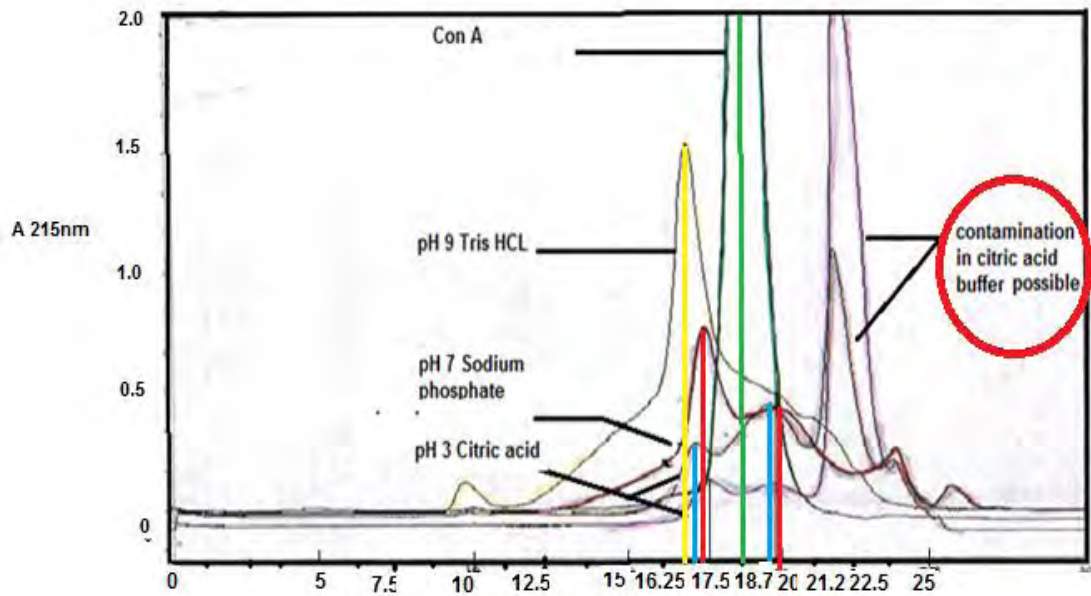


**Figure 4.12** Graphical representations of pH 9 Tris HCL buffer data. Area under the curve data was also manually calculated as seen in figure 4.10 and 4.11.

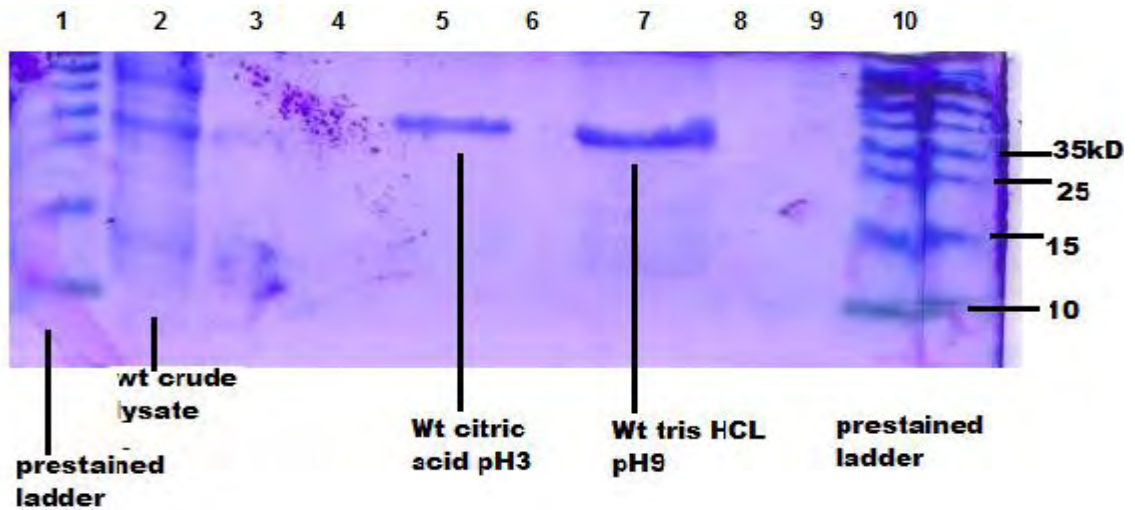
**Table 4.9** Fraction of each quaternary structure of wt PNA at pH 9.

<b>pH 9</b>	<b># blocks</b>	<b>Fraction %</b>
Tetramer	~4.5	0.166
Dimer	~ 16.0	0.59
Monomer	~6.5	0.24
Total	27	1.0

Data represented in figure 4.12 and table 4.9, is that of protein prepared at pH 9 with fractions displayed for each species. pH 9 data suggests the presence of predominantly dimeric species for this range. There exists less than 20 % of tetrameric form and nearly 25 % of monomeric protein. In order to investigate whether the large peak eluting at 22.5 minutes at pH 3 reflected some acid catalyzed autolysis of PNA, equivalent samples of PNA were pre-incubated at either pH 9 or pH 7 and then analysed by SDS PAGE.



**Figure 4.13** pH range comparison. This overlay again, highlighting the varying pH ranges but, specifically the presence of the smaller MW peaks of unknown identity.

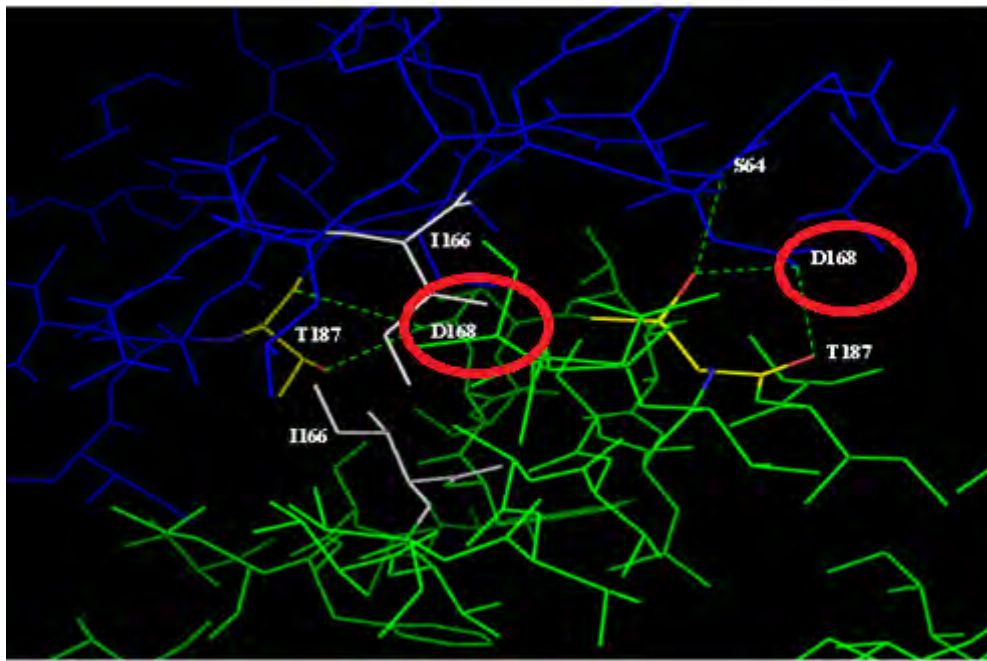


**Figure 4.14** SDS PAGE gel to determine presence of contamination. This gel was used to determine if the 22.5 minute peak in (Figure 4.13) was contamination of the pH 3 samples or degraded protein.

This SDS-PAGE gel revealed no other bands at pH 3 or pH 9 except the expected ~30kDa band. We therefore assumed that the aberrant 22.5 minute peak reflected a contamination in the citric acid buffer. A fresh batch of citrate buffer had been prepared for this gel as this experiment was run much later than the initial HPLC analysis since the previous batch of citric acid buffer had been discarded.

From the HPLC data shown above, a number of observations can be made.

At pH 9, the recombinant PNA used in this study is predominantly a dimeric species. The broadening of the main peak at both higher and lower elution times may suggest that a relatively rapid equilibration between dimer and either monomer or tetramer, but if this is true, the monomer and tetramer remain minor species at pH 9. Reasons for the discrepancy between the solution phase assays here and crystallographic data (which suggest PNA is a tetramer) remain unclear at present, but may reflect an artifact of the crystallization process. At pH 7, the proportion of PNA monomer increases, but the dimer is still the dominant species (see table 4.8). At pH 3, PNA is now observed predominantly in the monomeric form. This pH dependent switch from dimer to monomer between pH 7 and pH 3 is partially consistent with the protonation of a carboxyl acid residue in the dimer interface. One such candidate carboxylic acid residue is D168 in Figure 4.15, found in both monomers in the dimer interface. However, verification that this residue is indeed the culprit would require a site-directed mutagenesis study (e.g. to study a PNA (D 168N) mutant that could not be similarly protonated/ de protonated as a function of pH) that is beyond the scope of this Master's thesis. Interestingly though, D168 makes a number of hydrogen bonding interactions in the dimer interface and it would be reasonable to suppose that these might be disrupted on protonation at pH 3, thus potentially weakening the dimer interface and favoring monomer instead.



**Figure 4.15** The carboxylic acid residue at D168. Circled in red are the D168 residues found in both monomers in the dimer interface. (Swiss PDB viewer).

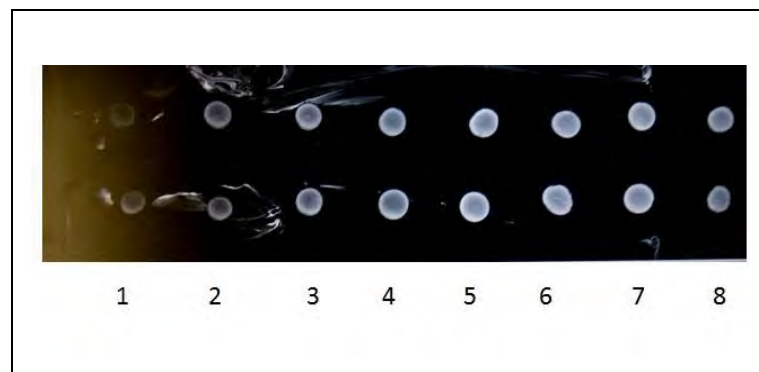
Interestingly, plant vacuoles, where PNA is thought to be processed into its mature form and stored, have an acidic pH so the data shown here may prove to have some strong biological relevance.

## Chapter 5: Affinity binding assays

Affinity binding assays were performed in order to establish whether peanut agglutinin was able to bind to streptavidin coated slides, derivitised with biotinylated forms of Lactose, N-acetylamine and the T-antigen.

### 5.1 Streptavidin coating of slides

Figure 5.1 shows an NHS-activated, PEG-coated glass slides successfully coated with streptavidin using the method described (section 2.2.6). The replicate spots seen above labeled 1 through 8 are Cy5-biotinylated BSA spotted across each of 16 wells (1 spot per well; 1ul/spot). The scanned image of this hand spotted slide showed that the slides were efficiently coated with streptavidin and that the Cy5-biotinylated BSA then bound efficiently to the slides. Figure 5.1 thus confirms that, whilst improvements in the uniformity of streptavidin coating could be achieved in principle, these streptavidin – coated slides should already be good enough for qualitative assessment of the binding of PNA variants to immobilized, biotinylated sugars.

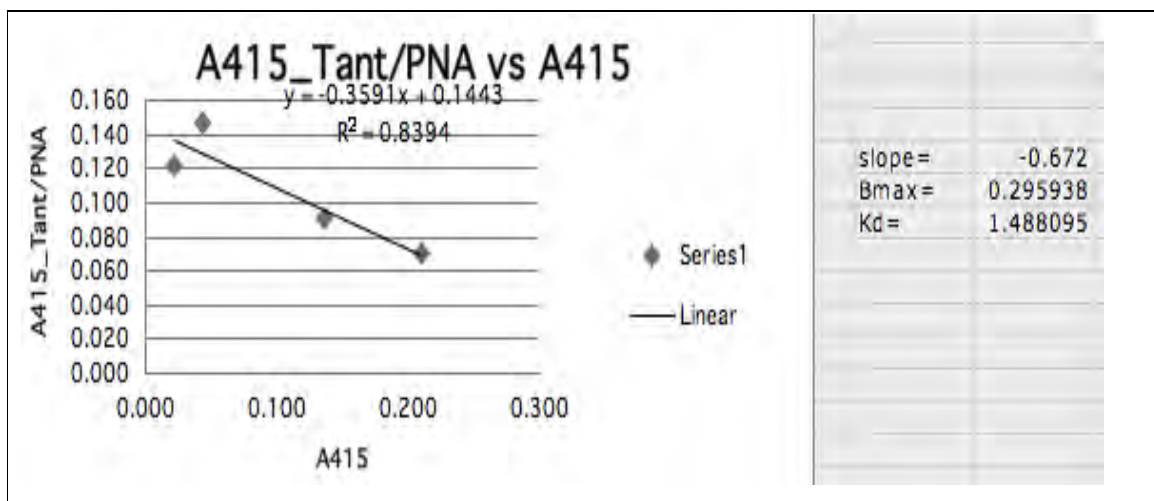


**Figure 5.1** Hand spotted slide. Cy5 Biotinylated BSA spotted across a glass slide.

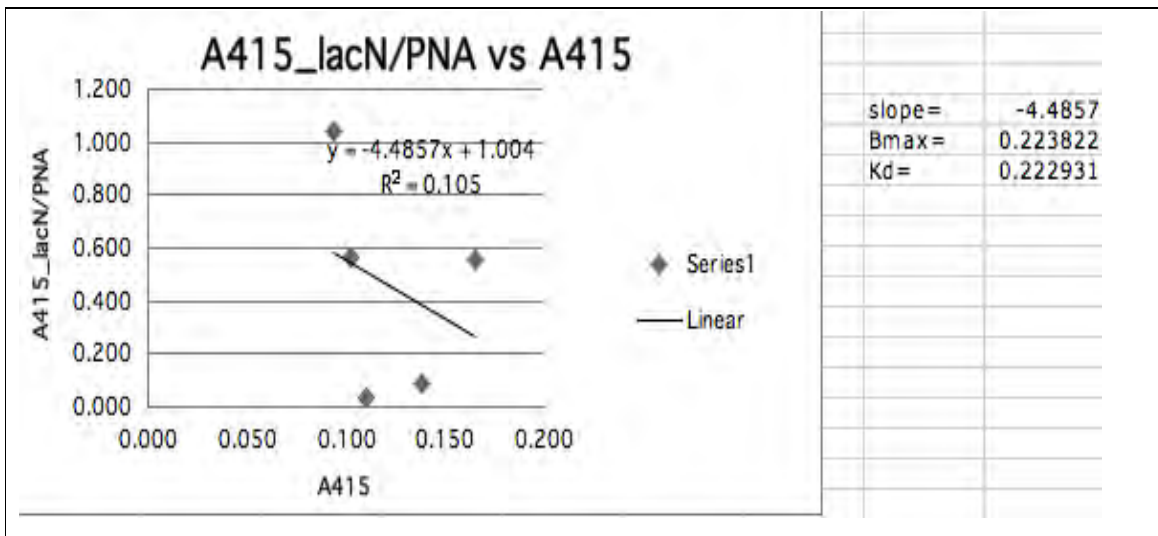


The use of two antibody concentrations and two incubation times in Figure 5.2 gave an indication as to which antibody concentration and incubation time would be best suited to the assay. Upon comparing the two antibody concentrations, the use of the 1uM concentration of proteins seemed to display very poor signal. The 100nm side of the slide also showed the highest variation in foreground as a function of PNA concentration and antibody concentration. However, this data did not appear to follow a logical trend and moreover proved poorly reproducible, with several replica slides showing only fluorescence of BSA and/or high backgrounds. Despite repeated attempts, it did not prove possible to create a reproducible, quantitative microarray based assay for PNA binding to immobilized sugars. Previous colleagues encountered a similar problem and it was suspected that the sugars were being masked underneath the PEG surface very quickly before assaying thereby making binding impossible. We switched to an ELISA format instead to try and establish whether there was a problem with the sugars and/or proteins as well.

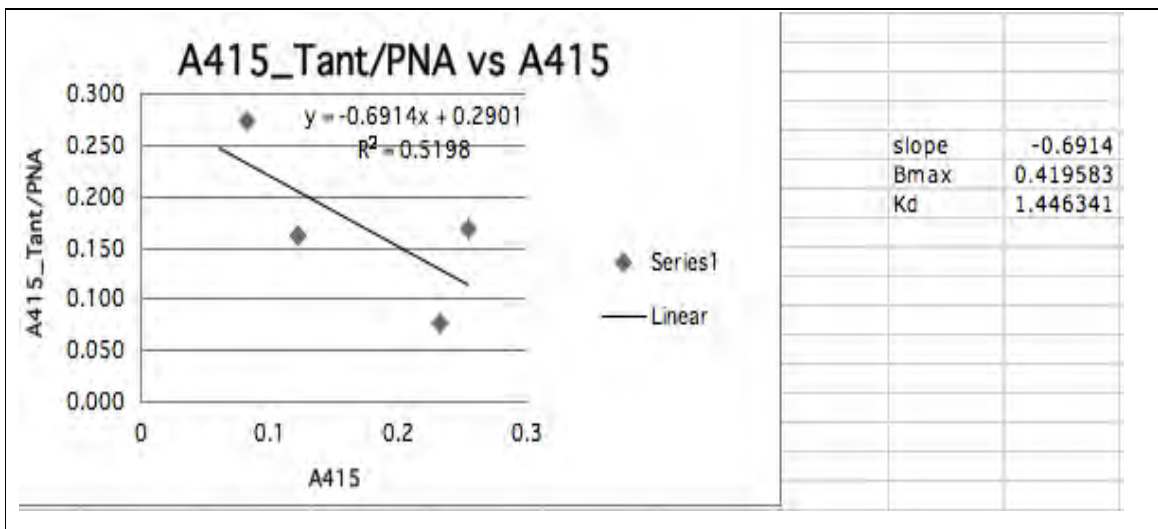
#### 5.4 ELISA enzyme linked immunosorbent assay



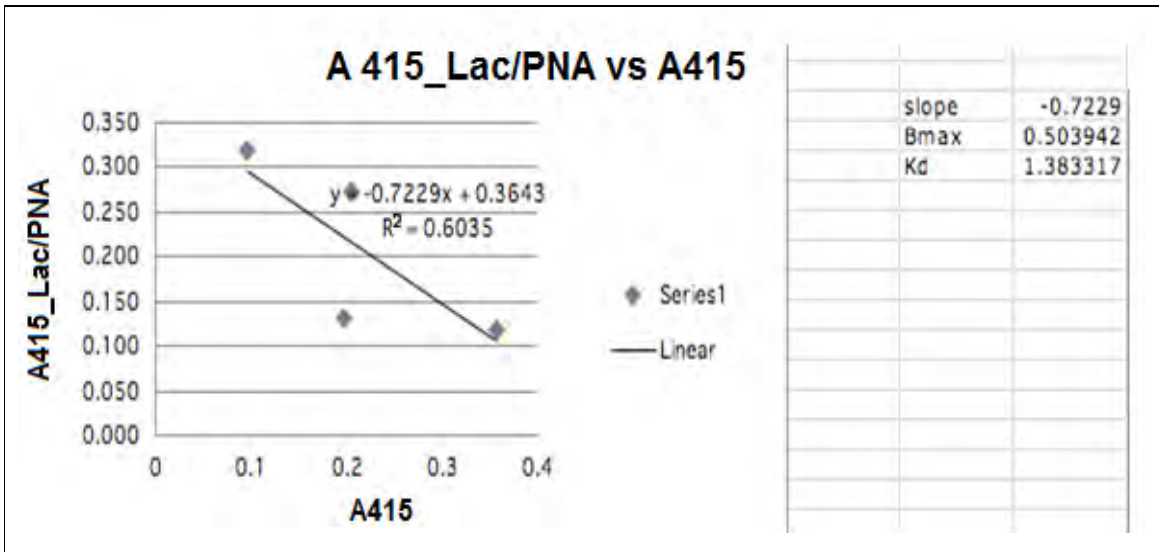
**Figure 5.3** Standard curve of wt PNA binding to T-antigen.



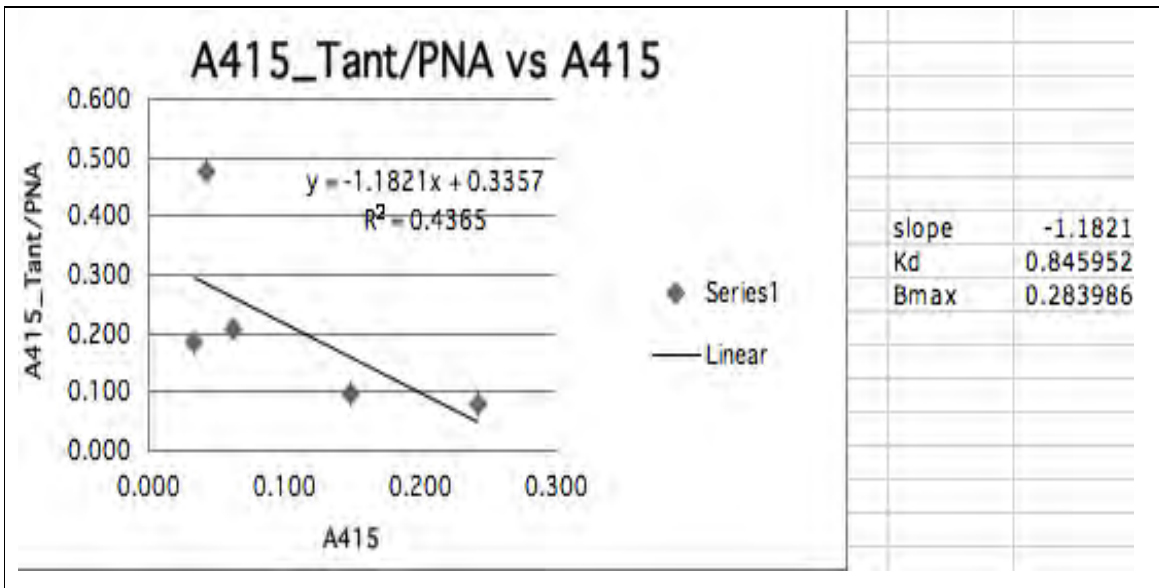
**Figure 5.4** Standard curve of wt PNA binding to LacNAc



**Figure 5.5** Standard curve of I101R PNA binding to T-antigen.



**Figure 5.6** Standard curve of I101R PNA binding to Lactose.



**Figure 5.7** Standard curve of N41R PNA binding to T-antigen.

**Table 5.1** Showing resultant ELISA data. By comparison the SPR data generated by Sharpe (2003) similar binding affinities are observed of wt PNA to the T-antigen, but less apparent distinctions between the wt PNA and I101R or N41R variants. This data suggests a modest decline in affinity.

<b>Carbohydrate (Biotinylated)</b>	<b>N20 (<math>K_D</math>)</b>	<b>I101R (<math>K_D</math>)</b>	<b>N41R (<math>K_D</math>)</b>
<b>T-antigen</b>	1.5 $\mu$ M	1.4 $\mu$ M	0.8 $\mu$ M
<b>LacNAc</b>	N/D	N/D	N/D
<b>Lac</b>	N/D	1.4 $\mu$ M	N/D

**Table 5.2** Predicted affinities of the wt PNA, N41R PNA, I101R PNA for the T-antigen using the bivalent analyte model (Sharpe, 2003).

<b>PNA species</b>	<b><math>k_{a1}</math> (1/Ms)</b>	<b><math>k_{d1}</math> (1/s)</b>	<b><math>K_D</math> (<math>\mu</math>M)</b>	<b><math>\chi^2</math></b>
wt PNA	4.49x10 <sup>3</sup>	2.72x10 <sup>-3</sup>	0.6 $\mu$ M	18
N41R PNA	2.86x10 <sup>3</sup>	5.91x10 <sup>-3</sup>	2.1 $\mu$ M	3.77
I101R PNA	3.5x10 <sup>3</sup>	1.46x10 <sup>-2</sup>	4.2 $\mu$ M	0.78

**Table 5.3.** Predicted affinities of the wt PNA and N41R PNA, for lactose using the bivalent analyte model (Sharpe, 2003).

<b>PNA species</b>	<b><math>k_{a1}</math> (1/Ms)</b>	<b><math>k_{d1}</math> (1/s)</b>	<b><math>K_D</math> (<math>\mu</math>M)</b>	<b><math>\chi^2</math></b>
wt PNA	9.43x10 <sup>3</sup>	1.81x10 <sup>-2</sup>	1.9 $\mu$ M	0.84
N41R PNA	1.80x10 <sup>3</sup>	3.31x10 <sup>-2</sup>	18.4 $\mu$ M	1.25

By comparison to the SPR data generated by Sharpe (2003) (see table 5.2 and 5.3) we

observed similar binding affinities of wt PNA to the T-antigen (Figure 5.3), but less apparent differences between the wt PNA and I101R or N41R variants (Figure 5.5 and 5.7). This data suggests a modest decline in affinity. Despite our best efforts to measure binding, the wt and the N41R mutant, was unable to provide decent binding affinities and no real discerning results by way of binding for the Lactose and LacNAc.

The resultant data is not of the best quality but the data can still be interpreted with caution. Our data suggests that there is not such a wide range of binding affinity for T antigen whereas data from Sharpe, 2003 shows about 7/8 fold affinity whereas our data shows about a 2 fold affinity. Our data also suggests binding of I101R PNA to lactose (Figure 5.6), whereas her work suggests no binding to Lactose for this mutant. Further work is needed here to confirm our data. Without having repeated SPR analysis, we are unable to verify whether the results obtained were due to technical failures/issues or if it is a genuine result. This experiment requires further investigation.

The ligand binding model used to plot A415\_sugar/ (PNA) vs A415 here was based on a 1:1 binding model generated from our Scatchard plot (see appendix, ligand binding assay). The fit of the data is based on the R squared values. Importantly though, whilst the ELISA data generated here is incomplete and the quality of the data not as high as it could ideally have been, we did observe here, binding of the I101R variant to the immobilized lactose, in contrast to the SPR data of Sharpe (2003) where no binding of this variant to lactose was observed. This suggests that further studies are needed to resolve the differences between the ELISA and previous SPR; in the absence of an SPR device locally, it was not possible to conduct this comparison here. However, our data does suggest that the dramatically increased specificity of the I101R variant for T-antigen based on SPR data may in fact not be true and further studies are clearly warranted.

## 5.5 Discussion

By means of expression and purification of the proteins, we were able to perform two experiments by which to address our initial experimental aims of the thesis:

1. To provide further experimental evidence for altered specificity of three variant forms of PNA that originated from the work of Sharpe (2003). Our experiments were aimed to address this specificity seen in previous work and showed some evidence of similarity to what was seen in that the PNA variants in that it showed no evidence of binding towards LacNAc. However, these results were interpreted with caution as the data was not of the best quality data. This could be due to differences in the technologies used, but without having repeated this experiment using SPR technology, it is impossible to verify whether the results obtained were true compared to her data. There could very well have been some technical failure on either side or it could be a genuine result generated. Further investigation on this is required to verify the results.

This work did show binding affinity towards to T-antigen as previously seen (Sharpe, 2003), but the fact that binding of I101R was observed for lactose needs to be further investigated as previous data shown by Sharpe, (2003) showed no binding of this mutant towards lactose.

2. To determine the multimeric state of PNA under varying experimental conditions. The experiments used to discriminate between the multimeric states of PNA under altered pH conditions were through the use of HPLC. At pH 9 it is seen that the recombinant PNA used in this study is predominantly a dimeric species. The broadening of the main peak at both higher and lower elution times indicate relatively rapid equilibration between dimer and either monomer or tetramer, but at this time, the monomer and tetramer remain minor species at pH 9. Reasons for the discrepancy between the solution phase assays here and crystallographic data (which suggest PNA is a tetramer) remain unclear at

present, but may reflect an artifact of the crystallization process.

At pH 7, the proportion of PNA monomer increases, but the dimer is still the dominant species (see table 4.8). At pH 3, the PNA is now observed predominantly in the monomeric form. This pH dependent switch from dimer to monomer between pH 7 and pH 3 is partially consistent with the protonation of a carboxylic acid residue in the dimer interface and one candidate could be the carboxylic acid residue at D168 in Figure 4.15, found in both monomers in the dimer interface. In order to verify that this residue is indeed the culprit would require a site-directed mutagenesis study but that is beyond the scope of this Master's thesis. Interestingly though, D168 makes a number of hydrogen bonding interactions in the dimer interface and it would be reasonable to suppose that these might be disrupted on protonation at pH 3, thus potentially weakening the dimer interface and favoring monomer instead. Interestingly, plant vacuoles, where PNA is thought to be processed into its mature form and stored, have an acidic pH so the data shown here may prove to have some strong biological relevance.

## **5.6 Conclusion**

In conclusion this work describes the use of the PNA variants in order to increase the selectivity of binding to the T-antigen. It also aims to confirm work done by Sharpe, (2003) where her work described the presence of primarily predominant dimeric species instead of tetrameric species as seen in literature. These findings need to be elaborated on in order to successfully use these variants for specifically selecting for the T-antigen. Looking at the aims, we found:

1. Both mass spec and HPLC confirm PNA is primarily a dimer in solution. These findings are based on the use of using neutral pH environments as well as varying pH ranges.
2. HPLC analysis confirms that the multimeric state of PNA in solution is pH-dependent and suggests D168 as a candidate residue mediating this.

3. Contrary to previous SPR data, PNA (I101R) does bind lactose and with similar affinity to T-antigen. Further studies therefore needed to explain why this variant was possibly selected through *in vitro* evolution experiments.

## References

1. Adhikari, P., Bachhawat-Sikder, K., Thomas, C. J., Ravishankar, R., Jeyaprakash, A. A., Sharma V., Vijayan, M. & Surolia, A. (2001). Mutational analysis at Asn-41 in peanut agglutinin. *J. Biol.Chem.* 276, 40734-40739.
2. Baldus, S. E., Zirbes, T. K., Hanisch, F.-G., Kunze, D., Shafizadeh, S. T., Nolden, S., Mönig, S. P., Schneider, P. M., Karsten, U., Thiele, J., Hölscher, A. H. & Dienes, H. P. (2000). Thomsen Friedenreich antigen presents as a prognostic factor in colorectal carcinoma. *Cancer.* 88, 1536-1543.
3. Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A. & Vijayan, M. (1996). Conformation, protein-carbohydrate interactions and a novel subunit association in the refined structure of peanut lectin-lactose complex. *J. Mol. Biol.* 259, 281-296.
4. Boyd, W.C. and Shapleigh, E. (1954) Specific precipitation activity of plant agglutinins (lectins). *Sci*, 119, 419-422.
5. Campbell Campbell, B. J., Finnie, I. A., Hounsell, E. F. & Rhodes, J. M. (1995) Direct demonstration of increased expression of Thomsen-Friendenreich (TF) antigen in colonic adenocarcinoma and ulcerative colitis mucin and its concealment in normal mucin. *J. Clin. Invest.* 95. 571-576.
6. Chelvanayagam, G., Heringa, J., & Argos, P. (1992). Anatomy and evolution of proteins displaying the viral capsid jellyroll topology. *J. Mol. Biol.* 228, 220-242.
7. Cook, G. M. W. (1986). Cell surface carbohydrates: Molecules in search of a

- function. *J. Cell Sci. Suppl.* 45-70, Suppl. 4.
8. Covell, D. G., Smythers, G. W., Gronenborn, A. M. & Clore, G. M. (1994) Analysis of hydrophobicity in the alpha-chemokine and beta-chemokine families and its relevance to dimerisation. *Protein Sci.* 3, 2064-2072.
  9. Chandra, N. R., Prabu, M. M., Suguna, K. & Vijayan, M. (2001). Structural similarity and functional diversity in proteins containing the legume lectin fold. *Protein Eng.* 14, 857-866.
  10. Dev, S., Khan, R. H., & Surolia, A. (2006). 2,2,2-Trifluoroethanol-Induced structural change of peanut agglutinin at different pH: A comparative account. *IUBMB life*, 58(8), 473-479.
  11. Dev, S., & Surolia, A. (2006). Dynamic light scattering study of peanut agglutinin: size, shape and urea denaturation. *Journal of biosciences*, 31(5), 551-556.
  12. Drickamer, K. (1995). Multiplicity of lectin-carbohydrate interactions. *Structural Biology.* 2 (6). 437-439.
  13. Drickamer, K. (1997). Making a fitting choice: common aspects of sugar-binding sites in plant and animal lectins. *Structure.* 5. 465-468.
  14. Drickamer, K. (1999). C-type lectin-like domains. *Curr. Opin. Struct. Biol.* 9, 585-590.

15. Elgavish, S. & Shaanan, B. (1997). Lectin-carbohydrate interactions: different folds, common recognition principles. *Trends Biochem. Sci.* 22, 462-467.
16. Hakomori, S. (2002). Glycosylation defining cancer malignancy: New wine in an old bottle. *Proc. Natl. Acad. Sci. USA.* 99, 10231-10233.
17. <http://spdbv.vital-it.ch/>
18. <http://web.expasy.org/sim/>
19. Inoue H., Nojima H., and Okayama H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96: 23-28.
20. Kilpatrick, D. C. (2000). *Handbook of animal lectins: Properties and biomedical applications*. John Wiley & Sons, Chichester. UK.
21. Kumar, K. K., Chandra K. L. P, Sumanthi, J., Reddy, G. S. (2012). Biological role of lectins: A review. *J. Orofac. Sci.* 4(1), 20-25.
22. Landon, L. A., Peletskaya, E. N., Glinsky, V. V., Karasseva., N., Quinn, T., P. & Deutscher., S. L. (2003). Combinatorial evolution of high-affinity peptides that bind to the Thomsen-Friedenreich carcinoma antigen. *J. Prot. Chem.* 22, 193-204.
23. Leatham, A. J. & Brooks, S. A. (1998). Light microscopy: Overview and basic methods. In *Lectin Methods and Protocols*, pp 3-21, (Rhodes, J. M., & Milton J. D., Eds.), Humana Press, Totowa, New Jersey.
24. Liang, P.-H., Wang, S.-K., and Wong, C.-H. (2007) Quantitative analysis of

carbohydrate-protein interactions using glycanmicroarrays: determination of surface and solution dissociation constants. *J. Am. Chem. Soc.* 128, 13668–13669

25. Loris, R., Hamelryck, T., Bouckaert, J., & Wyns, L. (1998). Legume lectin structure.
26. Manoj, N. & Saguna, K. (2001). Signature of quaternary structure in sequences of legume lectins. *Prot Eng.* 14(10), 735-745.
27. Morgan, W.T. and Watkins, W.M. (2000) Unraveling the biochemical basis of blood group ABO and Lewis antigenic specificity. *Glycoconj. J.*, 17, 501–530.
28. Neurohr, K. J., Bundle, D. R., Young, N. M. & Mantsch, H. H. (1982). Binding of disaccharides by peanut agglutinin as studied by ultraviolet difference spectroscopy. *Eur. J. Biochem.* 123, 305-310.
29. Peletskaya, E. N., Glinsky, V. V., Glinksy, G. V., Deutscher, S. L. & Quinn, T. P. (1997). Characterisation of peptides that bind the tumor-associated Thomsen-Friedenreich antigen selected from bacteriophage display libraries. *J. Mol. Biol.* 270, 374-384.
30. Peumans, W. J., & Van Damme, E. J. M. (1995a). Lectins as plant defense proteins. *Plant Physiol.* 109, 347-352.
31. Peumans, W. J., & Van Damme, E. J. M. (1995b). The role of lectins in plant defense. *Histochem. J.* 27, 253-271.

32. Peumans, W. J. & Van Damme, E. J. M. (1998). Plant lectins: Specific tools for the identification, isolation, and characterisation of O-linked glycans. *Crit. Rev. Biochem. Mol. Biol.* 33, 209-258.
33. Reisner, Y., Linker-Israeli, M. & Sharon, N. (1979). Separation of mouse thymocytes into two subpopulations by the use of peanut agglutinin. *Cell. Immunol.* 25, 129-134.
34. Rinderle, S. J., Goldstein, I. J., Matta, K. L. & Ratcliffe, R. M. (1989). Isolation and characterisation of amaranthin, a lectin present in the seeds of *amaranthus-caudatus*, that recognises the T-antigen (or cryptic-T) antigen *J. Biol. Chem.* 264. 16123-16131.
35. Slifkin M, Doyle RJ. (1990). Lectins and their application to clinical microbiology. *Rev.* 3, 197- 218.
36. Speight, R. (2000). The development of a new plasmid display technology for functional genomics. PhD Thesis, University of Cambridge, UK.
37. Speight, R. E., Hart, D. J., Sutherland, J. D. & Blackburn, J. M. (2001). A new plasmid display technology for the *in vitro* selection of functional phenotype-genotype linked proteins. *Chem. Biol.* 8, 951-965.
38. Springer, G. F. (1984). T and Tn, general carcinoma autoantigens. *Science.* 224, 1198-1206.

39. Springer, G. F. (1997). Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. *J. Mol. Med.* 75, 594-602.
40. Sotozono, M. A., Okada, Y. & Tsuji, T. (1994). The Thomsen-Friedenreich antigen-related carbohydrate antigens in human gastric intestinal metaplasia and cancer. *J. Histochem. Cytochem.* 42. 1575-1584.
41. Salunke, D. M., Swamy, M. J., Khan, M. I., Mande, S. C., Surolia, A., Vijayan, M. (1985). Arrangement of subunits in peanut lectin. *J. Biol. Chem.* 260, 13576-13579.
42. Sharpe, C. E. (2003) PhD Thesis. Applied molecular evolution of peanut agglutinin: Increasing binding specificity towards tumour cell markers. University of Cambridge, Cambridge
43. Sharon, N & Lis, H. (1972). Lectins: Cell-agglutinating and sugar-specific proteins. *Science*, 177, 949-959.
44. Sharon, N & Lis, H. (1990). Legume lectins--a large family of homologous proteins. *FASEB J.* 4(14), 3198-208.
45. Sharon, N & Li, H. (2004). History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiol.* 14(11), 53-62.
46. Swamy, M. J., Gupta, D., Mahanta, S. K. & Surolia, A. (1991). Further characterisation of the saccharide specificity of peanut (*Arachis hypogaea*) agglutinin. *Carbohydr. Res.* 213, 59-67.

47. Van Damme, E. J. M., Peumans, W. J., Pusztai, A., Bardocz, S. (1998). Plant lectins: A special class of plant proteins: In. Handbook of Plant Lectins: Properties and Biomedical Applications (pp.3-6). Chichester, England: Wiley and Sons Ltd
48. Vyas, N. K., Vyas, M. N., & Quioco, F. A. (1991). Comparison of the periplasmic receptors for L-arabinose, D-Glucose, D-Galactose, and D-Ribose-structural and functional similarity. *J. Biol. Chem.* 266, 5226-5237.
49. Vijayan M. and Chandra N. 1999. *Curr. Opin. Struct. Biol.* 9: 707–714.
50. Varki, A. (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology.* 3, 97-130.
51. Young and Oomen, 1992 Young, N. M. & Oomen, R. P. (1992). Analysis of sequence variation among legume lectins: A ring of hypervariable residues forms the perimeter of the carbohydrate-binding site. *J. Mol. Biol.* 228, 924-934.
52. Young, N. M, Johnstone, R. A. Z., Watson, D. C. (1991). The amino acid sequence of peanut agglutinin. *Eur. J. Biochem.* 196, 631-637.

## Appendix

### General methods

#### E.coli strains

*E.coli* JM 109 strain (New England Biolabs) genotype: *F' tra D36 pro A B LactoseIq (LactoseZ) MIS/ (Lactose-pro AB) glnV44 e14- gyr A96 rec A1 rel A1 end A1 thi hs dR17*

#### Primers

*6xHisfor*

[5'-CAT ATG CAT CAC CAT CAC CAT CAC GCC GAA ACA GTT TCC TTC -3']

*pceseg*

[5'-GCC GCC AGG CAA ATT CTG-3']

#### Sugars all at ~ 30kDa

BP Gal- $\beta$ -1,3-GalNAc - $\alpha$ -PAA-biot

BP Lac PAA biot

BP LacNAc-PAA biot

(PAA refers to the polysaccharide biotinylated derivatives of the sugar)

### General Methods

#### Preparation of LB Agar Plates

All reactions were performed under standard conditions unless otherwise stated and all solutions were prepared using Milli-Q water. All work was carried out using *E.coli* JM 109 strain.

LB agar was heated until molten and allowed to cool to ~50 °C prior to the addition of the appropriate antibiotic. Agar was poured into petri plates (sterile plates) and stored at

4 °C when set.

### **Growth of Bacteria on Solid Media**

*E.coli* JM 109 containing the appropriate plasmid was transferred using a sterile wire loop onto LB agar with Chloroamphenicol (Cam) at 30 ug/ml, the plates were inverted and incubated overnight at 37 °C or until single colonies were evident.

### **Growth of overnight cultures**

A single bacterial colony isolated on LB Cam agar (30 ug/ml) was picked up and inoculated into 5ml LB cultures with LB agar Cam for plasmid preparation and incubated at 37 °C overnight with vigorous agitation 225 rpm and then used as an inoculum into 500ml LB agar Cam for incubation overnight at 30 °C.

### **Preparation of Competent cells**

*E.coli* JM109 strain was used as a competent cell host for transformations using the Rubidium chloride method of preparation. A single colony of JM 109 was inoculated into a 5ml LB and grown overnight with shaking at 37 °C starter culture was then used to inoculate a 500ml LB culture and grown at 37 °C to OD<sub>595</sub> ~ 0.5 for ~3 hours. The culture was allowed to cool on ice for 15 minutes and then transferred to two 250ml centrifuge bottles and spun for 10 minutes at 4500 rpm in a Beckman centrifuge in order to pellet the cells. The supernatant was poured off and the pellets resuspended in 30ml total of Tfb I, and then placed on ice for 15 minutes. The cells were spun once again for 5 minutes at 4000 rpm and the pelleted cells resuspended in 6ml Tfb II and placed on ice. 100µl volume of competent cells was aliquoted into 1.5ml eppendorf tubes and stored at -80 °C.

### **DNA Manipulation**

Plasmid DNA was prepared from 5ml overnight cultures using Qiagen Miniprep Kit (QIAGEN) according to the manufacturer's protocol.

Sequencing was carried out by the Central Analytical Facility, Stellenbosch University

### **Agarose gel electrophoresis**

<b>Agarose concentration</b>	<b>Fragment ranges bp DNA ladder</b>
<b>0.8% - 1%</b>	<b>196-3109 bp</b>
<b>1.5-2%</b>	<b>90- 3109 bp</b>
<b>3%</b>	<b>90- 968 bp</b>

Electrophoresis grade agarose was added to 1x TBE buffer dependant on the percentage gel required as per fragment ranges. The suspension was boiled in a microwave and cooled to ~50 °C after which 5ul of EtBr of 10mg/ul added and the molten agarose was poured into a gel tray and allowed to set with appropriate comb size. The gels were immersed in 1x TBE running buffer until ready to be loaded

DNA ladder at a 2x concentration was loaded into one well to serve as molecular weight markers for estimation of DNA fragment sizes. Loaded gels were run at ~100 V and stopped just before the dye front reached the bottom of the gel. DNA Fragments were visualised on a UV trans-illuminator and the molecular masses assessed by comparison to the molecular weight markers.



## Scatchard Plot.

### Ligand binding assay

Protein<sub>free</sub> + sugar<sub>free</sub> = protein.sugar<sub>bound</sub>

$$K_d = \frac{[\text{protein}][\text{sugar}]}{[\text{protein.sugar}]} \quad (1)$$

Don't know [sugar<sub>free</sub>] but do know that

$$[\text{sugar}_{\text{total}}] = [\text{sugar}_{\text{free}}] + [\text{protein.sugar}_{\text{bound}}]$$

Substituting for [sugar<sub>free</sub>] in (1) gives

$$K_d [\text{protein.sugar}] = [\text{protein}] ([\text{sugar}_{\text{total}}] - [\text{protein.sugar}])$$

$$K_d [P.S] + [p][p.s] = [p][s_{\text{tot}}]$$

$$\Rightarrow [p.s] = [p][s_{\text{tot}}] / (K_d + [p])$$

Now, call [s<sub>tot</sub>] "B<sub>max</sub>"; call [p.s] "B"; and call [p] "F"

$$B = F B_{\text{max}} / K_d + F$$

$$B K_d + BF = F B_{\text{max}}$$

$$B K_d = F B_{\text{max}} - BF = F (B_{\text{max}} - B)$$

$$\Rightarrow B / F = B_{\text{max}} - B / K_d = B_{\text{max}} / K_d - B / K_d.$$

= scatchard plot.

Therefore a plot of B/F  $\propto$  B

Will have a gradient of  $-1/K_d$









680            690            700            710            720            730            740            750  
 GC T T T T T C C C T G T T G G G T T G G G T T G T T T G C C T T T T G G T T T G T T G T T G C C T T C C C C T T G G T T T T T T G G G G G T T T C

