A study of the structure and biological activity of the Eranthis hyemalis Type II Ribosome Inactivating Protein.

by

Marie-Therese McConnell

Canterbury Christ Church University

Thesis submitted

for the Degree of Doctor of Philosophy

2016

Table of Contents

List of figures

List of Tables

List of Abbreviations

Å	Angstrom
Arg	Arganine
Asn	Asparagine
ASP	Ammonium sulphate precipitation
AuNP	Gold Nanoparticle
BLAST	Basic Local Alignment Tool
BSA	Bovine Serum Albumin
CCP4	Collaborative Computational Project No. 4
Daf-c	Dauer abnormal formation – constitutive
DAP	1,3 Diaminopropanol
DLS	Dynamic Light Scattering
DNA	Deoxyribonucleic acid
DT	Direct treatment
DTE	Dithioerythritol
DMEM	Dulbecco's Modified Eagle Medium
Dyf-d	Dye filling - defective
EF1/EF2	Elongation Factor 1/2
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylenebis(oxyethylenenitrile)tetraacetic acid
EHL	Eranthis hyemalis lectin
FASTA	Fast-All

FITC	Flourescein Isothiocyanate
Fuc	Fucose
GalNAc	N-acetyl galactosamine
Gal	Galactose
Glc	Glucose
Glu	Glutamic acid
GlcNAc	N-acetylglucosamine
IdoA	Iduronic Acid
kDa	Kilo Dalton
L1/2/3/4	Larval stage 1/2/3/4
LUCA	Last Universal Common Ancestor
ML1	Mistletoe Lectin 1
Man	Mannose
MALDI-TOF	Matrix Assisted Laser desorption Ionsisation – Time of Flight
MS	Mass Spectrometry
MR	Molecular Replacement
MWCO	Molecular Weight Cut Off
MQ	MilliQ water
NAG	N-Acetly-D-Glucosamine Ligand code
NCBI	National Center for Biotechnology Information
Neu5Ac	Sialic Acid
NGM	Nematode Growth Media

$(NH_4)_2SO_4$	Ammonium Sulphate
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PES	Polyethersulfone
PHA	Phaseulus vulgaris agglutinin
RIP	Ribosome Inactivating Proteins
RIPA	Radioimmunoprecipitation assay buffer
RNA	Ribonucleic acid
RMSD	Root Mean Square Deviation
SDS-PAGE	Sodium dodecyl sulphate poly acrylamide gel electrophoresis
SPR	Surface Plasmon Resonance
T antigen	Thomson-Friedenreich antigen
Tn antigen	N-acetyl-galactosomine linkage to common core of T antigen
TEM	Transmission Electron Microscopy
Trp	Tryptophan
Tyr	Tyrosine
UV	Ultra Violet
Val	Valine
Xyl	Xylose

Acknowledgments

I would like to thank my supervisor Emilia Bertolo-Pardo for all of the support she has offered during my project, especially for always helping to find new ways of thinking about a problem, her endless patience and sense of humour, her belief in me which allowed me the freedom to explore areas of science that were beyond the boundaries of where I thought I could reach. I would like to thank my second supervisor Simon Harvey for his guidance, direction, patience and support in all things nematological especially for the lengthy experiments during which he supported me. I am most grateful for having had the support, guidance and belief from my supervisors to have taken the directions and opportunities which opened up before me. Thanks are given to David lisgarten for opening the door to the world of crystallography for me, through which I have stepped. I would also like to acknowledge the support of my panel chair Peter Vujakovic and Professor Adrian Holliday who made obstacles to progress always disappear, and of my fellow student, colleague and friend, Jana for walking this path along with me. To my dear friends Leanne and Philip, I am grateful for the encouragement, guidance and practical support given during my thesis writing. I am very humbled and eternally grateful to have been funded with a scholarship from Canterbury Christ Church University which allowed me to reach for my goals.

I would like to acknowledge the immense support given to me by members of the crystallographic community, specifically the CCP4 and BCA community, Oxford University Biochemistry department, The Wellcome Trust for Cell Biology at Edinburgh University, Birkbeck College, the Manchester Institute of Biotechnology and The Diamond Light Source at Harwell, and the organisers of The Northern Protein Structure Workshop where many an idea was learned in good company. The warmth, patience, knowledge exchange and practical support that was freely offered by members from other institutions and the strong hands that were extended to allow me to breach the gap will forever be held precious by me.

I would sincerely like to thank Professor Elspeth Garman and Dr Carol Trim for agreeing to examine my thesis.

Most of all I would like to thank my parents for their enduring love and support, and to my children for their unflinching patience, for never questioning the validity of my goal, for always accepting less than they deserved and for the sacrifices they accepted. For you my children, Freya and Killian, I hope you find your passion as I did mine, and thank you for allowing me to chase my dreams, as well as be your mother.

Freya and Killian, this thesis is dedicated to you both.

Abstract

Plant lectins and Ribosome inactivating proteins (RIPs) have been studied for well over a century.The first published work on the canonical RIP, Ricin, from *Ricinis communis* being characterised by Stillmark in his doctoral thesis in 1888. These proteins are characterised by the ability to selectively and reversibly bind to the carbohydrate groups of glycoconjugates, agglutinate erythrocytes and, in the case of RIPs, act as RNA glycosidases, depurinating a specific adenine residue within the highly conserved ricin/sarcin loop of the 28S ribosomal sub unit. Lectins and RIPs are used extensively in biotechnology and biomedical research. Winter Aconite Lectin (EHL) extracted from the corms of the *Eranthis hyemalis* plant was previously identified as having the characteristics of a Type II RIP, and therefore provided a novel target for functional characterisation studies. Homologous proteins have shown potential as immunotoxins and diagnostic tools in cancer research, mainly due to the changes in glycosylation patterns of tumour cells such as, expression of the Tn antigen rarely found in healthy tissue. As EHL has a glycomic binding profile that has highest affinity for N-acetyl-D-galactosamine, the defining carbohydrate unit of the Tn antigen, it therefore offers an interesting target for exploring the biological structure, function and effects on model organisms and cell lines.

The first aim of the research reported within this thesis was to establish an extraction and purification protocol, which would provide a homogeneous, pure and consistent yield of EHL in sufficient quantity to allow further study of the characteristics and structure. Once established, toxicity studies were carried out using *Caenorhabditis elegans* to examine the effect of EHL in a biological system. The data collected indicated that EHL binds to the

amphid neurons in *C.elegans* and causes a wide variety of concentration dependent effects including abnormal dauer formation.

In order to elucidate the structure of the protein, crystallographic screening was undertaken. A number of large single crystals produced diffraction data. A model structure was constructed using molecular replacement for phasing. The model is presented, as a work in progress with interesting observations as to potential glycosylation sites and active site residues.

A preliminary investigation into the viability of a conjugation of EHL to gold nanoparticles was undertaken in collaboration with The Bioscope group (University of Lisbon) and the conjugate was also tested for biological activity. Initial results indicated that the nanoparticles caused a cessation of the toxic effect of EHL.

In the final stages of the research presented herein, EHL was tested for antineoplastic activity against a number of cancer cell lines using a variety of methods to develop a viable protocol for future studies. Preliminary results indicate that EHL may act as a mitogenic agent on some cell lines, which also presents an opportunity for onwards research.

Chapter 1: Introduction

1.1 Winter Aconite

The *Eranthis hyemalis* plant or Winter Aconite is a late winter/early spring flowering perennial plant of the family *Ranunculaceae.* Though heavily cultivated by commercial growers, in a natural setting *E.hyemalis* can be found in deciduous woodland across central Europe. It is classified as an ephemeral plant due to the short life cycle which allows the plant to exploit higher light levels in winter when the overgrowing canopy has undergone autumnal abscission. The aerial parts of the plant emerge and a short period of growth culminates in a flowering period in late winter between January and March, when seeds are formed, foliage then dies back and the plant retreats below ground for the rest of the season (Royal Horticultural Society, 2016).

E.hyemalis is native to Europe and was introduced and thrived in North America and Canada, naturally occurring in woodland and commonly cultivated in gardens. A close relative *Eranthis cilica* extends the geographical range of the genus to Turkey and Afghanistan. *E. hyemalis* propagates easily by tuberous growth and seeding. Within the horticultural community, it is observed that this plant appears to be free from particular pests and pathogens (Missouri Botanical Gardens, 2016). Other members of this family include buttercups, marsh marigold and larkspur. As with many members of the buttercup family (*Ranunculaceae*), it is widely recognised as being a toxic plant if ingested and warnings of skin irritation are generally provided with any commercially available stock. Research first carried out in 1985 by Cammue, Peeters and Peuman (1985) indicated however that *E.hyemalis* also expresses a proteinaceous toxin of potentially more notable impact than causing skin irritation to various

substrates when isolated and purified. This toxin was found to cause clumping (agglutination) of erythrocytes as well as impacting on the fitness of some formidable herbivorous agricultural pests and plant viruses in later work by Kumar *et al.,* (1993). This toxin was named *Eranthis hyemalis* lectin or EHL. To date *E.hyemalis* is the sole representative of the *Ranunculaceae* to be reported to express lectin activity, and it is towards this activity that this research is focused.

1.2 Lectins

Lectins are a family of proteins found in plants, animals, bacteria and viruses which have the ability to selectively, specifically and reversibly bind to carbohydrate groups and can also be referred to as carbohydrate binding proteins (Sharon and Lis, 2007). This study will focus discussion on lectins found only within the plant kingdom.

Research into the biological properties of lectins began in the 19th century, mainly focusing on the observation that lectins agglutinate erythrocytes. This is widely believed to have been first reported in the doctoral thesis written by Peter Stillmark in 1888 (Stillmark, 1888; Sharon and Lis, 2004) on the highly toxic ricin, the hemagglutinin from the Castor bean plant (*Ricinis communis*) shortly followed by Abrin from *Abrus precatorius*. These early candidates were plant derived and were termed phytohemagglutinins. Karl Landsteiner built on the early work of Stillmark and made his subsequent own discovery that agglutination of blood cells from one individual by the sera of other groups of individuals could also be induced by the use of plant lectins, (Landsteiner, 1907) for which he won a Nobel Prize in 1930 (Schwarz and Dorner, 2003). In the following years, blood types had been well defined (for the ABO system) and many lectins were identified as having specificity for a particular blood group (Renkonen, 1948), however a new discovery published by Morgan and Watkins in 1953 identified

individual blood group determinants as being carbohydrate structures. This work presented the finding that lectins differentially agglutinated various animal blood cells, and that by competitive (hapten) inhibition with mono or oligosaccharides, agglutination could be inhibited, thereby demonstrating the specific interaction between lectins and the varying surface carbohydrates of specific types of erythrocytes (Morgan and Watkins, 1953). Understanding the differences in blood types had far reaching implications in medical science including increasing the successful outcome of blood transfusions and eventually organ transplants. The schematic (Figure 1.1) below shows a basic representation of the differing erythrocyte surface epitopes made up from the core fucose 'H' (O) antigenic determinant. The epitope is the region of a structure from where an antigenic response is elicited, so in the case of the A type the Anti-A antibody will respond to the GalNAc terminal residue. Lectins cause the agglutination of cells as a result of binding to the terminal glycans and as most lectins have more than one carbohydrate binding region (Berg *et al.*, 2002), this results in cross linkage.

Lectins which agglutinate human erythrocytes for groups A and O are readily found in nature, with blood group B specific lectins being rarer. This characteristic common to all lectins, is then the ability to bind to the carbohydrate moieties of cell surface glycoconjugates specifically and reversibly without introducing conformational changes to the mono- or oligosaccharides they to which they ligate (Sharon and Lis, 2004).

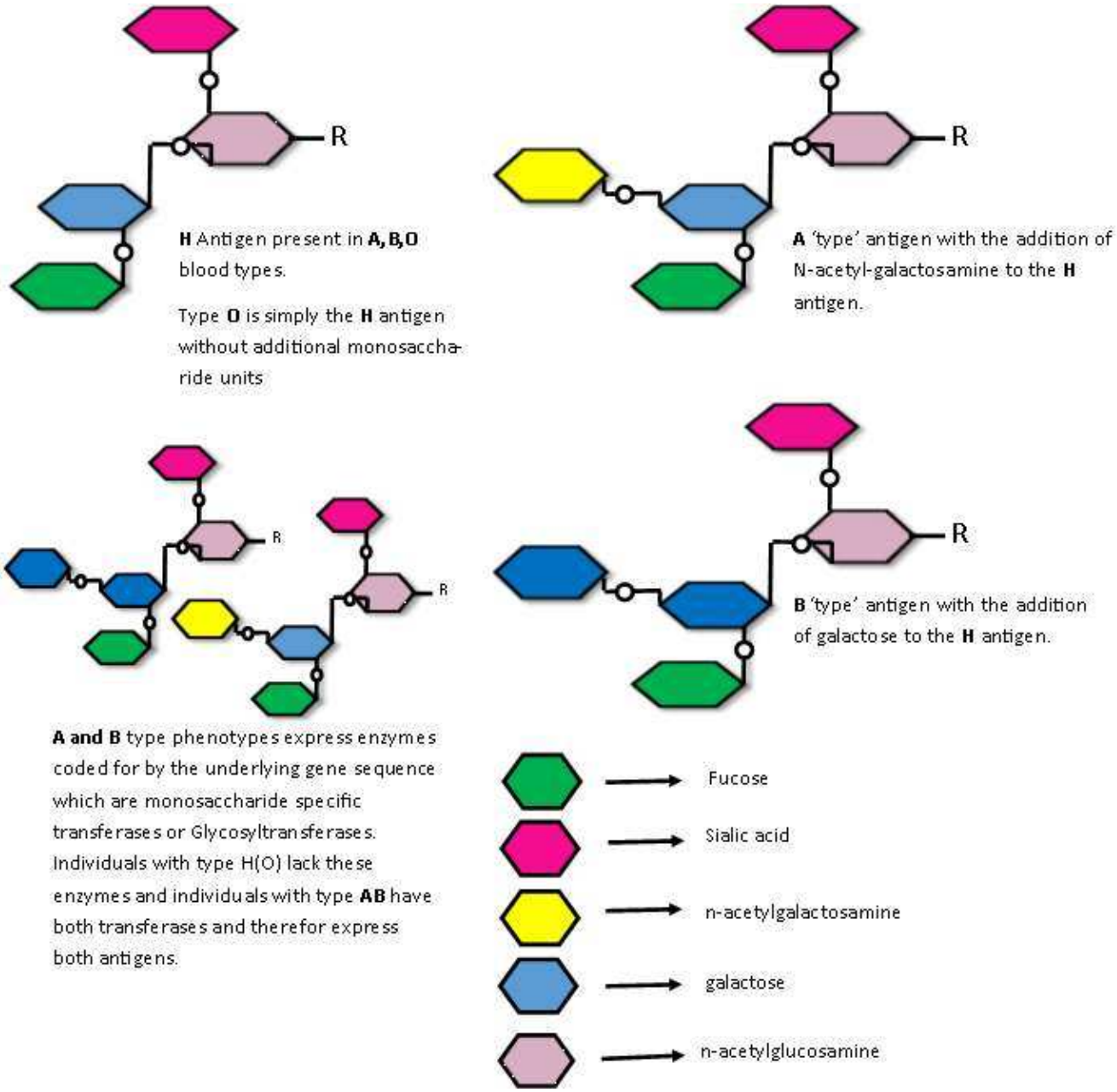


Figure 1.1 Schematic of the ABO surface antigen determinants.

Lectins are involved in a wide range of processes including carbohydrate transport, cell signaling and defence, and in potentially mediating symbiotic relationships between leguminous plants and their associated rhizobium (Sharon and Lis, 2004). Plant lectins have been proposed as playing a key role as plant defence mechanisms (Peumans and Van Damme, 1995), with many known to specifically bind to epithelial cells in herbivore and nematode guts

(Schubert *et al.*, 2012; Delatorre *et al.*, 2007). Insecticidal, antifungal and antiviral qualities have also been widely described (Edwards and Gatehouse, 2007; Peumans, Hao and Van Damme, 2001; Rao *et al.*, 1998). The empirical observations made by horticulturists that *E.hyemalis* has no reported pests perhaps hints towards the presence within its tubers of a lectin with toxic properties.

Some lectins have shown the capability of acting as a mitogenic stimuli to lymphocytes and other cells (Morgan, Ruscetti and Gallo, 1976). This results in a quiescent lymphocyte undergoing cell division and proliferation as a result of treatment by a mitogenic agent. Many lectins possess this characteristic but may only demonstrate it under the optimal experimental conditions with the right cell type (Kilpatrick, 1998). This property was discovered serendipitously during an experiment using a lectin from *Phaseulus vulgaris* (PHA) to separate leucocytes from red blood cells by agglutination and subsequent division of the whole blood products. The experiment indicated that the PHA was initiating mitosis among the leucocytes, a characteristic which was then exploited to maintain the lymphocyte culture (Nowell, 1960). With the finding that other lectins are also mitogenic, their use in biological applications was firmly rooted, and major advances such as the discovery of T Cell Growth Factor (interleukin-2) (Mier and Gallo, 1982) have resulted from their use both as serological and mitogenic tools. Commonly used lectins for mitogenesis are Concanavalin A, PHA and Pokeweed mitogen (Kilpatrick, 1998).

This thesis focuses on the lectin found in the tubers of Winter Aconite, *Eranthis hyemalis*, (EHL) which should be more accurately described as a ribosome inactivating protein (RIP). Earlier work has alluded to this fact, and with the structural work completed here, it is unreservedly the case. To-date, EHL is the sole representative of the *Ranunculaceae* to be

identified as a RIP. EHL preferentially binds N-acetyl-galactosamine, but also binds Galactose, Galacto-pyransosyl, D-glucose, and to a lesser degree D-ribose (Cammue, Peeters and Peuman, 1985). EHL is indeed a lectin but this description tells only half the story, and EHL should indeed be fully classified as a Type II Ribosome Inactivating Protein (RIP) based on the previously published structural and toxicity studies conducted (Cammue, Peeters and Peuman, 1985; Kumar *et al.*, 1993; George *et al.*, 2011) and the work herein described.

1.3 Ribosome Inactivating Proteins.

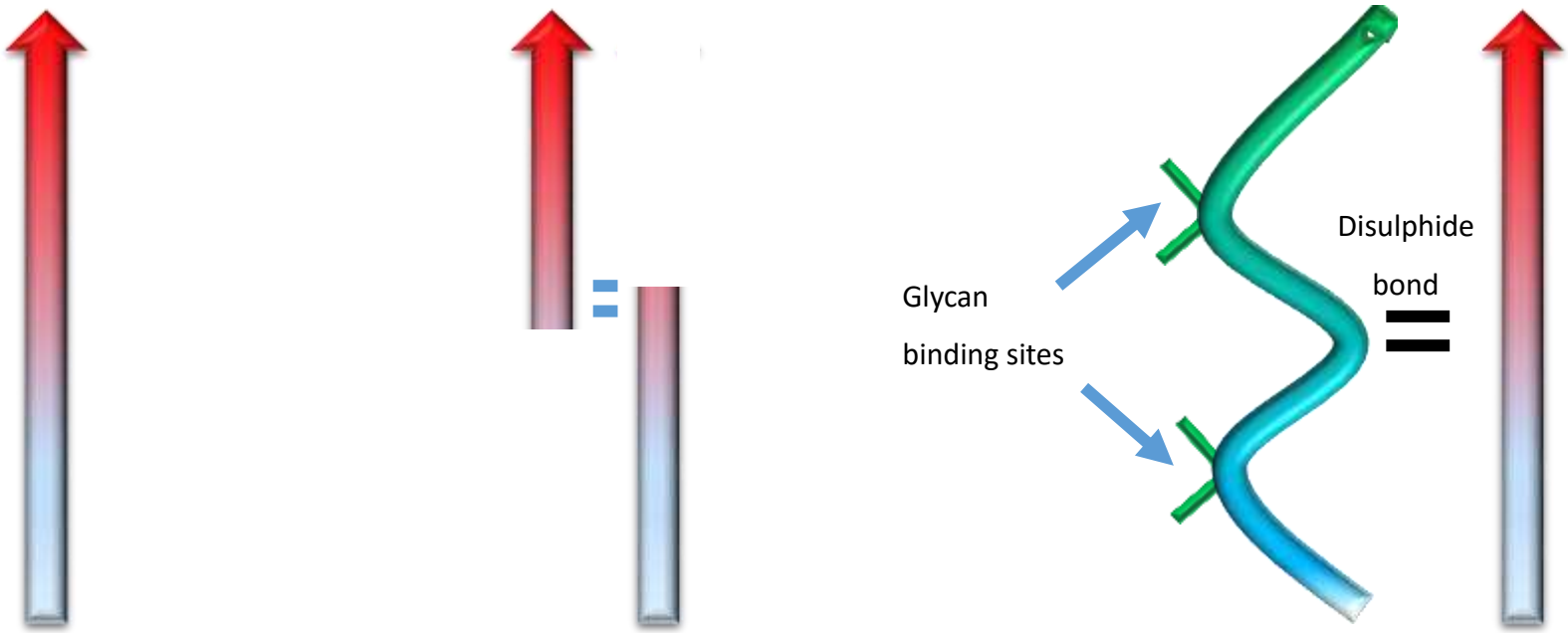
Ribosomes are the complex cellular organelles responsible for protein synthesis, by translating messenger RNA (the product of gene transcription) into polypeptide chains (Ben-Shem *et al.*, 2011). The eukaryotic ribosome is a complex of three or four ribonucleic acid molecules (dependent on taxonomy) and over 80 different proteins. The structure actually consists of a large and small subunit (Lodish, 2008) designated 60S and 40S respectively. These combine upon initiation of translation to form the 80S assembled ribosomal complex which binds to and deciphers the mRNA template. The S designation is a measure of the relative centrifugal sedimentation rate of each component measured in Svedberg units. This also applies to the rRNA molecules designated as 28S; 5S (also 5.8S in vertebrates) in the large 60S subunit and 18S in the small subunit. After the complex process of initiation has begun and the correct initiator methionine-charged tRNA (transfer RNA) has bound to the small subunit, addition of amino acids can proceed with the involvement of the elongation factors EF1 and EF2. When a stop codon is reached on the mRNA, release factors terminate translation and the mature peptide is released from the ribosome for processing (Lodish, 2008).

Ribosome inactivating proteins are a class of enzymes (EC 3.2.2.22) with a mode of action which results in the breakage of a glycosidic bond in the 28S rRNA in the 60S subunit of the ribosome, resulting in halt of protein synthesis and subsequent cell death. This characteristic mechanism, common to all RIPs, is the ability to cleave a single adenosine from a specific site in the ribosomal rRNA, A_{4234}. The Adenosine residue is situated in a highly conserved region known as the ricin-sarcin loop. The result of this glycosidic cleavage prevents EF1 from binding correctly. EF1 binding is essential for translation to proceed. Without EF1 binding appropriately, protein synthesis is terminated (Lodish, 2008). This N-glycosidase or glycosylase activity on rRNA is also reported to act on other deoxyribose nucleic acid substrates *in vitro* and therefore RIPs can be considered as polynucleotide deadenylating N-glycosidases (Barbieri *et al.*, 1997). RIPs are currently classified as Type I, Type II or Type III (see figure 1.2) based on the overall polypeptide chain composition and crucially, the presence or absence of lectin activity. Type I RIPs consist of a single RNA N-glycosidase domain which for ease of description will be termed an A chain. Most Type I RIPs consist of a single polypeptide of circa 30 kDa but may also assemble into oligomeric conformations. In addition to these Type I RIPs, some controversy exists as to the nomenclature of an out-group of RIPs which appear to be formed from proteolysis of a zymogen form of the 30 kDa polypeptide. This has resulted in two shorter peptide chains bound by noncovalent interactions referred to as either two chain Type I RIPs or more rationally Type III RIPs (Van Damme *et al.*, 2001). Type III RIPs continue to be classified due to their recent emergence in the research field, and will not be included in this discussion.

Type II RIPs consist of either one or two heterodimers (tetrameric form) linked by disulphide bonds; an A and a B chain. By convention, the B subunit or chain is a sugar specific lectin

containing the highly conserved ricin B domain which facilitates binding to extracellular glyconjugates. The mode of action of type II RIPs shows that this binding then mediates entry to the intracellular environment for the attached cytotoxic A-chain with the N-glycosidase activity by endocytosis (Virgilio *et al.*, 2010). The toxin is then subjected to retrograde transport via the Golgi complex, to the endoplasmic reticulum, where the disulphide bonds are reduced and the A chain is released in the cytosol to refold into an enzymatically active form (Hartley and Lord, 2004).

As previously mentioned the A chain of a Type II RIP acts as an inhibitor of eukaryotic protein synthesis by enzymatically breaking the glycosidic bond of a single Adenosine molecule at position 4234 ($A_{4234)}$ of the 60S rRNA subunit (Hartley and Lord, 2004). The B chain of a Type II RIP is a sugar specific lectin and contains multiple carbohydrate binding domains known as ricin-b domains. It is this binding mechanism which makes Type II RIPs more highly toxic than Type 1, despite the conserved A chain tertiary structures and sequence homology. Entry into eukaryotic cells is mediated by the binding of the lectin to cell surface glycoconjugates and receptors expressing particular glycomic profiles mirrored by the lectin's carbohydrate affinities. Type I RIPs exhibit extensive amino acid sequence conservation in the active site and overall similarity to the A chain of Type II RIPs.



Type I RIP: Single polypeptide chain with cytotoxic N-glycosidase activity. No lectin activity.

Type III or 2 chain Type I RIP: Single zymogen polypeptide chain or proteolytic product resulting in two truncated chains, cytotoxic.
No lectin activity.

Type II RIP: Cytotoxic A and carbohydrate binding lectin B chain with binding sites.

Figure 1.2 Schematic diagram of the over simplified structure of RIPs

In recent years the potential of lectins for use in cancer therapies has become a significant research focus due to their ability to preferentially bind to specific carbohydrates and differentiate between glycosylation patterns. Moreover, a number of plant derived lectins have been shown to have potent *in vitro* and *in vivo* anti-cancer effects (Voss *et al.*, 2006; Otsuka *et al.*, 2014) inducing autophagous and apoptotic pathways in tumour cells, and some are already used therapeutically. For instance, the recombinant mistletoe lectin rViscumin has been through phase 1 clinical trials and a number of other native mistletoe lectin preparations such as Lektinol and Iscador are prescribed widely throughout Europe as adjuvant therapies in chemotherapy, although their efficacy is not readily quantified (Horneber *et al.*, 2008).

Type II RIPs are specifically an area of increasing interest due to their antineoplastic properties, and their glycomic binding profile can be used to target specific glycans of biological molecules. For instance, the GalNAc specific RIP *Ximenia americana* (Riproximin) (Voss *et al.*, 2006)), *Sambucas sp* (Ferreras *et al.*, 2011)) and ML1 from *Viscum album* (Tonevitsky *et al.*, 1996) show higher binding affinity for tumour cells than for healthy cells, due to the over expression of particular surface saccharide groups in the changing glycomics of malignant cells (Bayer *et al.*, 2012). It is to this end that this thesis seeks to establish if EHL can demonstrate similar preferential binding to malignant cells.

The evolution of RIPs has been extensively studied and currently two main hypothesis are considered to be plausible. The classic hypothesis is that RIP genes were acquired by an ancestor of the seed plants, from which the Type I RIPs evolved. The acquisition of a lectin domain was then the subject of horizontal gene transfer from bacterial species and resulted in the Type II lineage. Subsequently a further deletion of the B-chain of these

divergent Type IIs giving rise to the Type III RIPs more recently (Peumans and Van Damme, 2010). However, according to Lapadula, Puerta and Ayub (2013) an alternative more parsimonious hypothesis is proposed. The fact that RIP domains have been found in both gram positive, gram negative bacteria and fungi presents the notion that paralogs of the RIP gene were present in the last universal common ancestor (LUCA.), and, were lost in some species and duplicated in others. This created the wide diversity of RIPs found, and expression of more than one type in some species such as those found in the *Iridaceae*, *Euphorbiaceae*, and *Cucurbitaceae* (Van Damme *et al.*, 2001). Lapadula, Peura and Ayab (2013) do however support the theory that the lectin domain was acquired and fused to the RIP domain in some species. This serves to illustrate the complexity of making a definitive phylogeny of the ribosome inactivating proteins. Furthermore, with the addition of newly identified RIPs to the databases and the discovery of the structure of novel divergent RIPs such as EHL providing more sequences on which to build a phylogeny, the discussion will continue into the future. The taxonomic lineage of *Eranthis hyemalis* can be seen in Table 1 and a characteristic example of a flowering *E. hyemalis* in figure 1.3.

Despite more than 40% and in some cases 50% homology to other RIPs, EHL would appear to belong to an outgroup of N-glycosidases as can be seen in figure 1.4. Using the NCBI BLASTp search for homologs shows that EHL is extremely divergent from Type I RIPs, and this is also reflected in the higher sequence divergence of the A-chain of homologous Type II RIPs. In contrast, the B-chain of EHL shows a much higher sequence similarity to other Type II RIPs. This is discussed in more detail in Chapter Three.

Table 1	
Kingdom	Plantae
Sub kingdom	Viridiplantae
Infrakingdom	Streptophyta
Superdivision	Embryophyta
Division	Tracheophya
Class	Magnoliopsida
Superorder	Ranunculanae
Order	Ranunculales
Family	Ranunculaceae
Genus	Eranthis
Species	hyemalis

Table 1. Taxonomy of *Eranthis hyemalis* (TSN 18775) from the Integrated Taxonomic Information System on-line database.

Figure 1.3. Cultivated Winter Aconite (*E. Hyemalis*) in flower.

1.4 Glycosylation

One of the main forms of peptide post translational modification is glycosylation. The study of glycomics, and in particular altered glycomics as disease markers is key to the progression of lectin research. It has been previously reported that in tumorogenesis alterations to the normal glycosylation pattern of the diseased cell, such as incomplete or truncated O-linked glycosylation, are one of the primary events which occur as an outcome of underlying genetic changes in glycosyltransferases leading to pathogenicity (Ghazarian, Idoni and Oppenheimer, 2011; Gorelik, Galili and Raz, 2001).

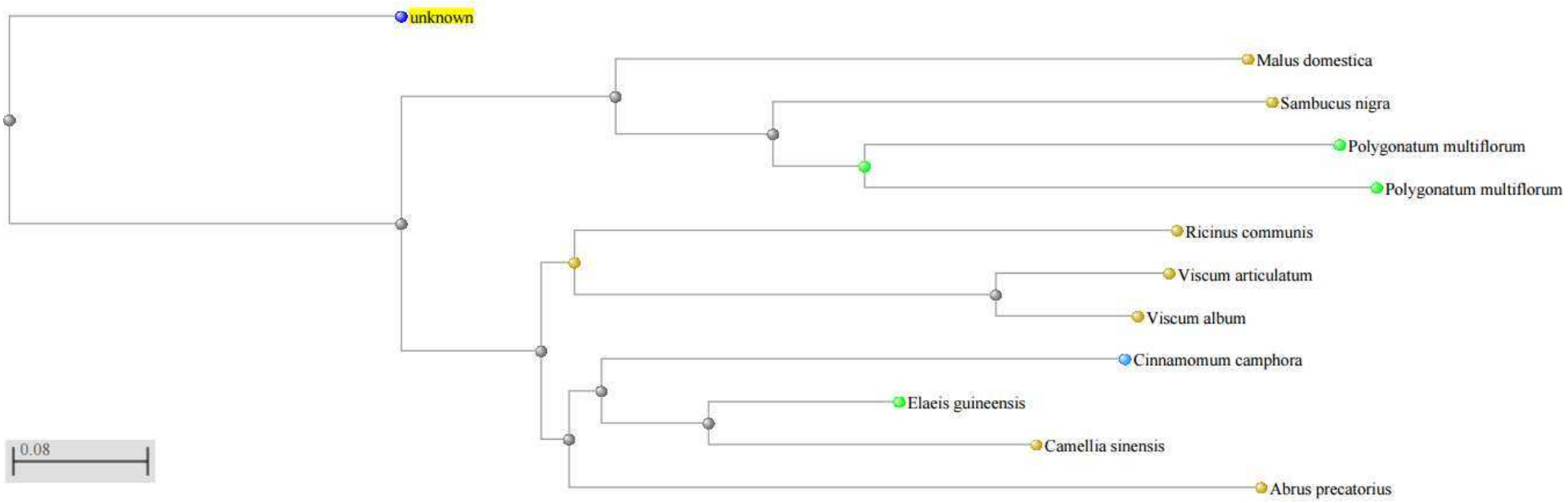


Figure 1.4 Neighbour Joining phylogenetic tree of EHL and other Type II Ribosome inactivating proteins. EHL is the 'unknown' node.

Of particular interest for future work with EHL is the cancer associated antigen, the Tn Antigen (Madariago *et al.,* 2014). N-linked and 0-linked glycosylation are the major pathways of interest in the biosynthesis pathways of glycoconjugates. N-linked glycans are N-acetylglucosamine units covalently bonded to an asparagine residue of a poly-peptide. O-linked glycans consist of a primary GalNAc unit covalently bonded to either a serine or threonine residue. These form the core of many glycoconjugation events, with further saccharide units being added to form a complex and vast array of possible branched or unbranched oligosaccharides (Gorelik, Galili and Raz, 2001). These are formed in mammals from a set of ten core monosaccharides, all of which can be derived from cellularly produced glucose. These are Glucose (Glc); galactose (Gal); N-acetylglucosamine (GlcNAc); N-acetylgalactosamine (GalNAc); Fusose (Fuc); Mannose (Man); Xylose (Xyl); glucuronic acid (GlcA), iduronic acid (IdoA), and 5-N-acetylneuraminic acid (Neu5Ac), more commonly known as sialic acid (Stowell, Ju and Cummings, 2015).

Amino acids undergo glycosylation as a form of post translational modification which attaches specific glycans to the peptide as can be seen in the N-acetylglucosamine unit attached to residues Asn 11 and Asn 355 in the current model of EHL. Biosynthesis of glycoproteins begins in the secretory pathway which compromise the endoplasmic reticulum and Golgi apparatus, and, ultimately signals the destination of the proteins in membrane localization or secretion to the cytosol. Whilst glycosylation of this nature is part of the normal cellular function, when underlying genes are mutated this can result in incomplete or mutated glycan structures being constructed or completely absent. As already mentioned, an example of this of particular interest is the Tn antigen which is recognised as an oncogenic marker (Ju, Otto and Cummings, 2011). Normally cells express the T antigen which is an O-linked structure with a

terminal GlcNAc unit and a GalNAc core. In antineoplastic disease, many enzymes can be affected. Specifically, the T synthase or N-acetylglucosamine transferase can be absent, caused by mutations in a gene such as *cosmc* (Hofmann *et al.*, 2015). The outcome of this is that the repetitive GalNAc core is exposed and produces the GalNAc clustered Tn antigen. This has now been classified as a biomarker of metastatic potential in cancers, with a correlation found between Tn density and the aggressiveness of a particular cancer. Lectins have been used to detect Tn antigen successfully *in vitro* and the hypothesis of increased binding to cancer cells by GalNAc specific lectins and RIPs is an area of current research (Madariago *et al.*, 2014).

1.5 Uses of lectins and ribosome inactivating proteins.

Due to their ability to bind preferentially to specific carbohydrates, plant lectins are now used in many aspects of immunology, biotechnology and biomedical research in both diagnostic and therapeutic roles. Over recent years lectin research has evolved as a means of gaining further understanding of many aspects of cancer and metastatic development due to the glycobiology of the cellular transformations. Lectins are useful in tumor cell recognition due to their binding affinity to antineoplastic biomarkers (Heinrich *et al.*, 2005). Furthermore, lectins are also used for their mitogenic and agglutinating properties in medical applications such as blood typing and cell culture (Mody, Joshi and Chaney, 1995). As previously mentioned, Mistletoe lectin extract is a widely used adjuvant therapy in Europe for patients undergoing chemotherapy, and a study conducted in 2002 with a group of breast cancer patients resulted in reduced primary treatment side effects and an increase in relapse-free time (Schumacher *et al.,* 2002).

Lectins and RIPs including Ricin are commercially produced for use in applications such as affinity chromatography, whereby the sugar specific lectin of choice is bound to a matrix, such as agarose or sepharose, to which the target glycoprotein will then bind in a column and be separated. Full lectin array kits are available in order to probe the expression of glycans on a substrate, and commercial mitogens are also in use. Fluorescent labelling of lectins has multiple applications in histochemistry and the relatively new concept of lectin immunoblotting can provide an economical alternative to traditional western blotting, and allow for the probing of specific sugar motifs on receptors (Dan, Liu and Ng, 2016). The symbiotic interaction between lectins and carbohydrates remains an area of research where potential applications are constantly emerging. Therefore, the contribution of this thesis in further characterizing the structure and activity of a lectin and ribosome inactivating protein with evolutionary divergence offers new possibilities.

1.6 Biological assays with *Caenorhabditis elegans* (*C.elegans*).

C. elegans is a free living nematode species used in laboratories worldwide for a variety of genetic and toxicology studies due to its well documented physiological and genetic characteristics which make it an ideal model organism. Due to the conserved nature of disease pathways between *C. elegans* and other organisms, including humans, it is an ideal model to carry out *in vivo* research and has led to many important discoveries. The culturing of *C. elegans* populations in the laboratory is an economical and accessible system and can yield extremely powerful statistical outputs due to the well-defined and annotated genotypes and phenotypes. The worms develop from an egg laid by a self-fertilising hermaphrodite adult and will reach reproductive maturity within only three days. An adult reaches around

1.3mm long and is capable of producing around 300 offspring, and is therefore ideal for studies involving reproductive fitness (Kaletta and Hengartner, 2006).

With an average lifespan of around 2 weeks, studies into aging and lifespan effects can also be conducted, with experimental results again producing powerful statistical outputs due to the large numbers of replicates involved. An example which illustrates this well is that of the recent findings of Stastna *et al.* (2015), who tested the hypothesis that calorific restriction is widely believed to increase lifespan. However, using a complex genetic panel of nematodes, they were able to establish that this effect is actually genotype dependent, and in some unfavourable genetic interactions (which they were able to map) it has the opposite effect (Stastna *et al.*, 2015). As such, *C. elegans* forms the system of choice for the toxicity assays carried out in this work, and indeed they were able to show a suite of complex effects (Chapter Five).

During the normal life cycle of *C. elegans,* the nematode passes through four larval stages (L1-L4), the first of which takes 16 hours and the second to fourth (L2-L4) stages taking 12 hours each under normal environmental conditions. Maturity is reached within 3 days as already mentioned. There is also an alternative larval stage, the dauer larvae, which may present upon increasing population density or restricted food availability. The dauer stage of *C. elegans* is illicited when environmental conditions indicate that they are unfavourable for reproductive success (Riddle *et al.*,1997).

The dauer larva is in a state of arrested development, sealed by a thick outer cuticle in order to survive harsh conditions, and it can only be entered into at L2 stage in response to chemosensory signals or pheromones detected in the L1 phase (Cassada and Russell, 1975). This dauer then has the ability to survive considerably longer than the normal two week life

cycle, and potentially up to eight times longer (Klass and Hirsh, 1976). Upon return of hospitable environmental conditions, the dauer larva can exit this state of arrested development by beginning to feed and will reach reproductive maturity as normal (Riddle and Albert, 1997). This mechanism is controlled by the amphid neurons of the worm which are responsible for chemosensation of dauer and other pheromones or the bacterial food source. The work presented in this thesis and the publication associated with it (McConnell *et al.*, 2015) explore a previously unreported dauer phenotype as a result of treatment with EHL.

1.7 Nanoparticles

Colloidal gold nanoparticles are not in themselves a new discovery and have long been exploited in both science and interestingly in art for their optical properties. Differing sizes absorb light of different wavelengths which produce a variety of coloured colloidal suspensions. Their applications in recent years have increased enormously and are used routinely in both material science and within biomedical sciences as therapeutic agents and drug delivery vehicles. The optical and electronic properties of nanoparticles are able to be manipulated by the alteration of surface chemistry or the use of differing sizes of particles. Gold nanoparticles are of particular interest in biomedical research mainly due to their inherently low toxicity and compatibility in a variety of biological applications, but also due to the optical properties referred to earlier. These are conferred by their localised surface plasmon resonance and the ability to be conjugated to a variety of useful substrates such as peptides and therapeutic drugs (Hutter and Fendler, 2004). Whilst a detailed explanation of the physics behind nanoparticles is beyond the scope of this thesis, the basic concept of note is that the surface plasmon resonance of a particle can be excited at a specific frequency of light depending on the size, shape, and surface texture of a particle. This results in strong and

characteristic light scattering and intense energy absorption. Localised surface plasmon resonance is the oscillation of charge density in metallic particles when subjected to excitation (Hutter and Fendler, 2004); a property used in tumour destruction through the technique of photothermal ablation (Hutter and Maysinger 2011). Gold nanoparticles have also shown extensive potential in conjugated drug deliver in cancer studies (Ahmad *et al.*, 2013).

Nanoparticles can be functionalised with ligands, with high affinity and specificity for target cells such as Tn+ which is where the ultimate interest in conjugating a RIP such as EHL presents an opportunity. With the ability to be functionalised with both therapeutic and imaging agents simultaneously, they are indeed a powerful tool in cellular studies (Ahmad *et al.*, 2013). The work presented herein with nanoparticles is a preliminary investigation into the viability of the conjugate and set of initial assays to establish if the EHL biological effect is altered by it, providing insight for future research. All of the synthesis, functionalisation and quality analysis work on the particles has been conducted by the Bioscope Group at the University of Lisbon with EHL supplied to them through this work. All biological work was conducted by the author.

1.8 Aim and objectives of this thesis.

The aim of this thesis was to conduct an in depth study of the structure and biological effects of the Winter Aconite (*Eranthis Hyemalis*) lectin (EHL). Other Lectins and RIPs have shown promise in biomedical research and have multiple biotechnological applications. Due to the conserved nature of the RIP family of proteins, it is likely that EHL could also have potential in similar applications. In order for that potential to be fully explored, the underlying structure and biological activity of EHL needed to be elucidated. As a first step, the existing extraction and purification protocol was refined. This was to allow increased and consistent protein

yields for downstream applications and further study including the conjugation of gold nanoparticles.

The investigation of the structure of the EHL protein was then carried out. This structural work had two aspects. Molecular biology techniques were employed in order to find the underlying gene sequence of the protein. Crystallisation screening was undertaken and the results of that were used to obtain diffraction data to establish the 3 dimensional structure of the protein.

A set of biological assays were conducted to explore the potential toxicity of the protein in living organisms. The biological assays initially focused on the model organism *Caenorhabditis elegans*. This work included using wild type N2 strains of the nematodes and also an exploration of the effects EHL has on genetically mutated strains; the work was carried out with the natural protein and with EHL conjugated gold nanoparticles. Finally, preliminary studies of the potential cytotoxic effect of EHL on mammalian cancer cell lines were conducted in order to develop a viable method for future research.

Specifically, the objectives of this thesis were:

1. To establish improved methods for the extraction and purification of *Eranthis hyemalis* lectin, and to carry out analysis to validate the success of these such as SDS-PAGE and agglutination testing. This work will be reported in Chapter Two.

2. To investigate the underlying gene sequence of EHL. This part of the research is presented in Chapter Three.

3. To investigate conditions that would allow crystallisation of *Eranthis hyemalis* to occur and use any resultant diffraction data to build the structure of the protein. Chapter Four contains the results of this investigation.

4. To investigate the effect of *Eranthis hyemalis* lectin on the model organism *Caenorhabditis elegans* and to characterise the effects in genetic mutants (Chapter five).

5. To establish the viability of a functionalised EHL nanoparticle conjugate and its potential effect on the cytotoxicity of EHL (Chapters Three and Five).

6. To develop a protocol for testing mammalian cancer cell lines with EHL (Chapter six).

In summary, *Eranthis hyemalis* or Winter Aconite is a late winter flowering perennial plant of the family *Ranunculaceae*. *E.hyemalis* expresses lectin activity and the lectin found in the tubers of the plant is a sugar specific (N-acetyl-D-galactosamine) Type II Ribosome Inactivating Protein. Type II RIPs have shown antineoplastic properties and hence have potential as therapeutic agents. A modified protocol for the extraction and purification of the *E. hyemalis* lectin (EHL) using affinity chromatography has been developed. The biocidal properties of EHL have been investigated against *Caenorhabditis elegans.* Crystallisation trials have yielded crystals which have produced diffraction data at 1.6 and 1.8 Angstroms resolution in two crystal systems, with space group $P2_1$ and $P2_12_12_1$ respectively. Molecular replacement has been used to extract phases and a structure solution generated for EHL using a RIP with 40% homology which has reported anti-cancer properties.

The first aim of the research undertaken within this thesis was to establish an extraction and purification protocol as mentioned, which would provide a homogeneous, pure and consistent yield of EHL in sufficient quantity to allow further study of the characteristics and structure. Once this was established, toxicity studies were carried out using *Caenorhabditis elegans* to examine the effect of EHL on lifespan and development, reproduction and the induction of abnormal dauer formation discovered in the first experiments conducted. The data collected indicated that EHL binds to the amphid neurons in *C.elegans* and causes a

wide variety of concentration dependent effects including the induction of a state of arrested development, the dauer state.

In order to elucidate the structure of the protein, crystallographic screening was undertaken. A number of large single crystals produced diffraction data. These data were used to construct a model structure using a molecular replacement solution. The model is presented, as a work in progress with interesting observations as to potential glycosylation sites and active site residues.

A preliminary investigation into the viability of a conjugation of EHL to gold nanoparticles was undertaken in collaboration with The Bioscope group (University of Lisbon) and the conjugate was also tested for biological activity. Initial results indicated that the nanoparticles caused a cessation of the toxic effect of EHL.

In the final stages of the research presented herein, EHL was tested for antineoplastic activity against a number of cancer cell lines using a variety of methods including the use of Resazurin, a redox reagent which is only reduced by metabolically active cells, to develop a viable protocol for future studies. Preliminary results indicate that EHL may act as a mitogenic agent on some cancer cell lines, causing proliferation, which also presents an opportunity for onwards research. Future work will continue with a view to completing a successful protocol for examining the effect of gold nanoparticle conjugation and EHL on various mammalian cancer cells.

Chapter 2: Extraction, isolation and purification of *Eranthis hyemalis* lectin (EHL) and the synthesis of EHL conjugated functionalised gold nanoparticles.

The Extraction and purification of EHL as reported here has been published in:

McConnell, M.T., Lisgarten, D.R., Byrne, L.J., Harvey, S.C. and Bertolo, E., 2015. Winter Aconite (Eranthis hyemalis) Lectin as a cytotoxic effector in the lifecycle of *Caenorhabditis elegans*. *PeerJ*, *3*, p.e1206.

The synthesis and characterization of the functionalised Gold nanoparticles was performed in the BIOSCOPE-PROTEOMASS Laboratory (FCT-UNL, Portugal) by Dr Javier Lodeiro and Ms Jamila Djafari. The BIOSCOPE group is headed by Prof. Carlos Lodeiro and Prof. Jose Luis Capelo. The BIOSCOPE group's method of synthesis and results of the work are reported here.

2.1 Introduction

Winter Aconite Lectin or EHL has been the target of previous investigation and extraction protocols were published by Cammue, Peeters and Peuman (1985), Kumar *et al.* (1993) and George *et al.* (2011). All three methods showed various levels of success, with yields ranging from 380 µg/ml to a maximum of approximately 1mg/ml. The aim of this work was to develop a method which returned consistently high yields, which would facilitate downstream applications; additionally, the aim was to obtain a product of high purity in order to conduct crystallisation studies. The extraction process of EHL was originally based on previously published work which stated that the protein could be eluted from the chromatography

column using distilled H_2O (Cammue, Peeters and Peumans, 1985). However protein yield was low and inconsistent, leading to the development of a new protocol which combined elements from all previously reported work to achieve the desired outcome.

Column chromatography is a method for the separation of complex mixtures of molecules based on various sets of characteristics. This can be separation based on size, such as in gel filtration columns (size exclusion), charge (ion exchange), hydrophobic interaction and as in this work, affinity chromatography which exploits a particular binding affinity between the stationary and mobile phase of the chromatography reagents. The stationary phase or matrix of the affinity column used in this work was an agarose bound glycoprotein fetuin from fetal calf serum (Sigma Aldrich). Bovine fetuin is glycosylated with a high percentage (>70%) of N-linked sialylated branched oligosaccharides with mostly GlcNac or Gal terminally sialylated residues (Green *et al.,* 1988). This branching allows for a great many lectin interactions and makes it an ideal choice for purifying carbohydrate binding proteins such as Type II RIPs. Work conducted in 1975 (Sela, Wang and Edelman, 1975) showed that a fetuin conjugated matrix could efficiently be used as a single affinity purification step for a variety of lectins with differing carbohydrate specificities, and thus was the matrix of choice in this research. In addition to the highly efficient binding affinity of fetuin-agarose, as the research progressed it became apparent that an additional purification step would be advantageous. The purpose of this additional purification step would aid in the maintenance and lifespan of the column matrix, and would help to achieve an even greater level of purity of protein for crystallographic studies. To this end an ammonium sulfate precipitation step was added. Ammonium sulfate ($(NH_4)_2SO_4$) precipitation of globular proteins works on the principle of volume exclusion, whereby the high concentration of ammonium sulfate molecules increase

the surface tension of the solvent. This in turn increases the hydrophobic interactions between protein and water, encouraging folding and reducing the solubility of the protein. This causes precipitation and aggregation of the proteins (Wingfield, 2001). As differing protein structures all have varying degrees of exposed hydrophilic surface amino acid residues, they will be 'salted out' by the $(NH_4)_2SO_4$ at different percentage saturations dependent on their solubility. This property can be exploited to fractionate crude protein mixtures. Ultimately, the reduction of interactions at the water/protein interface results in an increase in protein-protein interactions, resulting in precipitation. Exposing the combined $(NH_4)_2SO_4$ - protein solution to centrifugal forces results in the formation of a protein pellet (from the aggregated proteins), which can then be resuspended in buffer. It was observed that, over time, contaminating residues were retained on the column in the manual system; the useful lifespan of the matrix was shortened and yields reduced. Hence the ammonium sulfate precipitation step was introduced to reduce contaminant proteins and effectively add in a further purification step. An initial $(NH_4)_2SO_4$ saturation of 80% was successfully trialed, which subsequently facilitated the introduction of a graduated $(NH_4)_2SO_4$ cut. This has the advantage of allowing proteins above or below a particular solubility threshold being discarded, and a purer crude extract being prepared for affinity chromatography. EHL activity was monitored by testing for haemagglutination of erythrocytes.

The possibility of conjugating the purified EHL to gold nanoparticles was explored in collaboration with Bioscope group at the University of Lisbon (Portugal). Gold nanoparticles (AuNPs) have attributes which can be exploited for many aspects of cell biology research, such as their optical properties, which make them powerful imaging tools and they are also used as drug conjugated delivery systems (Ahmad *et al.*, 2013). Despite the low inherent cellular

toxicity of AuNPs, cancer cells are more vulnerable to destruction by nanoparticles than are healthy cells, and when combined with a chemotherapeutic agent, can increase localised drug concentrations whilst lowering overall systemic toxicity (Kodiha *et al.*, 2015). This makes them an interesting target to conjugate with a Type II RIP. One of the obstacles faced by cellular nanotechnology is initial cell specific surface binding. With EHL potentially having an affinity for the GalNAc repeating tumour specific Tn antigen (Ju, Otto and Cummings, 2011), the work here reports the successful conjugation of EHL to functionalised gold nanoparticles with a view to exploring the impact on the native protein and the biological effects of the conjugate.

2.2 Material and Methods

2.2.1 Preparation of affinity chromatography column

Initial protocols used a manually operated gravity fed Biorad econo column with a ratio of 1:20 width to height ratio (1.5 cm x 30 cm) which was packed at room temperature to prevent bed expansion with a custom chromatography medium. A bed volume of 8 ml was applied with 16 ml Fetuin-Agarose in solution with 0.5 M NaCl and immobilised on cross-linked 4 % beaded agarose (Sigma-Aldrich Company ltd, UK). The column matrix was allowed to settle with an additional head volume of 16 ml 0.5 M NaCl solution being applied to aid in bed compaction. The column was allowed to settle for 2 hours before being equilibrated with 200 ml of phosphate buffered saline (PBS). This column was connected to a UV path monitor, peristaltic pump and chart recorder to monitor peak absorbance. The process was mainly carried out in a cold room at 4°C until failure of the chart recorder occurred due to the low operating temperature. The arrival of the AKTA Pure protein purification system in the research labs meant that a column with greater pressure tolerances was required and so

initially a Pharma LC column was poured with the same media for use in conjunction with the AKTA. Subsequently a GE column was purchased and used due to its superior packing and running characteristics in conjunction with the AKTA Pure system.

2.2.2 Extraction of EHL

Dry mass of tubers from *E. hyemalis* ranged from 20g to 100 g and were supplied by Eurobulbs Ltd (UK). A final mass of 60g appeared to give the most consistent yield, as higher amounts resulted in higher levels of lipids and contaminants as well as aggregation and precipitation of the protein. The tubers were finely sliced and chopped before being homogenised with 250 ml of ice cold PBS and left to settle on ice for 30 minutes. The homogenate was removed and stored, and the remaining slurry was mixed with a further 250 ml of PBS. The two fractions were then combined and stirred at 4°C for 4 hours. The homogenised mixture was then centrifuged (Sorvall RC6 plus HSC) at 20,000 g for 30 minutes. The supernatant was retained and frozen at -80°C overnight in order to induce aggregation of any remaining lipid content in the sample. The sample was then defrosted and filtered through 3MM Whatman filter paper or muslin before undergoing a further centrifuge cycle of 20,000 g for 20 minutes. In the initial development of the protocol, the clarified supernatant was then applied to a Biorad chromatography column. Subsequently, as previously mentioned, the protocol was amended to include an ammonium sulfate precipitation step, in which solid ammonium sulfate ($(NH_4)_2SO_4$) was slowly dissolved in the crude filtered extract whilst under constant stirring at 4°C. The $(NH_4)_2SO_4$ was added to the crude extract according to the published ammonium sulfate precipitation tables dependent on volume and temperature. Lectin activity was found to be retained in the 40% ammonium sulfate fraction pellet (as seen in figure 2.5). This fraction was subsequently centrifuged at 10,000 g for 35 minutes. The resulting pellet was

suspended in a minimal volume of phosphate buffered saline and extensively dialysed (3 x 200x volume) reservoir changes over 20-24 hours) using a semi permeable dialysis membrane (20 kDa molecular weight cut off (MWCO)) to remove the ammonium sulfate. This step is essential when affinity chromatography is the downstream purification step as the salt molecules will also interfere with column binding due to hydrophobic interactions with the chromatography media.

2.2.3 Purification of EHL

In the first instance the crude extract was applied to the gravity column using a Watson Marlow SciQ400 peristaltic pump and allowed a contact time of 30 minutes to increase adsorption of target protein. The sample was also reapplied to the column to maximise specific binding. Non target proteins were then eluted consecutively with PBS and 1 M NaCl until absorbance at λ_{280} fell below 0.01. The column was then equilibrated with PBS and subsequently EHL was eluted with the previously published elutant of distilled water (Cammue, Peeters and Peuman, 1985). The resultant protein yield was extremely low, between 100 and 400 μg/ml, and as the column became less efficient due to residual crude contaminants, levels dropped to zero. A novel purification protocol (figure 2) was developed initially by trialing new elutants, such as lactose and n-acetyl galactosamine; however, the protein was retained on the column. However, elution with 40 mM 1,3 diaminopropanol (DAP) produced a sharp peak (Figure 2.1) and a Bradford assay was performed using Bovine Serum Albumin (BSA) as a standard, in UV micro-cuvettes (Fisherbrand). This confirmed a relative increase in yield of up to 1.3 mg/ml. As the project proceeded, a move towards using the Qubit flourometric assay for quantification of protein was taken. With the introduction of DAP, the protein initially appeared inactive as no agglutination of erythrocytes was observed.

Once the eluant was neutralised with 2-Amino-2-hydroxymethyl-propane-1,3-diol hydrochloride (Tris-HCl) at pH 7.0, activity returned. The arrival of a GE healthcare AKTA Pure protein purification system, facilitated the introduction of control over additional paramaters such as column pressure, flow rate and undiluted application of the sample to the new GE column. Elution of the target protein from the column was carried out and collected using the AKTA fraction collector in 0.5 ml aliquot. This allowed precision collection of defined peak fractions with no sample dilution from pre and post column washes. After neutralization as described with Tris HCl, samples were dialysed against PBS in 3 buffer changes. These each consisted of 200 x the sample volume, including a final overnight step. This became the foundation for the new and final protocol including ammonium sulfate precipitation and fractionation, dialysis, 0.45 and 0.22 µm filtration, use of the AKTA purification system with elution using 40 mM 1,3 diaminopropanol and subsequent neutralization with Tris HCl, buffer exchange with Pierce Slide-a-lyzer dialysis cassettes (20 kDA MWCO) and finally concentration as required using a Vivaspin PES device (30 kDa MWCO) A schematically represented overview of the process can be seen in Figure 2.

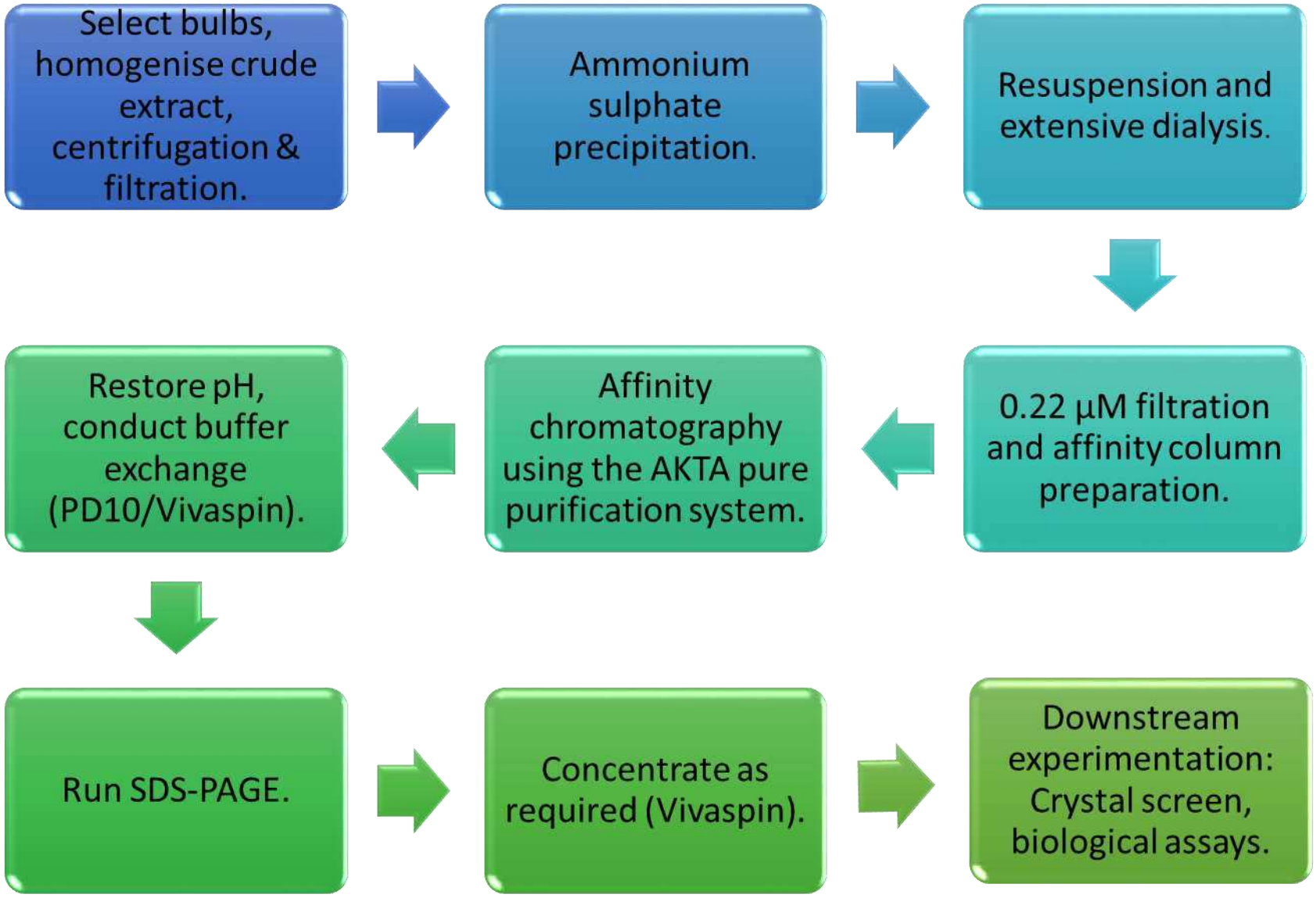


Figure 2. flow chart showing the process of extraction and purification of EHL

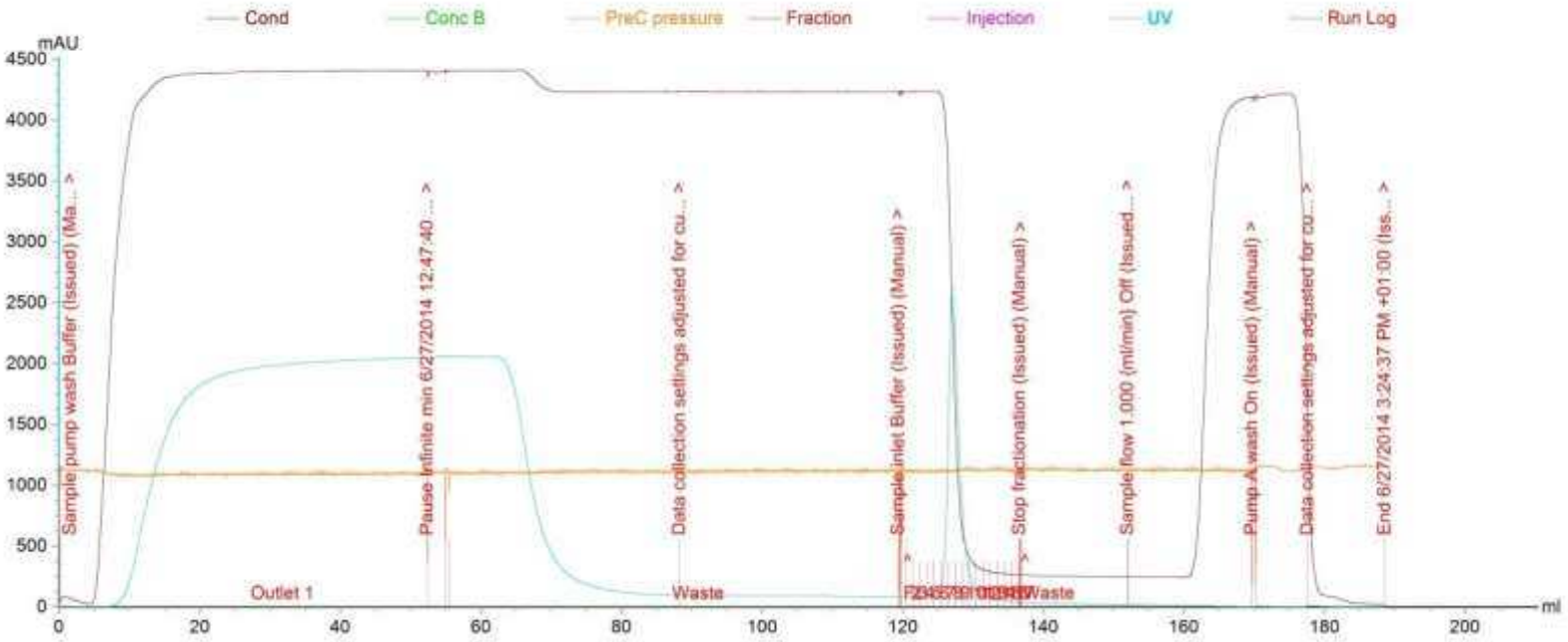


Figure 2.1 Typical AKTA Pure run with EHL using manual commands in Unicorn. Note the EHL peak at approximately 125 ml.

2.2.4 Analysis of EHL

Purified EHL was tested for agglutination activity using defibrinated rabbit erythrocytes (TCS Bioscience), with 20 µl of post column eluant added to a 20 µl sample of erythrocytes in a welled microscope slide. The purified EHL was also analysed by SDS-PAGE, with both reduced and non-reduced samples electrophoresed on 12 % protogel quick cast gels and subsequently stained with Coomassie Brilliant Blue. Initially concentration was measured by monitoring absorbance at λ280 nm using a ThermoFisher Evolution 201 UV spectrophotometer. Typically, absorbance was found to be outside the linear range. Consequently 1/20 dilutions were carried out to produce an absorbance within the linear range. This method was found to be prone to inconsistencies due to natural variation within yields, and thus a move towards quantification using chemofluorometry was made. A Qubit 2 fluorometer was then routinely used to measure specific protein concentrations in post column purified fractions. The Qubit fluorometer utilises bovine serum albumin (BSA) at concentrations of 0ng/µl, 200 ng/µl and 400 ng/µl to establish a concentration curve. Typical EHL yields are in the range of 1.3mg/ml to 2.4mg/ml. Differences can be attributed to moisture content within plant tubers and natural seasonal variation.

During attempts to concentrate the protein, it was discovered that Centricon Plus-70 Centrifugal Filter Units appear to retain the protein on the regenerated cellulose membrane, possibly due to lectin binding. Therefore EHL was concentrated successfully for downstream applications using Sartorius Vivascience vivaspin 6 (30,000 MWCO) Polyethersulfone PES membrane. The Vivaspin tubes were also used for buffer exchange after purification to remove DAP where appropriate. A further preliminary investigation into the specificity of EHL for agglutinating blood type specific erythrocytes was carried

out. The crude method which will undergo further development required the addition drop wise of four differing blood types to a 12 well plate, 50 µl of EHL was added to each and an incubation time of 5 minutes allowed prior to imaging (Figure 2.4).

2.2.5 Synthesis of EHL conjugated Gold nanoparticles

Gold nanoparticles functionalised with EHL (AuNPs+EHL) were synthesised by ligand exchange with citrate gold nanoparticles. The adsorption of protein onto gold nanoparticles to produce a bioconjugated nanosystem is presented as an excellent option to obtain a protein/nanoparticle hybrid system without any covalent modification to the protein which can result in modified activity. To minimize changes to the protein, citrate gold nanoparticles were used as the nanoparticle support. To this end Au+Citrate NPs were synthesised following a modified Frens synthesis (Frens, 1973; Shang *et al.*, 2007). For this synthesis, an aqueous solution containing 1 mM of tetrachloroauric acid (125 mL of MilliQ water) was heated until boiling and then 12.5 ml of citrate trisodium 1% (p/v) on boiling was added. The reaction mixture was allowed to boil for 5 minutes at which point the solution colour turns from blue to ruby red, and was then left stirring overnight at room temperature. The Citrate gold nanoparticles (AuNPs+Citrate) obtained were used without a washing process in order to avoid the loss of nanoparticles during the centrifugation and subsequent washing cycles.

AuNPs+Citrate were analysed by UV-Vis spectroscopy, Dynamic Light Scattering (DLS), Zeta Potential analysis and Transmission Electron Microscopy (TEM). Samples were prepared for TEM by allowing a 5 µl drop of sample dry on TEM copper grids which were coated with Formvar/carbon with a 400 mesh (pore size 42 µm). Zeta potential analysis provides a quantitative measure of the electrostatic charge between particles, and provides an

indication of the stability of a colloidal system. By UV-Vis, the red/pink colloidal solution of AuNPs+Citrate present a Surface Plasmon Resonance (SPR) band at ca. 519 nm. In TEM analysis it was observed that the Au+Citrate NPs present a quasi-espherical shape with an average size of 12.7 ± 2.5 nm (Figure 2.2). DLS studies indicated that Au+Citrate NPs have a z-average of 19.28 nm; z-average indicates the average radii of particles based on intensity measurements. The Zeta-Potential value obtained (-43.6 mV/cm) confirms the high colloid stability provided by citrate molecules adsorbed on the surface of the nanoparticle. Assuming that the gold reduction in gold nanoparticles was quantitative and of high monodispersity, the concentration of initial gold colloid is equal to 8.61×10^{13} NPs/mL. The Molar concentration was calculated using previously reported methods (Liu *et al.*, 2007; Fernandez-Lodeiro *et al.*, 2013). Assuming a spherical shape and uniform face centered cubic (fcc) structure, the molar concentrations of the gold nanoparticle (AuNP) solutions were calculated using Equations 2.1 and 2.2

$$N = \frac{\pi}{6}\frac{\rho D^3}{M}$$

(equation 2.1)

Where N is the number of atoms per AuNP, D is the average core diameter ((nm) of particles, ρ [g.cm^{-3}] is the density of face centered cubic (fcc) of gold (19.3 g.cm^{-3}), and M [g.mol^{-1}] is the atomic mass of gold (196.97 g.mol^{-1})

$$C = \frac{N_T}{NVN_A}$$

(equation 2.2) where C is the molar concentration of AgNPs or AuNPs, N_T is the total number of gold atoms added as $HAuCl_4$,

N is the number of NPs, V is the volume of the reaction solution in litres, and N_A is Avogadro's constant (number of atoms per mole).

From equation 2.1, the result is $N= 30.896D^3$

With a 12.7 nm diameter and 0.1375 L of total volume reaction, using equation (2.2), a concentration equal to $1.4x10^{-8}$ Molar is obtained.

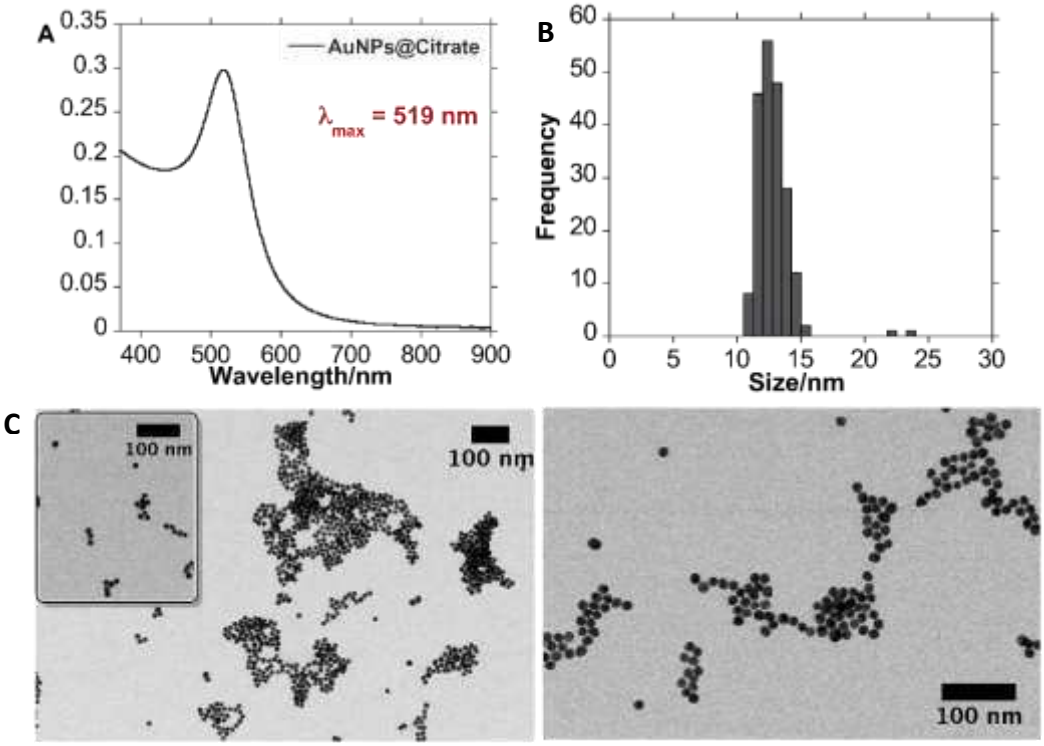


Figure 2.2: UV-Vis spectrum, histogram and images TEM of AuNPs+Citrate. (The histogram was generated counting 200 nanoparticles using the software package ImageJ).

The functionalisation of AuNPs+Citrate with EHL protein was facilitated by incubation of gold nanoparticles with a PBS+protein solution (the addition of Tris HCl buffer caused immediate aggregation of gold nanoparticles preventing proper bioconjugation). The

protein was suspended in PBS buffer by centrifugation and the concentration quantified using the Bradford technique indicating a composition of 1090 ± 40 µg/mL.

A series of six experiments were conducted examining differing protein solution amounts added onto AuNPs+Citrate (25 µL, 50 µL, 100 µL, 200 µL, 300 µL and 500 µL) in order to analyse nanoparticle stabilisation as a function of protein quantity. The amount of protein solution used for each sample was added to 5 mL of AuNPs+Citrate and left under vigorous stirring at room temperature in darkness for 2 hours to maintain protein integrity and promote effective functionalisation.

The NPs obtained were then centrifuged at 18187 x *g* for 25 minutes and suspended in PBS. A second centrifugation cycle was performed and the nanoparticles were suspended in milliQ water. The first supernatant was filtered (cellulose filter 0.22 micrometers) and quantified by the Bradford technique in order to determine the protein concentration at the NP's surface. Each set of EHL functionalised gold nanoparticles (AuNPs+EHL) were analysed by UV-Vis spectroscopy, DLS, Zeta Potential techniques and TEM. All the results obtained are indicated in Table 2.1

2.2.6 Differential agglutination test

EHL was applied in 30 µl aliquots to 4 differing blood types in a 12 well plate, the first well at the top of the images in figure 2.4 (see results) having 1 aliquot, then 2 in the second well and 3 in the third (total voume of 30,60 and 90 µl). The type B sample was applied to a single well due to limited quantity and was therefore given the maximum 3 aliquots. Samples were incubated for 5 minutes at room temperature and photographed.

2.3 Results

2.3.1 Purification and characterization of EHL

The concentration of natural yields of EHL obtained varied widely initially between 200ug/ml and 2.8mg/ml, with the improved protocol, AKTA system and the use of the Qubit flourometer however, a consistent and typical yield of around 1.5 mg/ml was established. An intense agglutination response of rabbit erythrocytes was exhibited when the purified extract was applied to cells and thus confirmed the presence of EHL (Figure 2.3a and b). SDS-PAGE analysis produced characteristic reduced protein bands at circa 31 and 28 kDa as well as an unreduced band circa 50-60kDa (Figure 2.3c); these values are consistent with those previously reported in the literature (Kumar *et al.,* 1993; Cammue, Peeters and Peuman, 1985; George *et al.*, 2011). A preliminary investigation as to whether EHL showed any specificity towards erythrocytes was also carried out. EHL shows greatest affinity towards type O, and a reduced agglutination towards type B. However as lectin specific to Type B are rare, this is of particular note for further research. Agglutination on addition of EHL was absent in the type A and AB samples (Figure 2.4). Although the agglutination titre assay requires significant modification to establish the exact relationship between the amount of EHL required to agglutinate a specific blood type more quantitatively, visual inspection is however sufficient to give initial guidance. The findings herein reported have value at an observational scale, to inform future experiments.

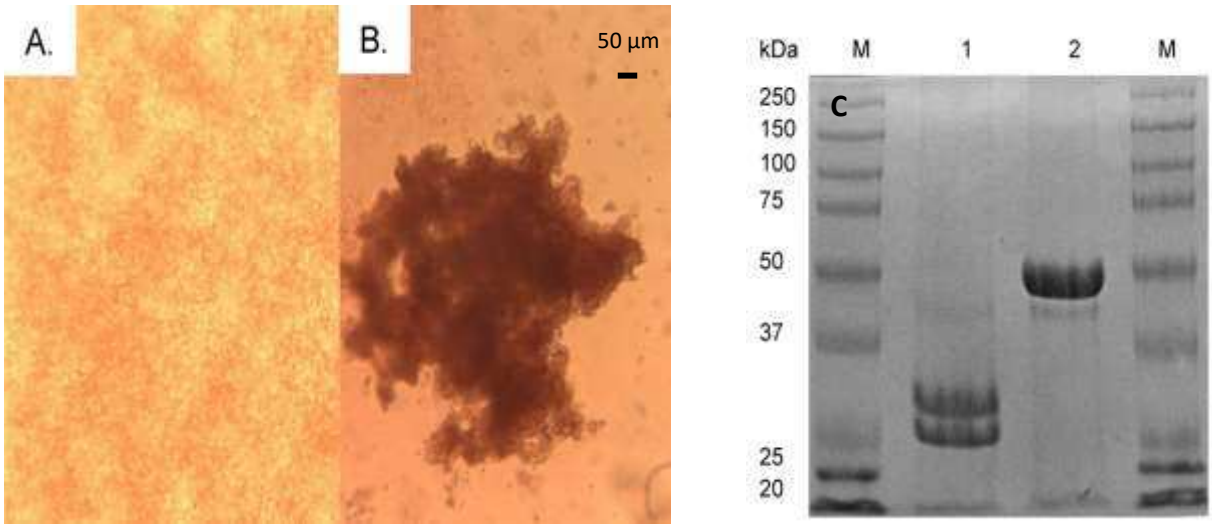


Figure 2.3 EHL is a heterodimeric Type II RIP which induces agglutination of erythrocytes (**A**) 100 x magnification, Rabbit erythrocytes treated with PBS, no clumping of erythrocytes observed (**B**) incubation with EHL results in intense agglutination. (**C**) 12% SDS-PAGE analysis of EHL. Lanes: M, Molecular weight marker; 1, β-mercaptoethanol reduced EHL; 2, non-reduced EHL.

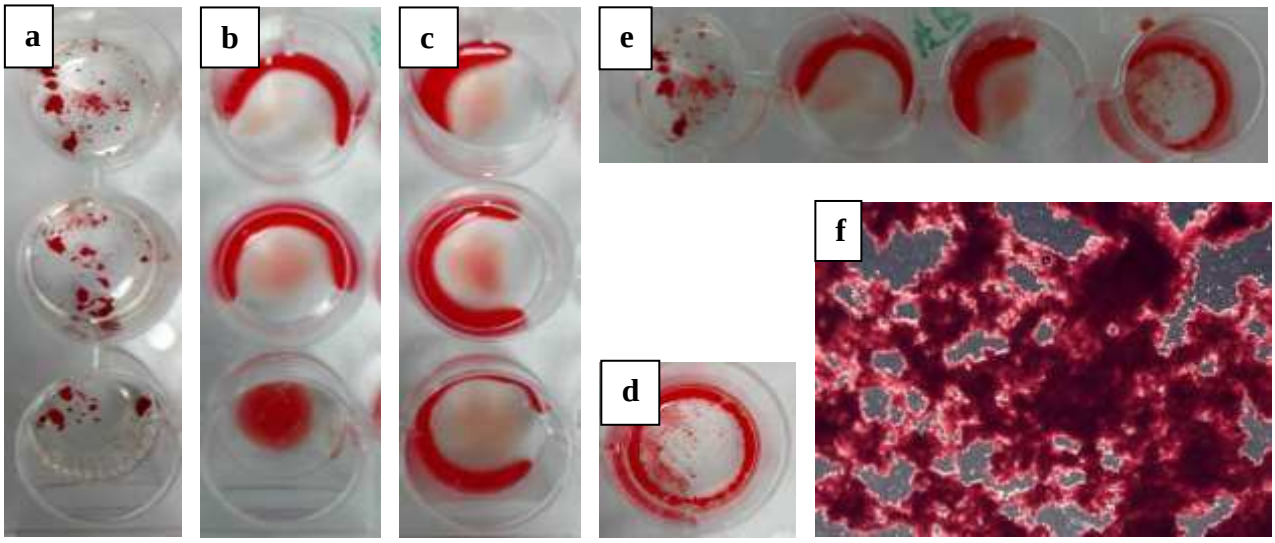


Figure 2.4 EHL shows blood type specific agglutination of erythrocytes. **a**) Blood Type (BT) O, **b**) BT A, **c**) BT AB, and **d**) BT B (sample availability restricted), **e**) left to right BT O; A; AB; B. **f**) Zoom image of BT O under 65 x magnification on a dissecting microscope with cross linkage of erythrocytes visible.

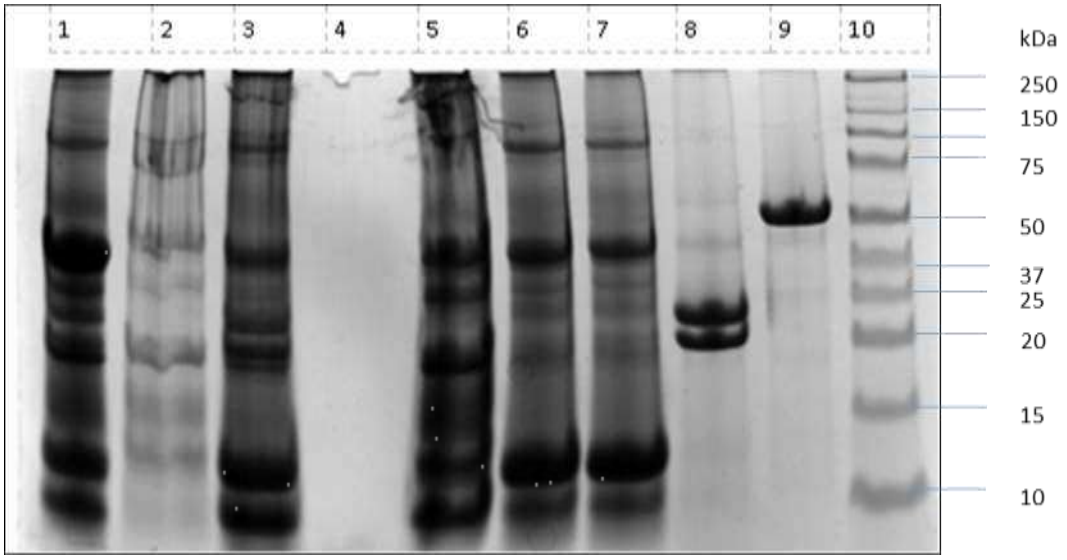


Figure 2.5 SDS PAGE 12% reducing conditions in lanes 1-8 showing extraction process fractions. Lane 1) Crude extract, 2)40% ASP supernatant, 3) 40% ASP Pellet, 4) 80% ASP supernatant, 5) 80% ASP pellet, 6) Affinity column flow through; unbound protein, 7) Affinity column flow through; non-specific bound, 8) Affinity column target protein (EHL) reduced with β-mercaptoethanaol, 9) non reduced EHL and 10) Molecular Marker.

2.3.2 EHL conjugated gold nanoparticle synthesis

The results presented in Table 2.1 show that adding lower quantities of protein (25 and 50 µL) is not sufficient to properly stabilize the AuNPs. Indeed, the high DLS values confirm the aggregation of AuNPs+EHL with a lower amount of protein (Figure 2.2). With the addition of a higher protein quantity (200, 300 and 500 µL), it was observed that there was a saturation on the NP's surface shown by the protein concentration values obtained in the supernatant. This demonstrates that there is an excess of protein in the medium. For these cases (200 and 300 µL), the first supernatant was not completely clear and still contained NPs which can produce an erroneous reading using the Bradford technique. This is because

in spectroscopic quantification the wavelength used may overlap with the Plasmon band of AuNPs, resulting in the appearance of increased protein values.

Volume Protein added (in PBS) (µL)	25	50	100	200	300	500
Volume total of reaction solution V_{tot} (µL)	5025	5050	5100	5200	5300	5500
Protein mass in the reaction V_{tot} (µg)	27.25	54.5	109	218	327	545
Protein concentration in the supernatant (µg/mL)	3.1 ± 0.1	7.4 ± 0.3	4.0 ± 0.1	40 ± 2	61.1 ± 0.4	84 ± 3
Protein concentration in the V_{tot} supernatant (µg)	15.6	37.4	20.4	208	323.8	462
Protein concentration into the functionalised NPs solution V_{tot} (µg)	11.7	17.1	88.6	10	3.17	83
Z-Average value (nm)	266.37	90.26	54.41	51.72	60.79	51.23
Zeta Potential (mV/cm)	-23.13	-19.83	-27.8	-24.63	-20.23	-29.37

Table 2.1: AuNPs+EHL solution composition for each experiment, DLS and Zeta Potential Values and protein amount on the nanoparticles.

These data indicate that a better functionalisation was achieved with the addition of 100 µL of protein (1090 ± 40 µg/mL) into 5 mL of nanoparticle solution. The AuNPs+EHL are more stable in this condition due to the high percentage of protein initially functionalised and the stability of the colloidal solution.UV-Vis spectroscopy confirms that the functionalisation of the nanoparticles is optimal with the addition of 100 µL of protein (Figure 2.6). A redshift in the SPR band from 519 nm to 528 nm indicates a compositional change of the NP's surface, which confirms a NP surface/ligand exchange between the citrate molecule and the EHL protein. The z-average obtained for AuNPs+citrate of 19.28

nm increased to 54.41 nm and hence demonstrates the absorption of protein on the gold surface.

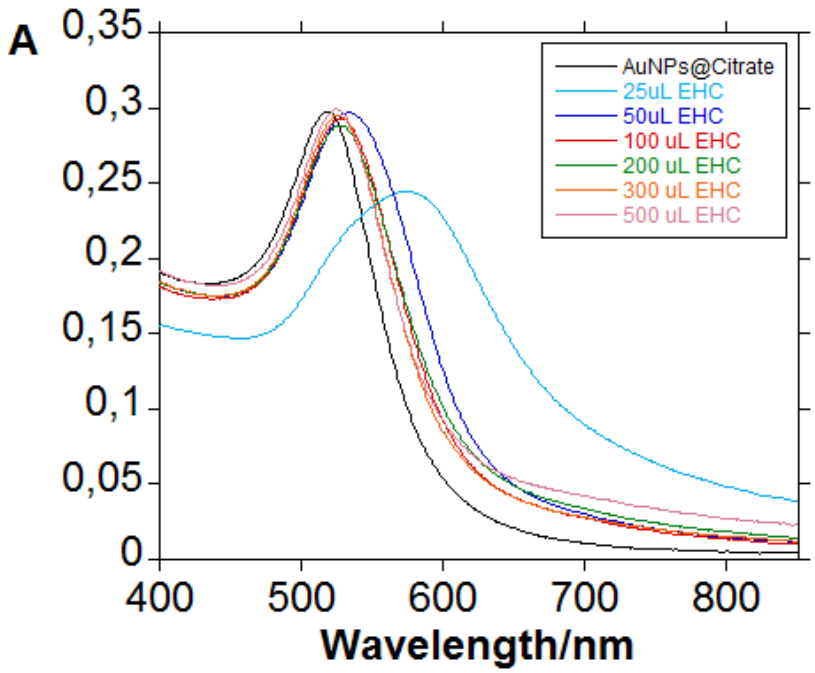


Figure 2.6: Au+EHL NPs UV-Vis spectra as a function of protein amount added (25µL, 50µL, 100µL, 200µL, 300µL and 500µL) in 5mL of Au+Citrate NPs.

Furthermore, 25 µL of protein did not allow the stabilisation of the colloidal system and induced nanoparticle aggregation due to partial functionalisation, with an SPR band centered ca. 573 nm. The presence of aggregate NPs was observed in TEM analysis (Figure 2.7).

Due to the dimeric structure of the protein, it can interact as a linker between two discrete NPs. Alternatively, due to incomplete functionalisation of the AuNP's surface, it can modify

the isotropy charge produced by the citrate adsorbed, resulting in the formation of aggregates. With the addition of 500 µL of EHL, the NPs also presentated aggregates (Figure 2.7). One possible reason for observing more aggregated NPs with 500 µL of protein solution might be as a resulting product of adding 10% v/v of phosphate buffer, giving high concentrations of salt which can then induce aggregation by the citrate blocking charges on the nanoparticle's surface during the adsorption of the protein.

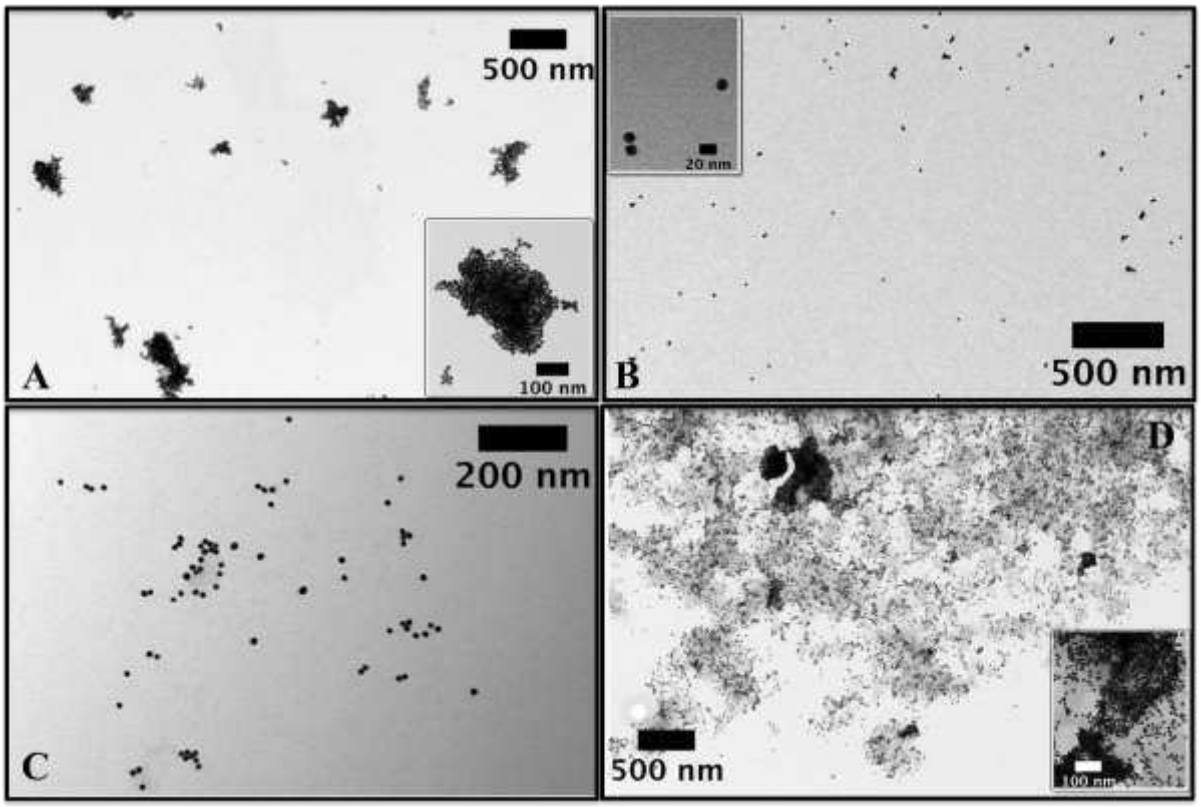


Figure 2.7: TEM images of AuNPs@Lectin with different amounts of protein added (A) 50 µL, (B) 100 µL, (C) 300 µL and (D) 500 µL. (The nanoparticles are submitted to two centrifugation cycles (18187 x *g* x 25 minutes) followed by resuspension in MQ water).

From the UV-Vis, DLS, Zeta Potential measurements and TEM image results, it was concluded that the amount of EHL protein needed for an optimal functionalisation of

AuNPs+Citrate was 100 µL. For lower values, protein quantity added was insufficient to ensure complete NP functionalisation, and for a value higher than 100 µL, due to the excess of protein added with the PBS volume used, the functionalisation was also not effective or complete.

A second batch of AuNPs+Citrate (denoted as AuNPs+Citrate-2) were synthesised, with functionalisation being completed without the centrifugation cycles and following the protocol described previously. The protein concentration used in this part of the study was 1.1 mg/mL in PBS to simulate the initial conditions in batch 1.

In UV-vis, the red/pink colloidal solution of AuNPs+Citrate-2 present a surface plasmon resonance band at ca. 519 nm (Figure 2.8). DLS studies indicated that AuNPs+Citrate give 21.56 nm in Z-average and a Potential zeta equal to – 39 mV/cm, thereby confirming the NPs stabilisation by citrate molecules. These NPs were functionalised with 100 µL of EHL solution (1.1 mg/mL) to ensure optimal functionalisation, centrifuged at 18187 x *g* for 25 minutes and suspended in PBS. A second centrifugation cycle at 18187 x *g* for a further 25 minutes was carried out and finally suspended in MQ water. AuNPs+EHL-2 presents an SPR band at ca. 527 nm and a Z-average of 51.52 nm.

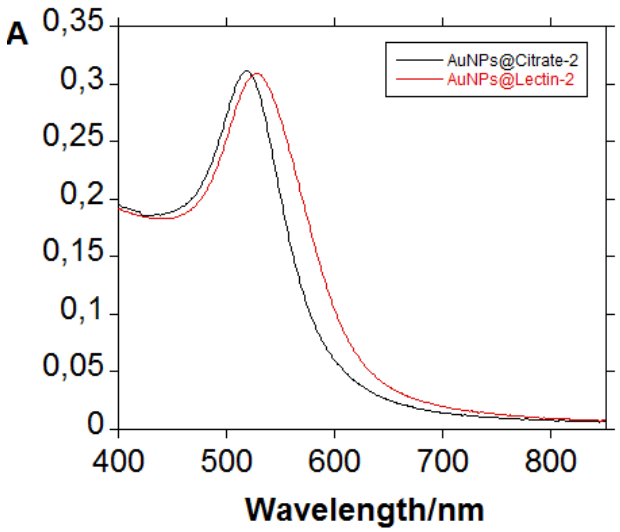


Figure 2.8: UV-Vis spectras of AuNPs+Citrate-2 and AuNPs+EHL-2.

The results obtained for this second batch confirm the reproducibility of the Au+Citrate NPs functionalisation and are summarised in Table 2.2.

Batch	**Batch 1**		**Batch 2**	
Nanoparticles	**AuNPs+Citrate**	**AuNPs+EHL**	**AuNPs+Citrate-2**	**AuNPs+EHL-2**
SPR Band (nm)	519	528	519	527
Z-Average (nm)	19.28	54.41	21.56	51.52
Zeta Potential (mv/cm)	-43.6	-27.8	-39.8	-24.9

Table 2.2: UV-Vis, DLS and Zeta Potentials obtained for the two batches synthesised.

To demonstrate that the protein is present at the surface of the NPs, a study was carried out by adding 200 µL of NaCl 2M into the colloidal system. When NaCl is added to the AuNPs+Citrate-2 system, as a result of reduced separation between particles due to the blockage of the citrate surface charge (and therefor repulsive force) by the sodium anions, aggregation occurs. This also results in an increased zeta potential of the NPs which were previously stabilised by the electrostatic charge blockage of the citrate molecules. This then showed a redshift of the SPR band and a base line increase (Figure 2.9). With the addition of NaCl into the AuNPs+Lectin-2 system, no destabilisation was observed, which confirmed that the EHL was present on the NPs surface.

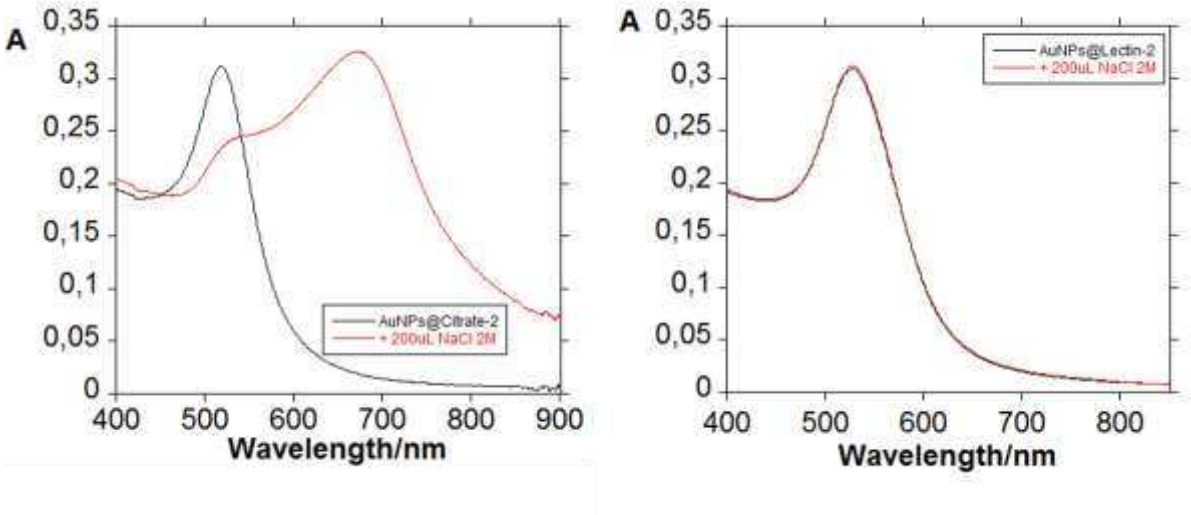


Figure 2.9: UV-Vis studies of the addition of 200 uL NaCl 2M in the NPs obtained Au+Citrate NP-2 and Au+EHL NP-2

2.4 Discussion

The method published in McConnell *et al.* (2015) and reported in detail herein has effectively achieved the objective of the work, which was to extract and purify EHL to a quality and quantity to allow further experimental work to be carried out as required.

Towards the conclusion of this research and following crystallisation and all biological assays, the introduction of a protease inhibitor cocktail (Sigma Aldrich) was added in the initial homogenization step of some of the final extractions in an effort to further reduce any heterogeneity of the sample prior to affinity chromatography. Overall this later addition of a protease inhibitor cocktail in the natural crude extract appeared to have a negative impact on downstream studies, with vastly reduced agglutination activity apparent. This potentially indicated that a proteolytic event contributes to the functional activity of the protein such as that seen in the post translation processing of Ricin (Peumans and Van Damme, 2010) but further investigation is required to establish the exact nature of the change in activity. Further support for this hypothesis was that the ammonium sulfate precipitation (ASP) process also appeared to be altered. Lectin activity was retained in the 80% ASP fraction rather than the 40% albeit at minimal levels. This could possibly indicate that with highly variable temperatures in the laboratory as well, the protein may be salted out on the boundary between 40 and 50%. Further investigation of this observation will aim to identify the precise percentage saturation that applies to a protease inhibited sample. The success of the nanoparticle conjugation by the BioScope group opens a new opportunity for EHL research, both in terms of the structural and functional impact on gold nanoparticles on the native protein, and for the potential of a bioconjugate with biomarker specificity.

Chapter 3: Studies investigating the genetic and amino acid sequence of EHL

De novo sequencing was carried out at the Proteomics laboratory at the Bioscience Technology Facility, Department of Biology, University of York on a commercial basis.

3.1 Introduction

Protein purification using a natural crude extract has, as described in Chapter Two, additional challenges which can hinder progress with structural work. The availability of plant tissue and natural variation in the condition and growth stage of the plant can affect the levels of protein expression which can impact on consistent output and experimental control. The purification protocol requires a relatively long period of approximately 7 -10 days in total. Ultimately, the objective would be to design a construct for a recombinantly expressed protein. The underlying gene sequence is required in order to design such an expression system. One major benefit of recombinant expression would be the possibility to add a histidine tag to the protein for ease of purification. A His-tag is achieved by the insertion of a DNA sequence which codes for a histidine rich repeat of between six and nine residues attached to the N or C terminal of the vector expressed protein. The resultant protein can be purified with accuracy and specificity using affinity chromatography with immobilised metal ions which bind the histidine tag. (Arnau *et al.*, 2006). These columns can be stripped and cleaned and reused many times aiding in an economically improved protocol. A His-tagged recombinant protein would also open the possibility of conducting assays with anti-His-tag antibodies, which can support more in-depth studies and imaging techniques, or alternatively His-tags can be easily removed by a further purification step

(Arnau *et al.*,2006). Ultimately this would aid the speed and accuracy of the extraction and purification protocol.

A second reason for requiring knowledge of the gene sequence of the protein is of course to inform the crystallographic data and complete the model which has been constructed from the electron density maps (see Chapter Four). The diffraction data which have been collected are both complete and of a reasonably high resolution. However an atomic resolution of approximately 1.5 Å (at which point hydrogen atom begin to become distinguishable) is required to assign certain amino acids to the electron density. This is due to the ambiguity in the structures of the side chains i.e. valine/threonine only differ by a methyl/hydroxyl group (Lesk, 2010). Thus, despite the quality of the crystallographic data obtained, in the absence of the peptide sequence or gene sequence, the structure can only be partially assigned. Completion of the structural model would be enabled by the elucidation of the gene sequence, and would serve to inform the evolutionary phylogeny of ribosome inactivating proteins further.

With the resources available to this study, two approaches were taken in order to obtain the sequence. In the first instance, work focused on the DNA of the plant. Based on a small section of previously published fragments of amino acid sequence (Kumar *et al.,* 1993) (See figure 3.1), four degenerate primers were designed. When a sequence is unknown, degenerate primers become a useful mechanism to explore the genomic DNA with. A degenerate primer consists of a set of oligonucleotide sequences which contain a number of possible bases at various points. This is to accommodate the nature of the triplet coding redundancy i.e. more than one combination of nucleotides in a codon may code for one amino acid (Iserte *et al.*, 2013). These primers are designed such that for the target protein,

all possible combinations of nucleotide sequence are contained within the primer species. Primers with high degeneracy will increase the possibility of unspecific annealing of the designed primers as they become less sequence specific and potentially target closely related non target sequence. As a starting point in primer design, a back translation of the amino acid sequence must be undertaken. There are many online tools available to help the researcher with this and related tasks in designing primers which observe the required conditions for optimum amplification of target sequence. In this work, the peptide fragments were back translated using the Colorado State University Molecular toolkit, the results of which are presented in figure 3.1.

Once the primers were designed, genomic DNA was required from the target species (see figure 3.2). Winter Aconite DNA was successfully extracted using the spin column method with a geneJETplant extraction kit (Thermo Scientific) and the sequence probed under various temperature conditions with the designed primers.

Despite the degeneracy in the primers, the Polymerase Chain Reaction did return positive amplification of an appropriately sized DNA band when the correct conditions were found (figure 3.3). A number of steps were taken to improve the definition of the band such as running gradient PCR, and using a commercially available PCR clean up kit to remove primer dimer bands from the gel. Ultimately, PCR products were sent for Sanger sequencing which works by adding nucleotides to a single stranded DNA template using DNA polymerase. The template specific oligonucleotide primer anneals to the template and the chain is extended by the polymerase from the 3’ hydroxyl group of the primer. The deoxynucleotides added to this chain are base-pair matched to the template (Sanger, Nicklen and Coulson 1977). The free nucleotides in the reaction include dideoxynucleotides

which have the hydroxyl group absent, so that elongation of the sequence chain is terminated when a dideoxynucleotide is incorporated. Only one species of dideoxynucleotide is present in each of four nucleotide reaction vessels and strands will be generated of all possible lengths, due to the millions of template strands in the reaction. These randomly terminated strands of all lengths are then effectively separated on the basis of size and ordered on an electropherogram (see figure 3.4) which allows the complementary sequence to be read. A number of factors can affect the success of this process such as quality of DNA supplied. Whilst the initial set of sequencing based on the original primers appeared to work reasonably well, a second set of sequence specific primers based on that output failed. Unfortunately, the sequencing results from this second set of primers was poor, and whilst Basic Local Alignment Tool (Altschul *et al.*, 1990) hits had initially been promising, the second set of sequence returned showed no known similarity to any sequences in the database.

With many techniques available to try to improve sequence signal such as a better primer design based on homology alignment or using the EHL peptide sequence, the decision to move towards sequencing the protein was taken. A purified protein sample was sent for commercial *de novo* sequencing at the University of York. The results obtained were limited to a small number of confidently assigned fragments, but revealed a strong homology to other known type II RIPs.

The work reported here has yielded the crystallographic structure of EHL to 1.6 Å, which combined with the gene sequence, would allow the structure to be fully refined and deposited in the Worldwide Protein Data Bank (PDB). Without the underlying DNA/mRNA sequence, ambiguity in assigning some structural amino acid residues would remain and

would hold the structure solution progress back. As the sole representative so far defined as having RIP activity from the family *Ranunculaceae*, sequencing the genome would provide an additional branch in the phylogenetic tree of these remarkable toxins and further inform the understanding of the evolutionary convergence of the genes involved. Obtaining the gene sequence and whole genome would allow the completion of the project from a structural and molecular perspective.

Fortunately, towards the end of the project, additional funding was secured in order to sequence the plant genome. Genomic DNA was extracted for a second time under conditions applied by the University of Edinburgh commercial sequencing facility (Edinburgh Genomics). At the point of writing this thesis, the quality standards had been achieved and the sequencing of the EHL genome was underway. It is not within the scope of this thesis to present these data, and it will form the basis of future work on this protein. The raw 'reads' of primary sequence which were returned will require considerable processing and bioinformatics to create the contiguous fragments that will ultimately be assembled to present the genome. This will be the source of the gene sequence which will be mined using RIP homologs using the highly conserved regions within the EHL model.

3.2 Materials and Methods

3.2.1 Genomic DNA extraction

Using the GeneJET Plant Genomic DNA Purification Kit (Thermo Scientific) genomic DNA was extracted using the following method: firstly 350 µL of Lysis Buffer A from the kit was added into a 1.5 mL pre sterilised microcentrifuge tube. Previously frozen plant tissue (-80°C) weighing 100 mg was ground using a pre cooled and sterilised micro pestle and centrifuge tube, itself placed in a falcon tube with pre-cooled ethanol (-80°C) as an

alternative to liquid nitrogen. The ground tissue was immediately transferred into a 1.5 mL microcentrifuge tube containing 350 µL of Lysis Buffer A. This was then vortexed for approximately 20 seconds. Subsequently 50 µL of Lysis Buffer B and 20 µL RNase A was added to the tube. For the genome sample sent to Edinburgh Genomics, a further reagent was added to this ste: 1 µl of Riboshredder RNase blend (Epicentre) with an amendment to the published protocol. This was to ensure the level of purity required for genomic sequencing. The sample was then incubated for 10 min at 65°C using a shake cycle on an Eppendorf thermomixer. Once the incubation was complete, 130 µL of Precipitation Solution was added and mixed. The mixture was then incubated for 5 min on ice. The tube was centrifuged for 5 min at 17,572 x *g* and the supernatant collected and transferred to a clean microcentrifuge tube. Next, 400 µL of Plant gDNA Binding Solution and 400 µL of 96% ethanol was added and mixed. Half of the sample (600µL) was added to the spin column and centrifuged for 1 min at 7084 x *g*. The flow-through solution was discarded and the remaining mixture applied onto the same column which was also centrifuged for 1 min at 7084 x *g*. 500 µL of Wash Buffer I containing guanidinium hydrochloride was added to the column and centrifuged for 1 min at 11,068x *g*. The flow- through was again discarded and the column placed back into the collection tube. Wash Buffer II (500 µL) was then added to the column and subject to a centrifuge cycle of 3 min at 17,572 x *g*. The collection tube was emptied and the purification column placed back into the tube and spun for a further 1 min at 17,572 x *g*. The column was transferred to a sterile 1.5 mL microcentrifuge tube and the genomic DNA eluted by adding 100 µL of elution buffer (10 mM Tris HCl, pH 9.0, and 0.1 mM EDTA) to the centre of the column membrane with a 5 min room temperature incubation. This was then centrifuged for 1 min at 11,068 x *g* and a second elution step

using 100 µL Elution Buffer performed. The purified DNA was stored at -20°C for PCR and genomic sequencing.

3.2.3 Polymerase Chain reaction

Many differing combinations of parameters were used to optimize the PCR reaction, including changing the amount of primers added, the concentration of stock DNA and the annealing temperature. The method used 2 x concentrated ReadyMix Taq PCR Reaction Mix with MgCl2 (SigmaAldrich), which contains. 20 mM Tris-HCl, pH 8.3, 100 mM KCl, 3 mM MgCl2, 0.002 % gelatin, 0.4 mM dNTP mix (dATP, dCTP, dGTP, TTP), stabilizers, and Taq DNA Polymerase (10 µl in a total reaction volume of 20 µl). To this was added 2 µl each of forward and reverse primers, template DNA, and 4 µl of sterile dH_2O.

3.2.4 *De novo* Sequencing Analysis

As performed and reported by The University of York protein identification service.

In-gel tryptic digestion was performed after reduction of an EHL SDS PAGE gel band with Dithioerythritol (DTE) an epimer of DTT, and S-carbamidomethylation with iodoacetamide. Gel pieces were washed two times with 50% (v:v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, then once with acetonitrile, and dried in a vacuum concentrator for 20 min. Sequencing-grade, modified porcine trypsin (Promega) was dissolved in the 50 mM acetic acid supplied by the manufacturer, then diluted 5-fold with 25 mM ammonium bicarbonate to give a final trypsin concentration of 0.02 µg/µL. Gel pieces were rehydrated by adding 10 µL of trypsin solution, and after 10 min enough 25 mM ammonium bicarbonate solution was added to cover the gel pieces. Digests were incubated overnight at 37°C.

3.2.2 Primer design

```
1       MetValSerGluAlaAlaArgIleArgTyrIleGluHisLeuValArgArgSerThrArg

        ATGGTTTCTGAAGCTGCTCGTATTCGTTATATTGAACATCTTGTTCGTCGTTCTACTCGT

             CAGC  G  C  CA C  CA C  C  C  G  CT C  CA CA CAGC  CA C

             A  A     A  A  A  A  A     A        A  A  A  A  A  A  A

             G  G     G  G  G     G              G  G  G  G  G  G  G


2       IleIleGlyArgAspGluLeuValAspValAspGlyGlnAsnAsnAsnGly

        ATTATTGGTCGTGATGAACTTGTTGATGTTGATGGTCAAAATAATAATGGT

          C  C  CA C  C  GT C  C  C  C  C  C  G  C  C  C  C

          A  A  A  A        A  A     A     A              A

                G  G        G  G     G     G              G


3       MetIleTyrAspThrAspSerAlaValProAspAlaThrVal

        ATGATTTATGATACTGATTCTGCTGTTCCTGATGCTACTGTT

             C  C  C  C  CAGC  C  C  C  C  C  C  C

             A        A     A  A  A  A     A  A  A

                      G     G  G  G  G     G  G  G
```

Figure 3.1 Primer design. Amino acid sequence entered as fragments as seen in Kumar *et al.* (1993), back translated using:
http://www.vivo.colostate.edu/molkit/rtranslate/index.html

A 1 µL aliquot of each peptide mixture was applied to a ground steel MALDI target plate, followed immediately by an equal volume of a freshly-prepared 5 mg/mL solution of 4-hydroxy-α-cyano-cinnamic acid (Sigma) in 50% aqueous (v:v) acetonitrile containing 0.1%, trifluoroacetic acid (v:v).

Positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III in reflectron mode, equipped with a Nd:YAG smart beam laser. MS spectra were acquired over a range of 800-5000 *m/z*. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg^1-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu^1-Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929.). Monoisotopic masses were obtained using a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) and a S/N threshold of 2.

For each spot the ten strongest precursors, with a S/N greater than 30, were selected for MS/MS fragmentation. Fragmentation was performed in LIFT mode without the introduction of a collision gas. The default calibration was used for MS/MS spectra, which were baseline-subtracted and smoothed (Savitsky-Golay, width 0.15 m/z, cycles 4); monoisotopic peak detection used a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a minimum S/N of 6. Bruker flexAnalysis software (version 3.3) was used to perform spectral processing and peak list generation.

De novo sequencing of tandem mass spectra was performed by hand, with *a-, b-, b^0-, y-, y^0- and y**-ions considered as possible fragment ions. *De novo* derived peptides sequences were matched to homologous protein sequences using the on-line MS-BLAST service provided by Washington University,(St louis, Missouri,USA) (Shevchenko *et al.*, 2001).

Peptide fragmentation and analysis were carried out using MALDI-TOF/ MS/MS and *de novo* sequencing. The peptide mixture produced was analysed by MALDI-TOF with the ten most intense peptide ions selected for MS/MS fragmentation. The resulting fragmentation spectra were searched against the NCBI database restricted to *viridiplantae* using the Mascot search program but no matches were produced.

Manual inspection of the fragmentation spectra showed that they were of sufficient quality to expect a match from database searching, if the corresponding protein was present in the database. Database searches did not produce any matches.

In an attempt to identify the protein, *de novo* sequencing was performed on the fragmentation spectra generated. *De novo* sequencing was undertaken by hand at York University with annotated spectra produced subsequently in BioTools or DataAnalysis software packages.

The resulting suggested *de novo* sequences were submitted to MS-BLAST search using the on-line service provided by Washington University. The MS-BLAST search allows for more divergence between the theorised sequences and those held in the database, by allowing for: amino acid substitutions, deletions or insertions. Where a match is made across multiple *de novo* sequenced peptides, there is confidence that the protein matched has good homology with the protein sequenced (Proteomics Laboratory at the Bioscience Technology Facility, Department of Biology, University of York).

3.3 Results

3.3.1 Genomic DNA Extraction

DNA extraction carried out using a geneJET plant genomic DNA extraction kit (Thermofisher) as described, the results of which can be seen on the 1.5% agarose gels below in figure 3.2. Qubit fluorimetry (Invitrogen) showed a concentration range of 28 -32 µg/ml. Sample A was used for PCR work and Sample B was sent to the Edinburgh Genomics centre for sequencing towards the conclusion of this work.

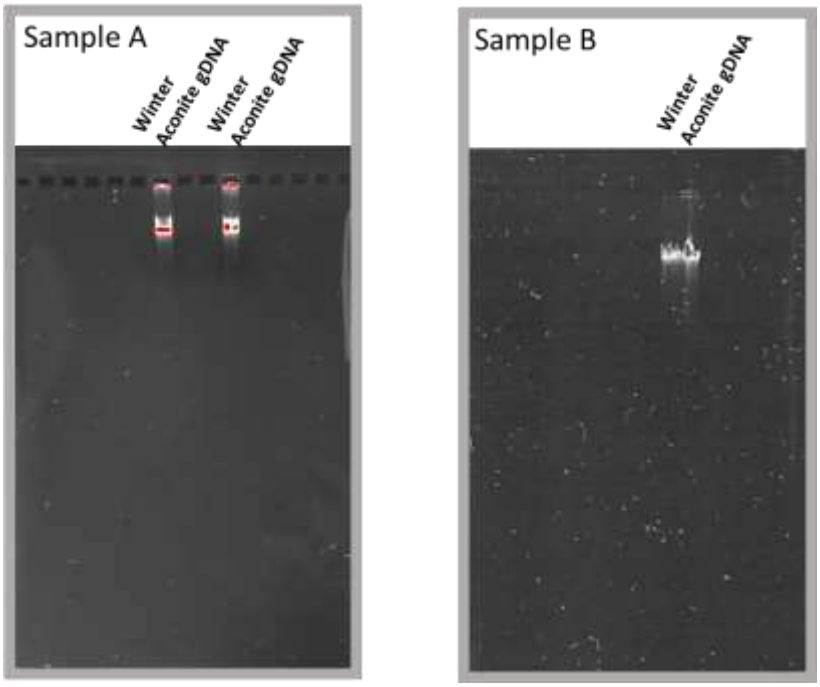


Figure 3.2 Genomic DNA extractions for PCR and genome sequencing.

3.3.2 Polymerase Chain Reaction

A large number of reactions were carried out under varying combinations of temperature and concentration of primers. A gradient PCR produced the sharpest bands at 52°C as seen in figure 3.3. This product was sent for sequencing.

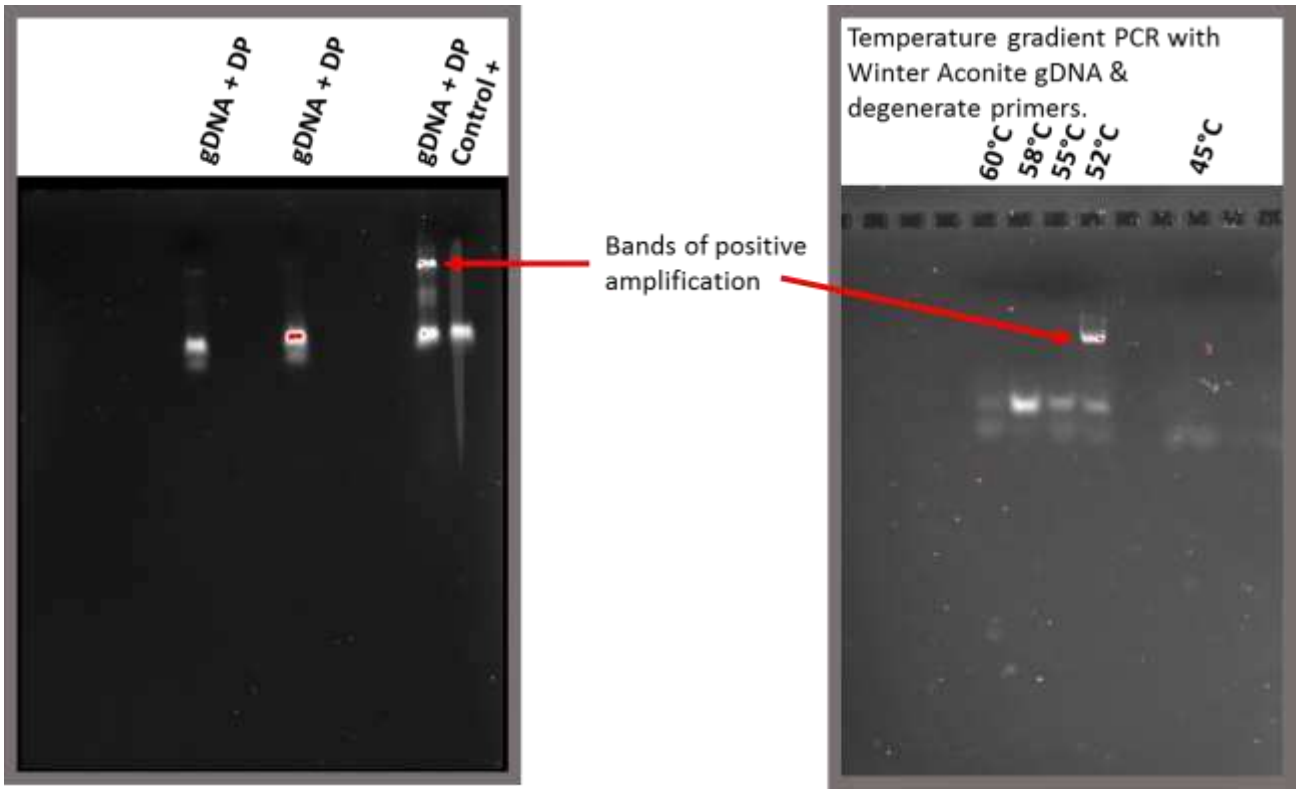


Figure 3.3 PCR amplification of region of Winter Aconite gDNA from primers and with a 15°C temperature gradient between 45 and 60°C.

3.3.3 Sanger sequencing

The results of Sanger sequencing were limited and in the interest of pursuing other experimental studies, this approach was postponed early in the project with a view to achieving funding for entire genome sequencing of EHL. The electropherogram shown in figure 3.4 demonstrates the low background noise signal from the successful sequencing output.

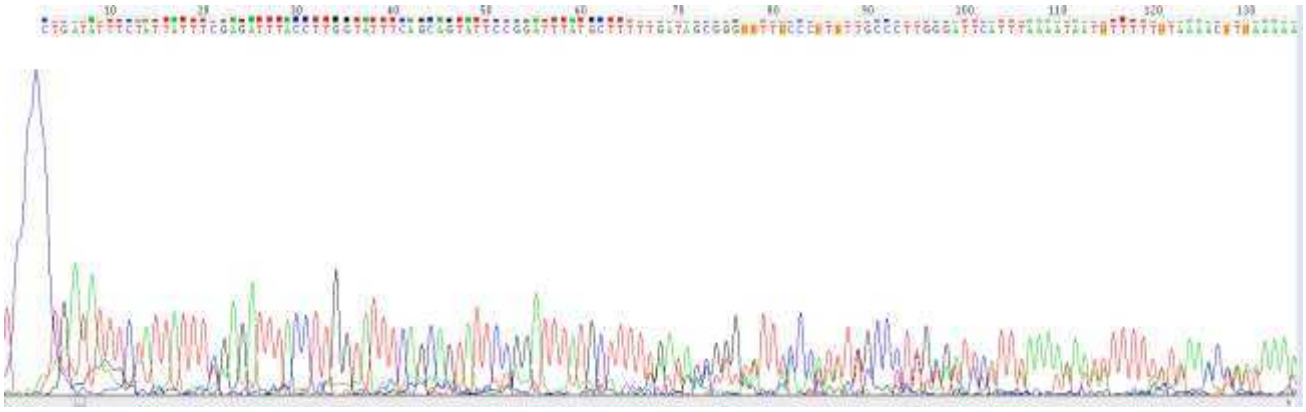


Figure 3.4 Electropherogram of the PCR product sequenced by the Sanger Method (Sanger, Nicklen and Coulson, 1977) at Source Biosystems.

3.3.4 *De novo* sequencing

De novo sequence analysis of fragmentation spectra was carried out and suggested two confidently assigned peptide fragment sequences as shown in Figure 3.5 (a and b), with the following tags, QQWA[L/I]YSDST[L/I]R and NWNNNGNP[L/I]Q[L/I]WQCTQQQNQR. Peptide fragment homology searches produced matches to various Type II RIPs all within the ricin-b domain (carbohydrate binding) regions.

This result confirms the status of EHL as a Type II RIP, as has been reported in previously published sequence data (Kumar *et al.*, 1993). Sequence tags QQWA[L/I]YSDST[L/I]R and NWNNNGNP[L/I]Q[L/I]WQCTQQQNQR were matched to two regions in Nigrin-b (SNA-V; *Sambucas nigra* agglutinin-V) UniProtKB P33183.2 correlating to residues 470-481 and 325-345 respectively using the FASTS search facility against the SwissProt database. Close homology to other Type II RIPS was also matched within the ricin b domain of Abrin-a (P11140) and Ricin (P02879). PDB structures were searched using the SSearch algorithm

and also showed matches to SNA-II (3C9Z) (*Sambucas nigra* agglutinin II), Abrin-a (1ABRB) and Ricin (1RZOB) in the ricin b domain as well as ML1 from *Viscum album* (1QNKB).

The MS-BLAST search for the suggested *de novo* sequences gave strong matches to plant ribosome inactivating/ricin-agglutinin proteins. The top three matches of which were:

- gi|7141131|gb|AAF37219.1|AF213984_1 ribosome inactivating protein RIPt [Polygonatum multiflorum]/
- gi|255585867|ref|XP_002533608.1| ricin-agglutinin family protein [Ricinus communis]//:gi|223526509|gb|EEF28777.1| ricin-agglutinin family protein [Ricinus communis]
- gi|255582840|ref|XP_002532193.1| ricin-agglutinin family protein [Ricinus communis]//:gi|223528125|gb|EEF30196.1| ricin-agglutinin family protein [Ricinus communis]/

These matches also help to give further evidence to the suggested *de novo* sequences.

The results of these were consistent with the in-house homology search results conducted on confidently assigned sequences using the University of Virginia UVa FASTA Server. The FASTS and SSearch algorithms were used for homology searching against the SwissProt (NCBI) and PDB databases. The results of crystallographic studies presented in Chapter Four indicate an amino acid sequence which is in agreement with the peptide fragments published by Kumar *et al.,* (1993) calculated for 71/72 residues. This confirmation is based on the electron density maps calculated from the diffraction data. The *de novo* sequencing fragments have been docked within the model sequence also where possible, but only utilizing the confidently assigned peptides to do so and even so with a truncation of ambiguous regions where the ion series is not corroborated.

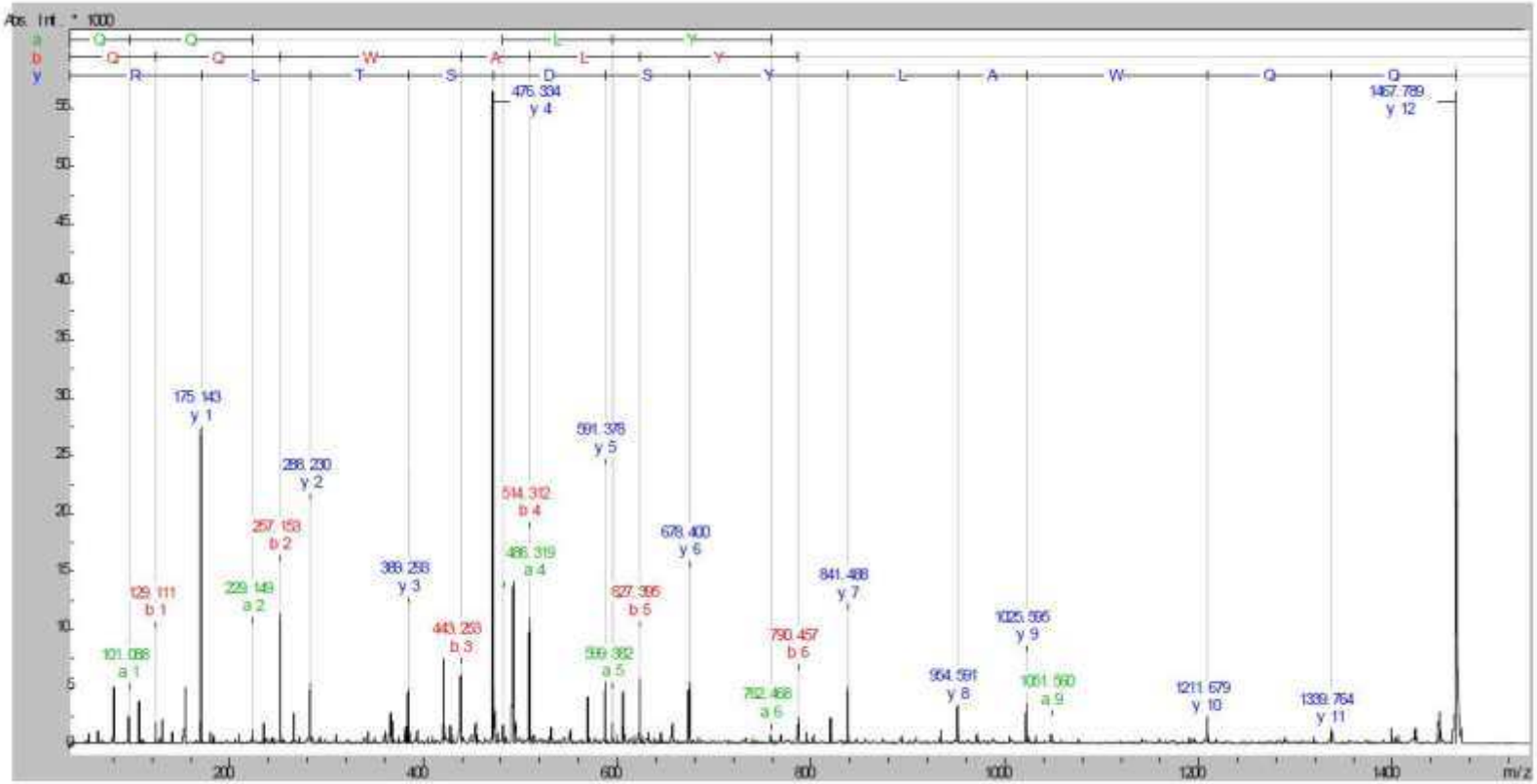


Figure 3.5a *De novo* peptide sequencing based on the MALDI-TOF MS/MS fragmentation spectra of EHL in-gel tryptic digestion after reduction with DTE and S-carbamidomethylation with iodocetamide. A) **QQWA(L/I)YSDST(L/I)R –** The most confidently assigned peptide in the batch, with a good level of overlapping *y*- and *b*-ion series.

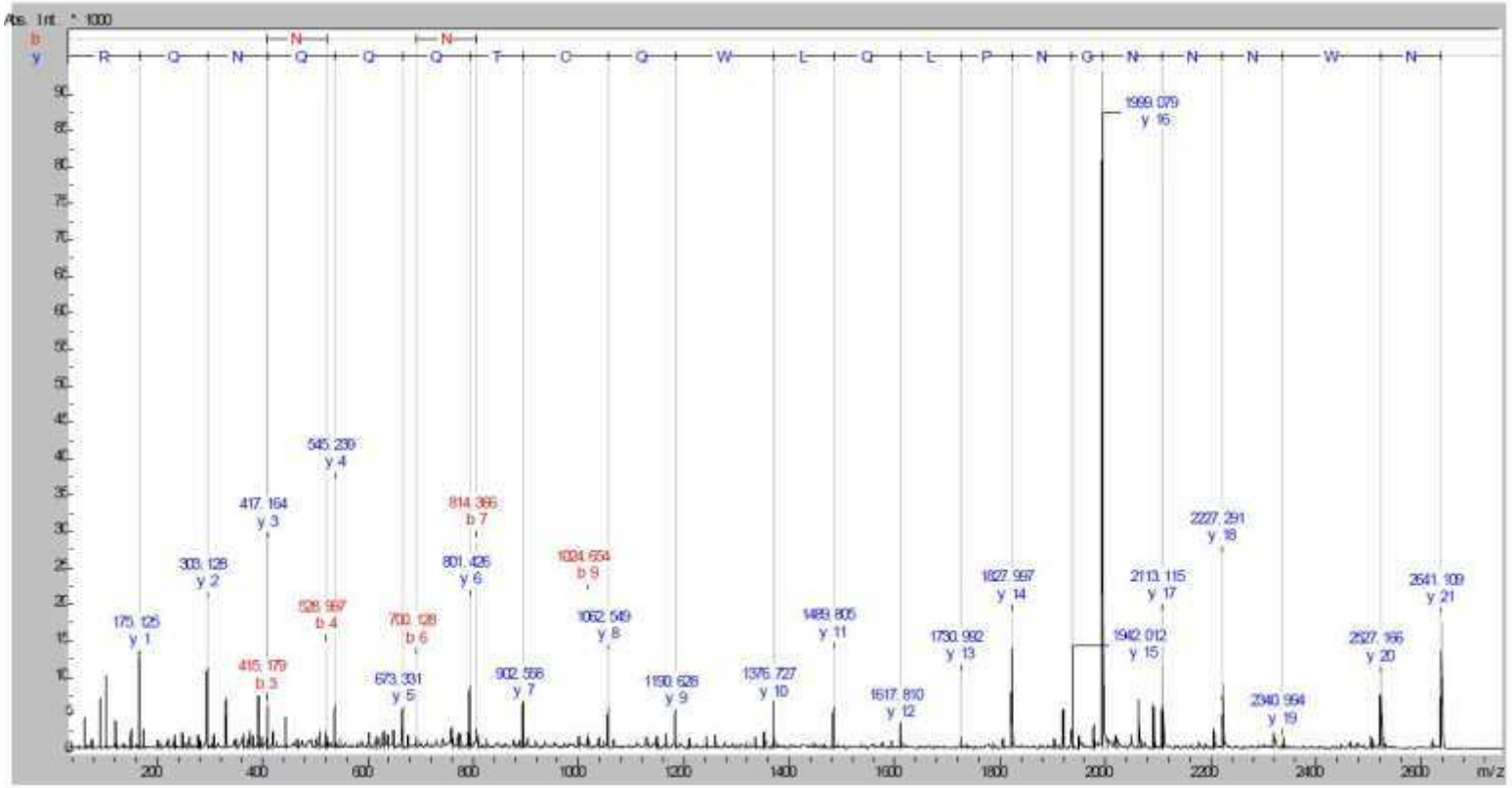
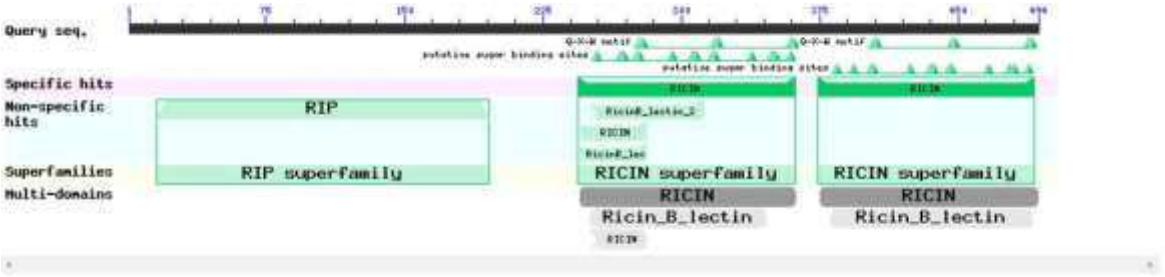


Figure 3.5b *De novo* peptide sequencing based on the MALDI-TOF MS/MS fragmentation spectra of EHL in-gel tryptic digestion after reduction with DTE and S-carbamidomethylation with iodocetamide. B) **NWNNNGNP[L/I]Q[L/I]WQCTQQQNQR** – A strong *y*-ion series is observed throughout the peptide

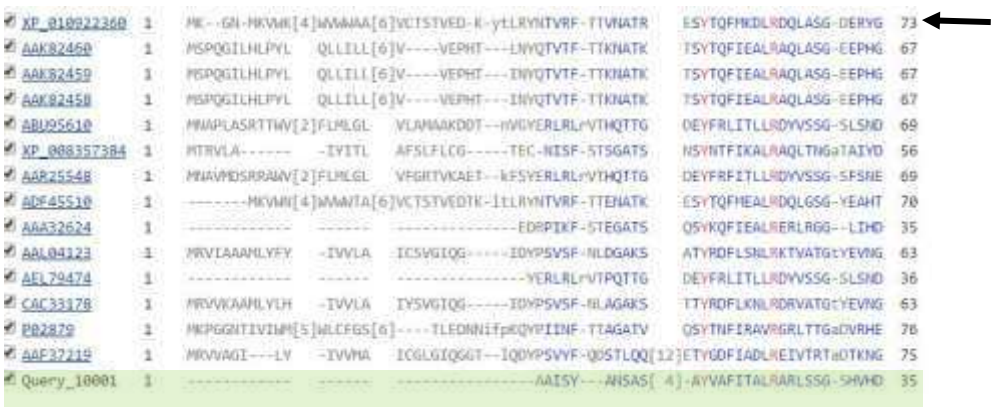
These are identified as: A-Chain - LSSGSHVHDFLPQR (note the leucine and glutamine are inverted and the proline interjects between the residues in the model rather than precedes them as in the experimental ion series) and B-Chain - QQWALYSDSTIR; NNNGNPIQLWQCTQQQNQR. Combined with the Kumar fragments and the electron density maps, the model can be viewed with particular confidence in 115 residues that have been independently corroborated using a variety of methods. The resulting amino acid sequence was used to conduct a BLASTp search using the NCBI server and database (Altschul *et al.*, 1990). The results of which (Figure 3.6) confirm the identity of EHL as an evolutionary divergent Type II RIP. A sequence alignment confirmed the conservation of motifs present in other RIPs such as the $(QxW)_3$ regions highlighted in Figure 3.7.



A-Chain
AAISYANSASPNVVAYVAFITALRARLSSGSHVHDFPQLRAPSALPSACRFILVDLENGGGHTITVPIDVFNAA
VGGYLVGAAFLYFSCGPPAAAAIFPAAVSRSNGFGGGYGALGARAAESLGHAALLSAILALFYSYSQRTSFLTI
IQMVSEAARIRYIEHLVRRSTGASGDFLPDPSAACLELCWDPLSTQIQACGGSGVFTRPTWCAACCYAAVAI
LAIGEVNSGAALALLLFG

B-Chain
MTTTRVIIGRDELCVDVRDGENNNGNPIQLWQCTQQQNQRWTFAIDGTIRSNGKCLTTYGYAAGAYIMIY
DCDSAVPDAVVWACSGNGTISNPASGLALTASSCSPGTTLTCQTLINASREAWTVGQYSQPATTSYISGFRA
MCMEANDLNTLTWDMSCASGQECQQWALYSDSTIRVFCNPSLATTSGGASASLVIGIASGSGWGLQRW
AFSANGVILLPDACLTMDVRGSNVSLRQIILYEPTGNPNQQWLAYS

Figure 3.6 EHL BLASTp results and sequence used**.** EHL is confirmed as a Type II RIP with conserved domains, $(QxW)_3$ motif and 14 putative sugar binding sites identified (Marchler-Bauer *et al.*, 2014).


Sequence residue number

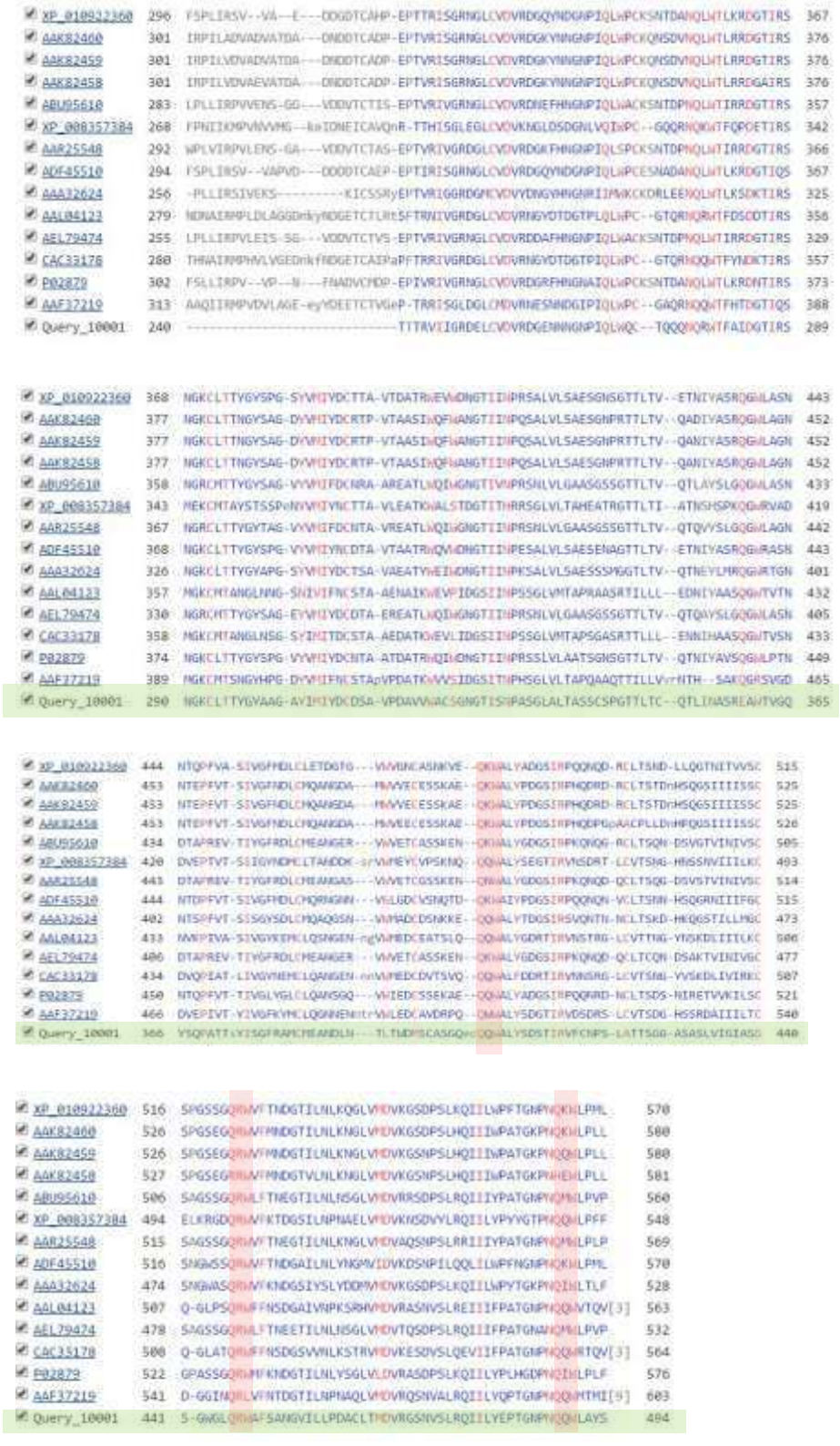


Figure 3.7 BLASTp multiple sequence alignment. EHL is identified as Query_10001 shaded in green. Areas in pink boxes show the conserved $(QxW)_3$ motif.

Key to sequence used in BLASTp alignment figure 3.7:

- XP_010922360.1 Elaeis guineensis
- AAK82460.1 type 2 ribosome-inactivating protein cinnamomin III (*Cinnamomum camphora*)
- AAK82459.1 type 2 ribosome-inactivating protein cinnamomin II (*Cinnamomum camphora*)
- AAK82458.1 type 2 ribosome-inactivating protein cinnamomin I (*Cinnamomum camphora*)
- ABU95610.1 articulatin precursor [*Viscum articulatum*]
- XP_008357384.1 nigrin b-like [*Malus domestica*]
- AAR25548.1 lectin precursor [*Viscum album*]
- ADF45510.1 type 2 ribosome-inactivating protein precursor [*Camellia sinensis*]
- AAA32624.1 abrin-a [*Abrus precatorius*]
- AAL04123.1 ribosome-inactivating protein [*Sambucus nigra*]
- AEL79474.1 ribosome-inactivating protein precursor [*Viscum ovalifolium*]
- CAC33178.1 ribosome-inactivating protein precursor [*Sambucus ebulus*]
- P02879.1 Ricin (*Ricinis communis*)
- AAF37219.1 ribosome inactivating protein RIPt [*Polygonatum multiflorum*]
- lcl|Query_10001 EHL (*Eranthis hyemalis*)

3.4 Discussion

The aim of this part of the work was to find the gene sequence of EHL. Whilst progress was limited initially, the combined result of all the techniques used has resulted in a confident and plausible protein sequence, which can be used for bioinformatics analysis.

A comparison and analysis of the sequence of EHL shows that the B chain is most conserved amongst homologous RIPs with sequence identity of 53% with Mistletoe Lectin (ML1) and 48% similarity with Ricin. Using the sequence as it currently stands for alignment. The A chain appears much more divergent with only 35% similarity with Mistletoe and 38% with Ricin. These comparisons identify EHL as belonging to an outlying group in the RIP

phylogeny which have been identified in approximately 100 plant species (Van Damme *et al.*, 2001) (Figure 3.8).

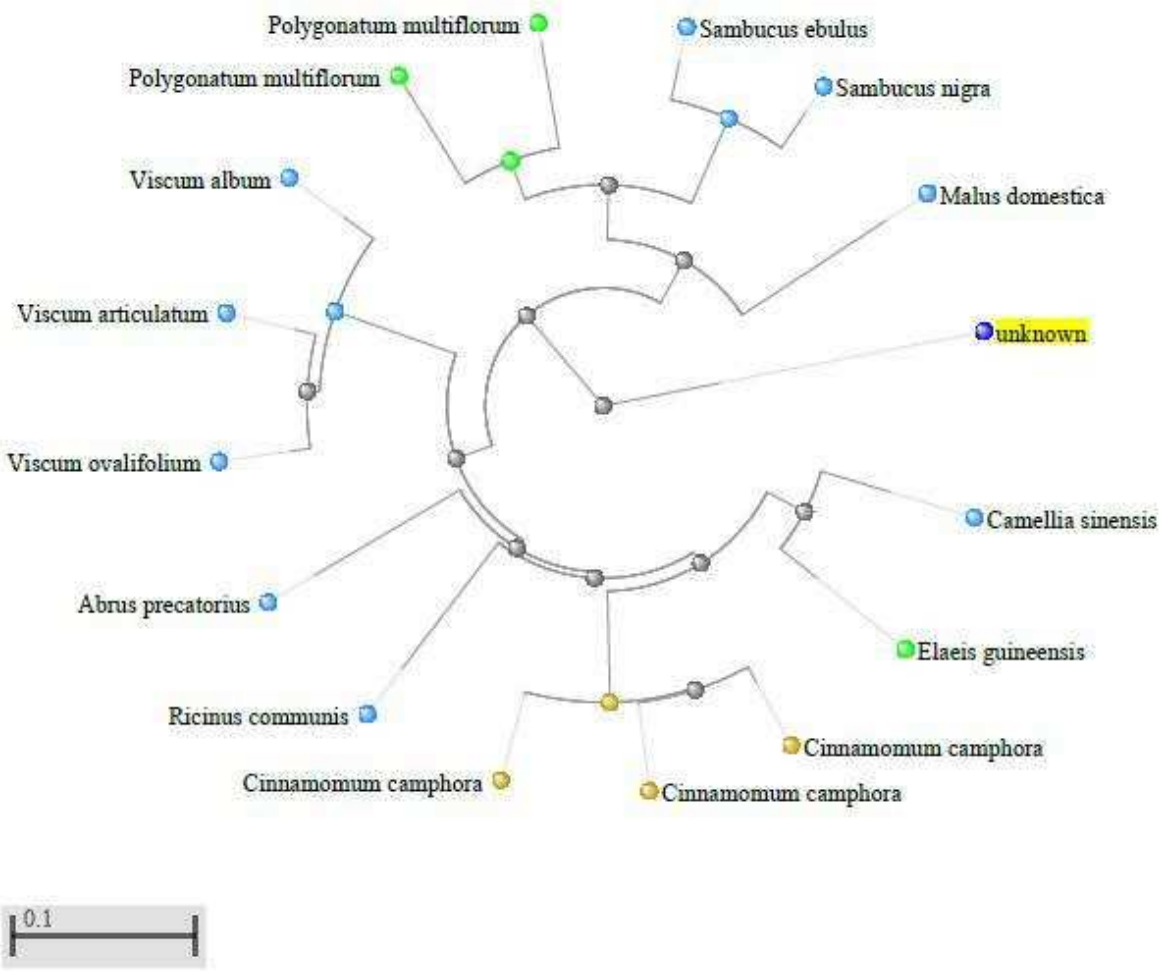


Figure 3.8 COBALT Neighbour Joining phylogenetic tree from Blastp alignment of the whole EHL sequence made using COBALT Software by Papadopoulos and Agarwala (2007).

This small sequence alignment analysis already demonstrates the absence of a region of sequence at the N-terminal of EHL which is present in other RIPS such as ricin, corresponding to a 35 residue signal peptide (Van Damme *et al.*, 2001). This highlights once

again the requirement for the whole gene sequence so that any signal peptide and linker regions (which may be apparent in unbuilt electron density) can be identified.

Once the structure is completely refined and the genomic sequence available, this phylogeny will of course be subject to further analysis, which may cause this apparent evolutionary divergence to be reassessed.

Chapter 4: Crystallographic studies of EHL

Diffraction data collection was carried out at Diamond Light Source on Beamline 104-1 (Beamtime allocation number MX8997)

Crystal data were collected by Dr. Colin Levy from the Manchester Institute of Biotechnology for which the author is grateful.

Strategies for density modification and structure building were supported by senior academic staff including Dr. Edward Lowe at The Oxford University Biochemistry Department.

Support from Dr Clare Naylor at Birkbeck College in establishing the data processing strategy is gratefully acknowledged.

4.1 **Introduction**

> "X-Ray crystallography has been an essential technique for structural analysis over the last one hundred years. It has allowed us to view the world on a scale that was previously impossible, giving 3 dimensional information and insight in the structural mechanism and functions of important materials and biological molecules" (Brooks-Bartlett and Garman, 2015).

Key to understanding the function or mechanism of action for any protein is finding the three dimensional folded structure. Macromolecular X-ray crystallography provides a means to elucidate the underlying structure and help to define the function and biological activity of a protein and other molecules. The interpretation of the diffraction of X-rays by crystals (X-ray crystallography) is a technique that has been used for over one hundred

years and has provided structures of biologically fundamental molecules such as Insulin (Adams *et al.*, 1969), Vitamin B12 (Hodgkin *et al.*, 1956), Penicillin (Crowfoot *et al.*, 1949), and the double helix structure of DNA (Watson and Crick, 1953). With more recent advances in reducing the rate of radiation damage from the intense synchrotron X-ray beams used such as cyro-cooling (Garman, 2003), increasingly more complex structure determinations have been made possible, such as that of the ribosome for which a Nobel prize was awarded to Venekatraman Ramakrishnan, Thomas Steitz and Ada Yonath in 2009 (Brooks-Bartlett and Garman, 2015).

Unlike photons with energy in the visible light spectrum which reflect from the external outline of an object, X-rays penetrate through the whole of a molecular structure providing detail on the internal three dimensional structure. The electrons of the atoms are excited by the X-rays which leads to the emission of X-ray photons. When these are of the same wavelength as the incoming X-rays, elastic (cohesive) scattering results. The scattering from all atoms will interfere to produce a diffraction pattern according to Bragg's law (equation 4.1) which is captured on a detector. The scattering can be expressed as structure factors which contain both the amplitude and the phase of the diffracted photon. The mathematical principle of the Fourier transform allows a conversion of these structure factors into a density map of the positions of electrons in the structure (Blow, 2002).

Within the crystal lattice is contained the smallest translatable repeating unit, or unit cell; the smallest volume of space that can be used to fully describe the crystal. By convention, the unit cell has an origin at the bottom left hand corner, from which distances a, b and c are measured in Angstroms. The values of a,b and c define the volume of the unit cell. Alpha (α), beta (β) and gamma (γ), define the angles of the unit cell edges between b-c,a-c

and a-b respectively (Blow, 2002). The origin of the unit cell defines the parameters from which all measurement indices (Miller indices) (*h,k,l*) of the reflections are generated. The lattice planes or Bragg planes are planes within the crystal from which X-rays reflect such that Bragg's law is obeyed. The Bragg planes generate a set of indices (*h,k,l*) for each reflection based on the spacing and direction of the planes such that a set of coordinates is generated for each diffraction spot on the detector; critically all atoms in the unit cell contribute to all intensities in the *h,k,l* reflection. The intensities of the diffraction spots are proportional to the square of the structure factor's amplitude. The structure factor contains information on the number of electrons in each atom (the amplitude) and its position (the phase). It is not possible to measure the phase directly from the structure factors and a variety of methods can be applied to do so indirectly (Blow, 2002). In this work, initially phases were extracted from a homologous protein using molecular replacement software PHASERMR (McCoy *et al*., 2007). This exploits the difference between the experimentally observed structure factors and the calculated structure factors from a homologous model. With phases derived from the molecular replacement model, the Fourier transform then allows an electron density map to be assembled. This forms the basis for building and refining the structure.

W.L Bragg proposed the idea in 1913 (Bragg and Bragg, 1913) that there can be considered to be planes of atoms within the lattice, and that some of these planes will produce strong and weak reflections if the angle of incidence of the X-ray is equal to the angle of reflection. Where the difference between reflections is a whole number of wavelengths, Bragg's law will be observed.

In order for Bragg's law to be obeyed, i.e. the incident beam for a given order of diffraction to be reflected from a plane and in phase with the reflected beam in successive planes in the same h,k,l family which contribute to the intensity of a reflection *h,k,l,* the following condition must be fulfilled (Hammond and Hammond, 2009);

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

Equation 4.1 Braggs Law (Bragg and Bragg, 1913)

That is to say, the distance between the incident and reflected beam of two parallel photon streams from a monochromatic source is a whole number of wavelengths apart (*n* is an integer) (Blow, 2002). Where this condition is not fulfilled, the resultant reflected beam will not be in phase with the first beam and so will result in destructive interference and no signal.

A characteristic diffraction pattern from the EHL crystal can be seen in Figure 4.1 with the rings showing the resolution and the diffuse reflections in the region between 3 and 4 Å indicative of solvent scattering. The spots observed are the reflections from the Bragg planes in the crystal and are termed the reciprocal lattice. The resolution of the data relates to the spacing between successive planes in a family of *hkl* reflections i.e *d*=1.6 Å ($n\lambda = 2d_{hkl}$ sinθ) (Blow, 2002).

The crystallisation process begins with purifying the protein to a high and homogeneous level and then finding the conditions which allow the protein to crystallise. This is done through screening a theoretically infinite number of conditions. In practice, with the introduction of commercial screens which allow a rational and logical sequence of

screening, this has resulted in, on average, screening 300-500 conditions to provide leads in globular proteins likely to crystallise (Chayen and Saridakis, 2008).

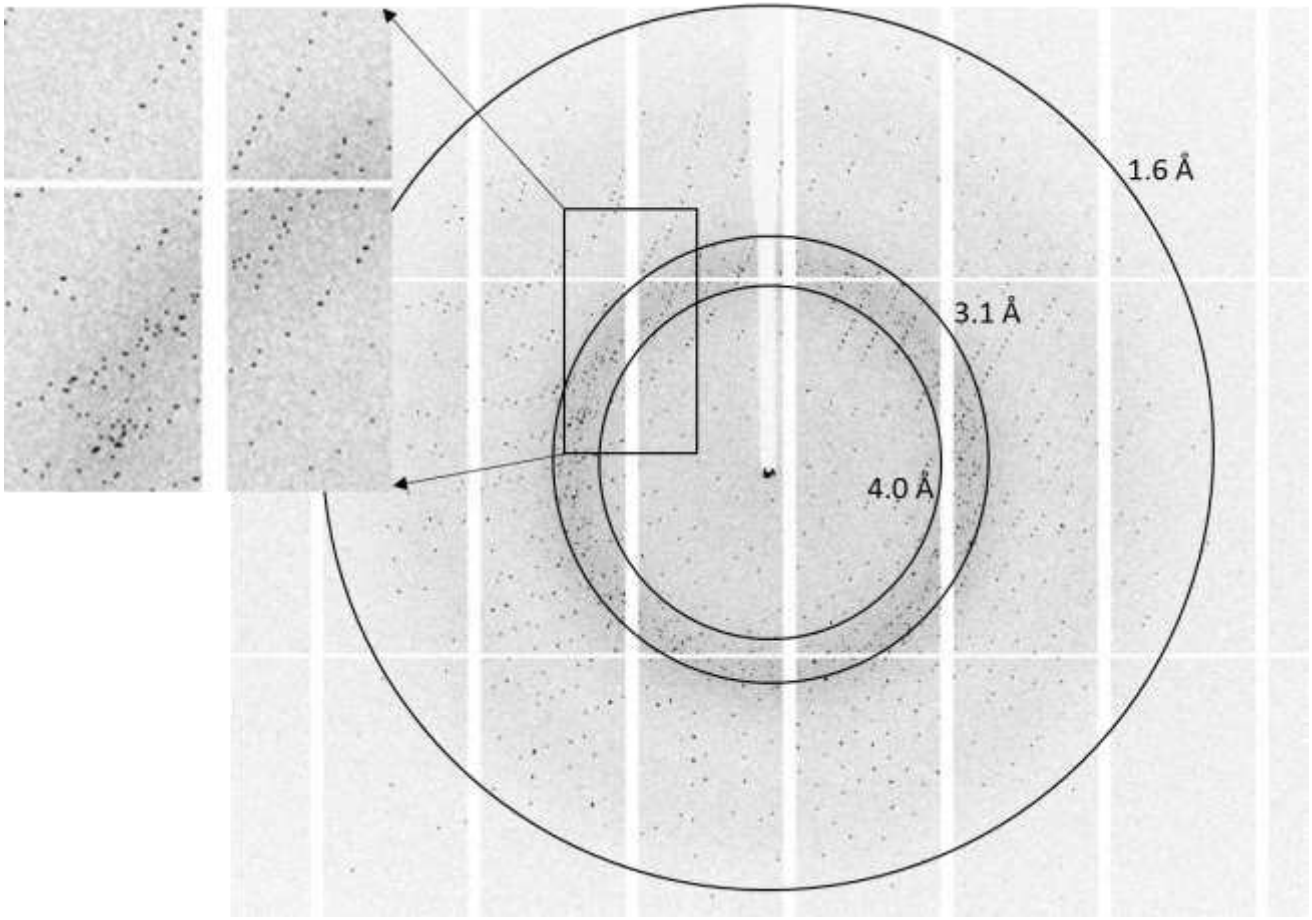


Figure 4.1 Diffraction pattern of EHL at 1.6 Å (outer resolution limit marked by ring) and zoomed in image of reflections. Inner resolution rings marked to show the shadow region between 3.1 and 4.0 Å indicative of disordered solvent from mother liquor and solvent channels within protein and crystal.

In order for protein to crystallise, it must undergo a phase transition, which requires in the first instance the protein to reach a point of supersaturation in order to spontaneously nucleate. If the solution is too highly supersaturated, then the protein will form a precipitate, if it is under saturated it will remain soluble and will not crystallise. Once nucleation has occurred, the saturation will reduce slightly and allow the protein to enter what is termed the metastable zone. Here crystal growth may occur, and the solution is in equilibrium with the reservoir of the crystallisation condition (Chayen and Saridakis 2008). The crystallographic method of vapour diffusion relies on the initially under saturated protein passing through the metastable zone into the nucleation zone and returning without forming a precipitate to the metastable zone. Commercial screens provide a systematic and economical method for sampling the crystallographic space. Promising leads or 'hits' can then be optimized by varying the parameters of the hit, such as adding a pH gradient vs precipitant or salt concentration gradient of the additive.

In the majority of modern structures, the underlying gene sequence and resultant amino acid sequence of the protein would be known. The protein would likely have been cloned, tagged and overexpressed in an expression system such as *E.coli* or yeast. After purification, crystallisation and diffraction experiments, a homologous structure can be used with molecular replacement programmes to extract phase information and is then used for structure building and refinement using the crystallographic software suites available, this work uses CCP4 (Collaborative Computational Project, Number 4) (Winn *et al.*, 2011). Whilst as reported in Chapter Three, some attempts were made to achieve this using molecular biology techniques such as degenerate PCR and *de novo* protein sequencing, positive results were limited. Research funding was granted at the end of this

project to sequence the entire genome of the Winter Aconite plant, and this will ultimately at a future date confirm the amino acid sequence of EHL in its entirety. All discussions of the model built and herein presented are underpinned with the caveat that the structure cannot and is not considered complete until such time as the protein sequence can be verified by other methods. For this reason, solvent content such as water molecules have not been added to the model and the hydrogen bonding not resolved as even at the reasonably high resolution of 1.6 Å individual atoms are still ambiguous. The local environment of a particular residue may indeed give further weight to a particular side chain being assigned, however caution has been taken not to overstate the structure. The method used which is discussed more in depth later was to minimize model bias, in the absence of sequence data and to prevent an overreliance on the molecular replacement solution. As discussed in Hou *et al.* (2007) whilst high resolution data and the resulting electron density maps can be used to inform sequence, ambiguity however cannot be resolved for residues such as Asparagine/Aspartic Acid, Glutamine/Glutamic acid, Valine/Threonine, and in some cases due to disordered side chains this is also the case for Cysteine/Serine, Lysine/methionine and potentially Isoleucine.

Whilst the RMSD for bond angles and the R factors of the refined model, though higher than would be ideal, are in good agreement, and the Ramachandran plot shows a reasonable standard of geometry, the model of EHL presented here, is not a complete structure. The interesting features of the model are highlighted and with the consideration that these were built from a main chain poly-alanine backbone trace in shelxE4 (Sheldrick, 2010), the model does at this point certainly confirm the RIP nature of the protein as well

as its active site homology to various other rRNA N-glycosidase proteins such as Mistletoe Lectin and Ricin (Krauspenhaar *et al.,* 2002, Weston *et al.,*1994).

In the case of EHL the strategy employed was first to screen for crystal growth using the sitting drop vapour diffusion method with four differing commercially available screens. The resultant crystals of those initial experiments diffracted to reasonably high resolution so that optimisation was at this time not required, as the primary aim of the crystal is to provide good diffraction data (Chayen and Saridakis 2008).

Literature suggests starting crystallisation studies with a purified protein with a concentration upwards of 5 mg ml^{-1} (Bergfors, 2009, Chayen and Saridakis 2008). EHL fractions were pooled and concentrated in an attempt to reach this ideal concentration. However it was found that above 4 mg ml^{-1} using the stated protocol, EHL formed a precipitate and this presented a requirement for a methodological change which was not viable at that time. A decision to proceed with screening of protein at 4 mg ml^{-1} was taken.

4.2 Materials and methods

Preliminary crystal screens were set up to identify favourable crystallisation conditions using 4 commercial screens covering 384 conditions at two different temperatures (4°C and ambient temperature).

- JCSG-plus
- PACT premier
- MIDAS
- Morpheus

Primary screening was carried out using a strategy developed by Newman *et al*. (2005). This strategy combines a systematic grid screen of pH, anion and cation testing (PACT) and the sparse matrix screen developed by the Joint Centre for Structural Genomics (JCSG). The

JCSG+ screen is based on previously successful crystallisation parameters, which maximises coverage of crystallographic space with no redundancy (Jankarik and Kim, 1991). This combined strategy is designed to minimise the amount of conditions required to be screened. The MIDAS screen (modern intelligent dynamic alternative screen) is based on alternative polymeric precipitants developed by Dr Clemens Grimm and was also used due to its non-reliance on polyethylene glycols (PEG) as precipitants. MIDAS also contains a restricted range of pH and salt concentrations designed around physiological parameters (Grimm *et al*., 2010). The final screen used was Morpheus, which also incorporates a cryo-protectant in each condition and contains a selection of low molecular weight ligands previously found in many PDB structures. Morpheus has been designed by Fabrice Gorrec, in collaboration with the Medical Research Council Laboratory of Molecular Biology (LMB) at Cambridge (Gorrec, 2009). Each screen was produced under licence and purchased from Molecular Dimensions Ltd. The chemical conditions of each screen as stated by Molecular Dimensions can be found in Appendix 2.

The screens were manually set up using 48 well MRC maxi sitting drop plates (Medical Research Council supplied by Molecular Dimensions) with a reservoir volume of 200 µl and a 2 µl drop volume with a 1:1 ration of protein to well solution. Plates were checked daily to begin with and then weekly after a period of six weeks. Initial hits were identified in the JCSG+ condition H4 in the form of a series of microcrystals and needles (Figure 4.2). Other crystals grew over a period of between 8 weeks and over 12 months as outlined in the results section.

Data were collected on PACT D10 crystal at the Diamond Light Source on beamline I04-1 under proposal MX8997-34 . The crystal was cryoprotected with a 1:4 ratio of PEG 400 to

well solution. Data were collected at a wavelength of 0.9795 Å and reflections recorded by a Pilatus 6M detector . The crystal diffracted to a resolution of 1.6 Å. A Crystal grown in PACT condition B12 was also cryoprotected and data were collected to a resolution of 1.8 Å at an oscillation angle of 0.10 degrees. A third crystal (PACT E5) was analysed using a Rigaku FRE+ Superbright copper rotating anode X-ray source with an incident wavelength of 1.54 Å (CuKα). The crystal was cryoprotected with combination of glycerol and well solution. The detector used was a Saturn 944+ and data with resolution of 1.9 Å were collected.

Data processing beginning with spotfinding, indexing and integration of the diffraction data were completed with iMOSFLM (Battye *et al.* 2011). Scaling and merging of data were completed through SCALA (Aimless, Pointless and ctrunctate) (Evans, 2006; Evans, 2011) using the CCP4i GUI. PHASERMR (McCoy *et al.,* 2007) generated a single solution using 1M2T.pdb used as a molecular replacement search model. A poly-alanine back bone trace was created with SHELXE4 (Sheldrick, 2010). CHAINSAW (Stein, 2008) was used to create a poly-alanine model from the molecular replacement solution. The MR model was superposed in COOT (Elmsley *et al.*, 2010) to the SHELXE4 trace with residues subsequently built into the MR solution where density from the shelxe4 maps supported it. Four rounds of refinement were carried out in REFMAC5 (Murshudov, 2011; Murshudov, Vagin and Dodson, 1997). Previously published residues (Kumar *et al*., 1993) were manually docked into the model using CCP4 (Winn *et al.,* 2011) Molecular Graphics (CCP4MG) sequence viewer.

4.3 Results

4.3.1 Crystal screening

The first EHL crystals were observed in the 4°C JCSG screen which appeared to have yielded needle form crystals (Figure 4.2a) as well as the larger cluster seen in Figure 4.2b. In the 20°C screen, a spread of microcrystals was observed (Figures 4.2 c and d). All scale bars are estimated due to equipment limitations. The crystallisation conditions are summarised in Table 4.1 and a full account of the screens in Appendix 2. A second set of larger crystals was observed in the JCSG screen after approximately 20 weeks which can be seen in Figure 4.3. Further crystals grew in the PACT screen and were observed after approximately 24 weeks, a large rhomboid form crystal grew in PACT condition D10 at 20°C and also a crystal found at 4°C in PACT condition B12 as seen in Figure 4.4. Data was collected on both crystals at Diamond Light Source synchrotron beamline I04-1 to 1.6 and 1.8 Å respectively. A Final series of crystals was observed after 15 months in various PACT conditions at 20°C as seen in Figure 4.5. Data were collected on crystal B grown in PACT condition E5 using a Rigaku FRE+ source and Saturn 944 CCD detector at Oxford University Chemistry department to 1.9 Å

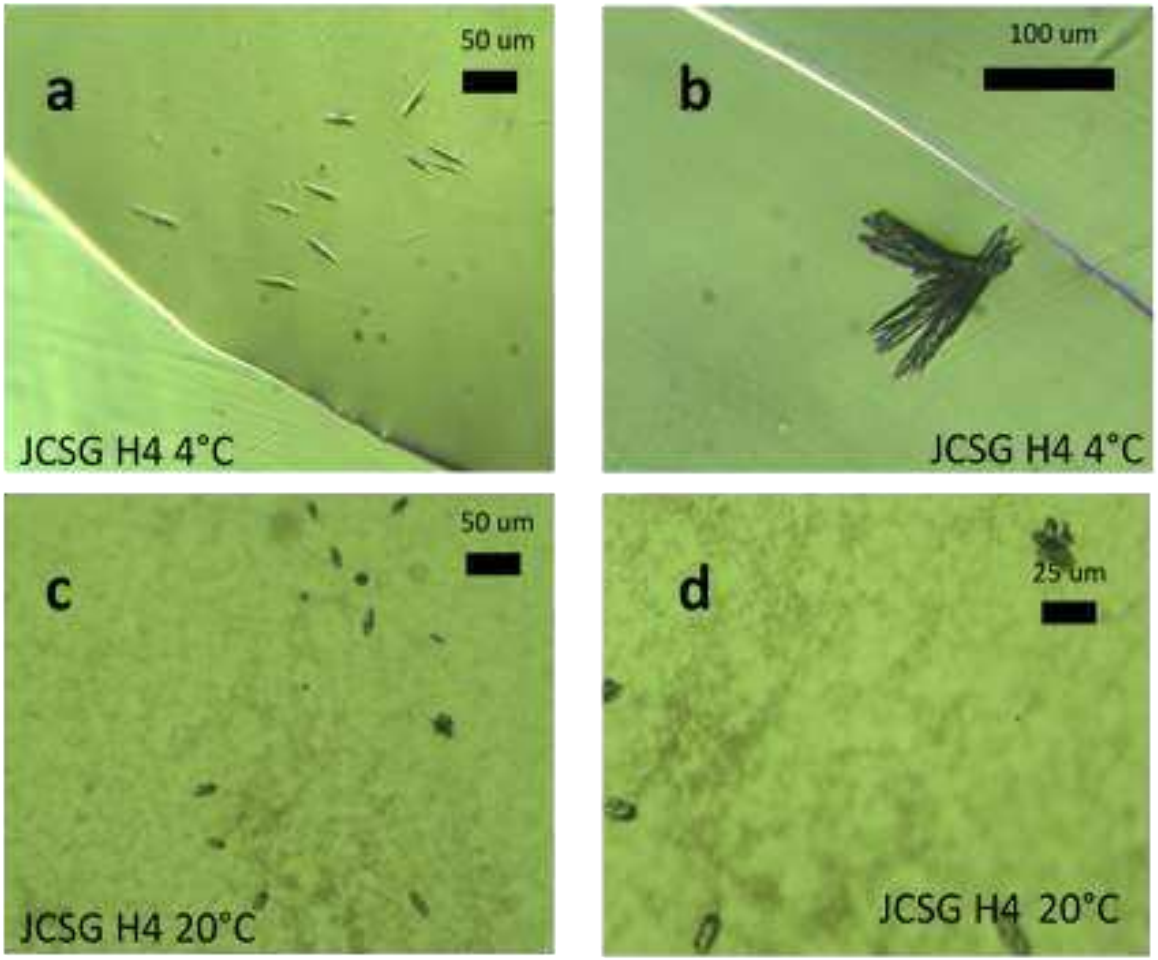


Figure 4.2 First crystals observed for EHL **a**) needle form crystals JCSG H4 **b**) cluster of needles JCSG H4 **c**) microcrystals in JCSG at 20°C **d**) microcrystals in JCSG H4 at 20°C.

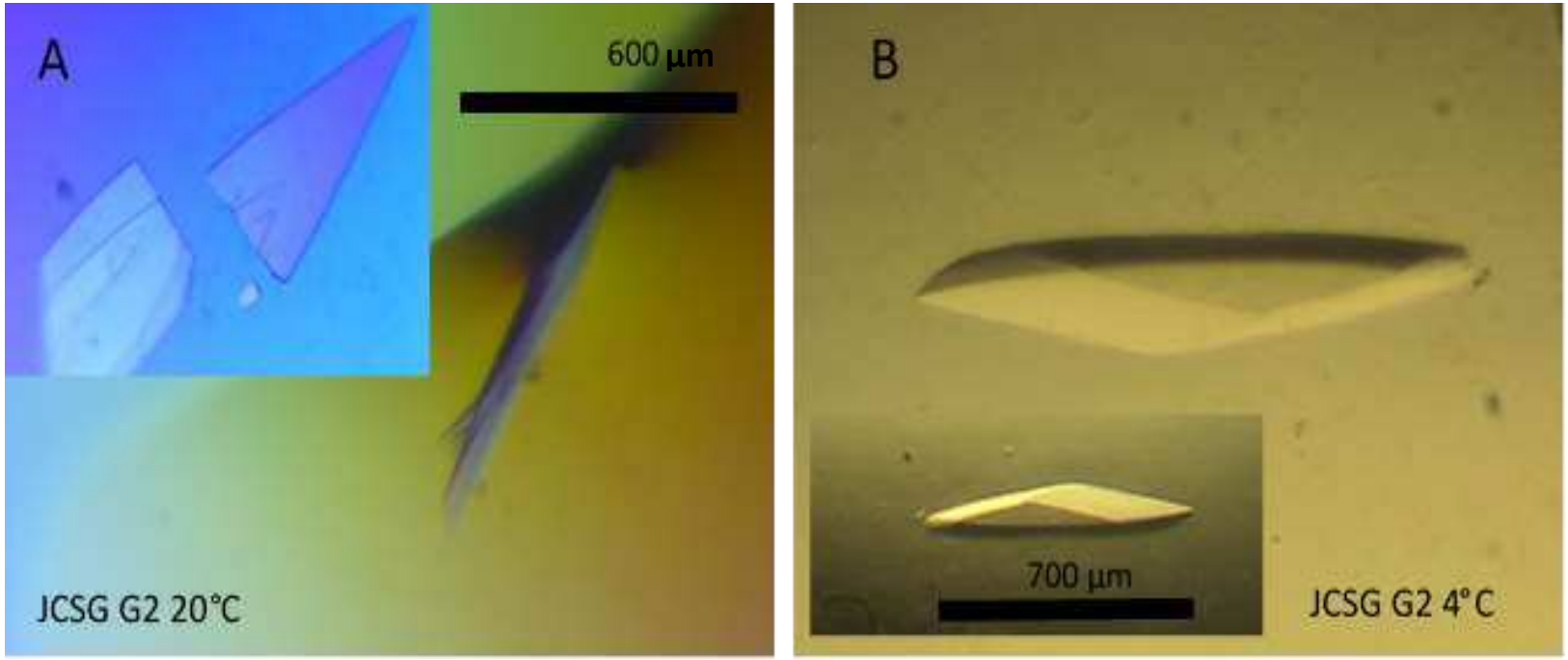


Figure 4.3 Second set of crystals observed after approximately 20 weeks. **A**) JCSG G2 crystal was used to assist in expertise and practice manipulation techniques. **B**) JCSG G2 at 4°C a large crystal was observed and data were collected at synchrotron to 3.2 Å.

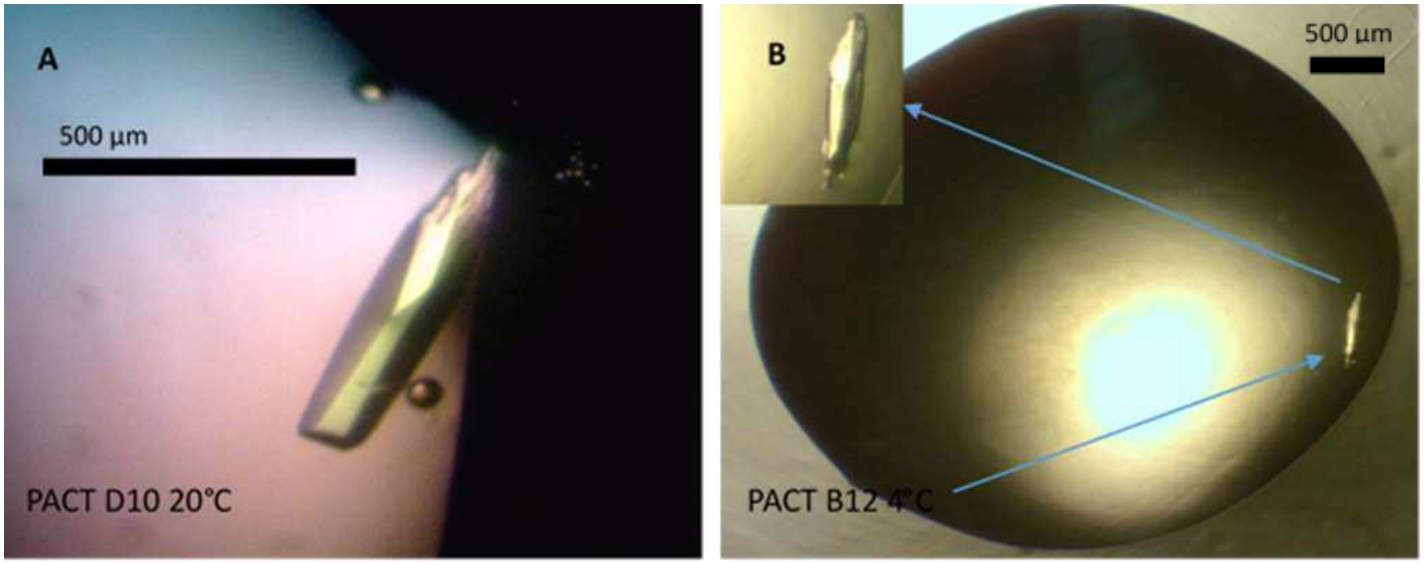


Figure 4.4 Crystals in PACT screen observed after approximately 24 weeks **A**) rhomboid form in PACT condition D10 at 20°C **B**) Similar morphology crystal found at 4°C in PACT condition B12. Data collected on both crystals at Diamond Light Source synchrotron beamline I04-1 to 1.6 and 1.8 Å respectively.

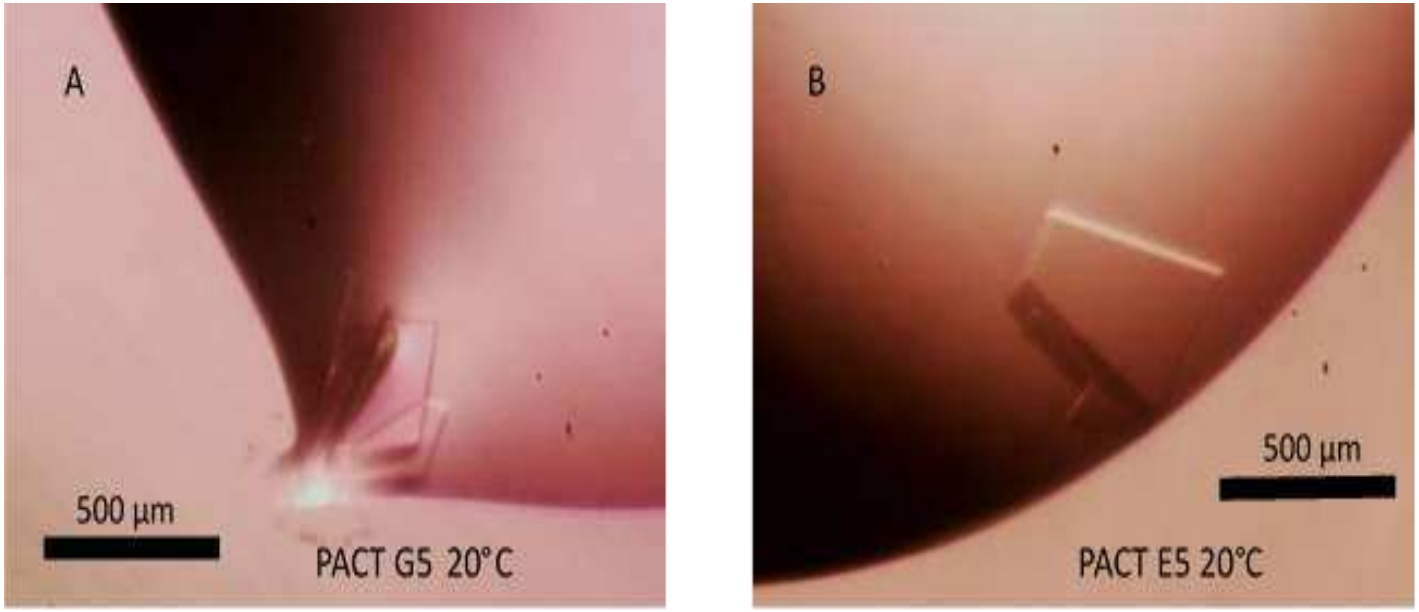


Figure 4.5 Final series of crystals observed after 15 months in PACT conditions at 20°C, data were collected on crystal B grown in condition E5 using a Rigaku FRE+ source and Saturn 944 CCD at Oxford University Chemistry department to 1.9 Å **A**)PACT G5 multiple sheet like crystals similar to those seen in JCSG G2 **B)** unusual morphology in PACT E5.

4.3.2 Data collection and analysis

Data were collected at the Diamond light source Synchrotron on beamline I04-1 on the crystal grown in PACT premier condition D10. It was characterized as belonging to the monoclinic space group $P12_11$ with unit cell dimensions a = 46.84 Å b = 94.39 Å c = 71.12 Å and α = 90° β = 96.47° γ = 90°. Based on these cell parameters the contents of the crystallographic asymmetric unit is a single heterodimeric molecule with a Matthews Coefficient (Matthews, 1968) of 2.60 and a solvent content of 53% (Collaborative Computational Project, Number 4. 2011). Data collection statistics are shown in Table 4.2 for this crystal. These data were used for model building . Data were also collected on PACT B12 at the DLS synchrotron (MX8997-34 on I04-1) and E5 crystals at the University of Oxford Chemistry department using a Rigaku FRE+ Superbright copper rotating anode X-ray source details of which can be found in Table 4.1.

The SHELXE4 (Sheldrick, 2010) model as described in the methods section underwent many cycles of building and refinement in COOT (Elmsley *et al.*, 2010) and REFMAC (Murshudov *et al.*, 2011, Murshudov, Vagin and Dodson, 1995) until the final model which can be seen in Figure 4.7 was arrived at.

A Ramachandran plot of the refined model was generated in COOT (Elmsley *et al.*, 2010) (Figure 4.6) which shows the energetically favoured protein backbone torsion angles of phi and psi (the angles between N-Cα and Cα-C respectively) for a given amino acid residue sequence. Residues should lie in favoured regions on the plot when a geometrically plausible model has been built. A percentage of between 1-2 % outside of allowed regions is normal (Rupp, 2009)

Table 4.1 EHL crystal hits obtained from initial commercial screens JCSG and PACT

Screen	Well	[Conc}	Morphology	Time to appear	Resolution	Source	Space group
JCSG +	**H4** 0.2 M Calcium Chloride dehydrate, 0.1M BIS-Tris at pH 5.5 in 45% v/v 2,4-methyl pentanediol (MPD)	4.0 mg/ml	Needles and microcrystals	8-10 weeks			
JCSG +	**G2** 0.02M Magnesium chloride hexahydrate 0.1M HEPES buffer at pH7.5 with precipitant 22% w/v Poly (acrylic sodium salt) 5100	4.0 mg/ml	Prism/sheets	12 weeks	3.2 Å	DLS Synchrotron	$P2_12_12_1$
PACT	**D10** 0.2M Magnesium Magnesium chloride hexahydrate 0.1M Tris, pH 8.0 20% w/v PEG µ0	4.0 mg/ml	Multifaceted prism rough end sections	24 weeks approx.	1.6 Å	DLS Synchrotron	$P12_11$
PACT	**B12** 0.01M Zinc Chloride, 0.1M MES pH 6.0 20%w/v PEG 6000	4.0 mg/ml	Multifaceted prism rough end sections		1.8 Å	DLS synchrotron	$P2_12_12_1$
PACT	**E5** 0.2M Sodium Nitrate, 20% w/v PEG 3350	4.0 mg/ml	prism with embedded sheet	12 months +	1.9 Å	Rigaku FRE+ Superbright copper rotating anode.	P2
PACT	**G5** 0.2M Sodium Nitrate, 20% w/v PEG 3350, 0.1M BIS-Tris Propane, pH7.5, 20% w/v PEG 3350	4.0 mg/ml	sheets	8 months approx.			
PACT	**H5** 0.2M Sodium Nitrate, 20% w/v PEG 3350, 0.1M BIS-Tris Propane, pH 8.5, 20% w/v PEG 3350	4.0 mg/ml	irregular prism	12 months +			
PACT	**F6** 0.2 M Sodium Formate 0.1M BIS-Tris Propane pH6.5 20% w/v PEG 3350	4.0 mg/ml	irregular prism	12 months +			

Table 4.2 Data collection statistics for EHL crystal PACT D10 from which model was built.

Crystal	**EHL**	**PACT D10**	**20°C**
Data Collection			
Source	DLS I04-1	Pilatus Detector 6M	
Wavelength (Å)	0.92819		
Space group	$P12_11$		
Cell dimensions			
***a,b,c* (Å)**	46.84, 94.39,71.12		
α,β,γ (°)	90° 96.47° 90°		
	Overall	**Inner Shell**	**Outer Shell**
Low Resolution (Å)	47.19	47.19	1.63
High Resolution (Å)	1.60	8.76	1.60
R_{merge}	0.107	0.038	0.678
I/σI	11.2	24.0	4.7
Completeness (%)	99.6	99.6	99.3
Multiplicity	3.1	3.2	3.2
Refinement			
Resolution (Å)	70.667 - 1.6		
No. reflections	80904		
No. non Hydrogen Atoms	3664		
R factor	0.2381		
R free	0.2597		
R.m.s deviations			
Bond lengths (Å)	0.0264		
Bond Angles (°)	2.1953		

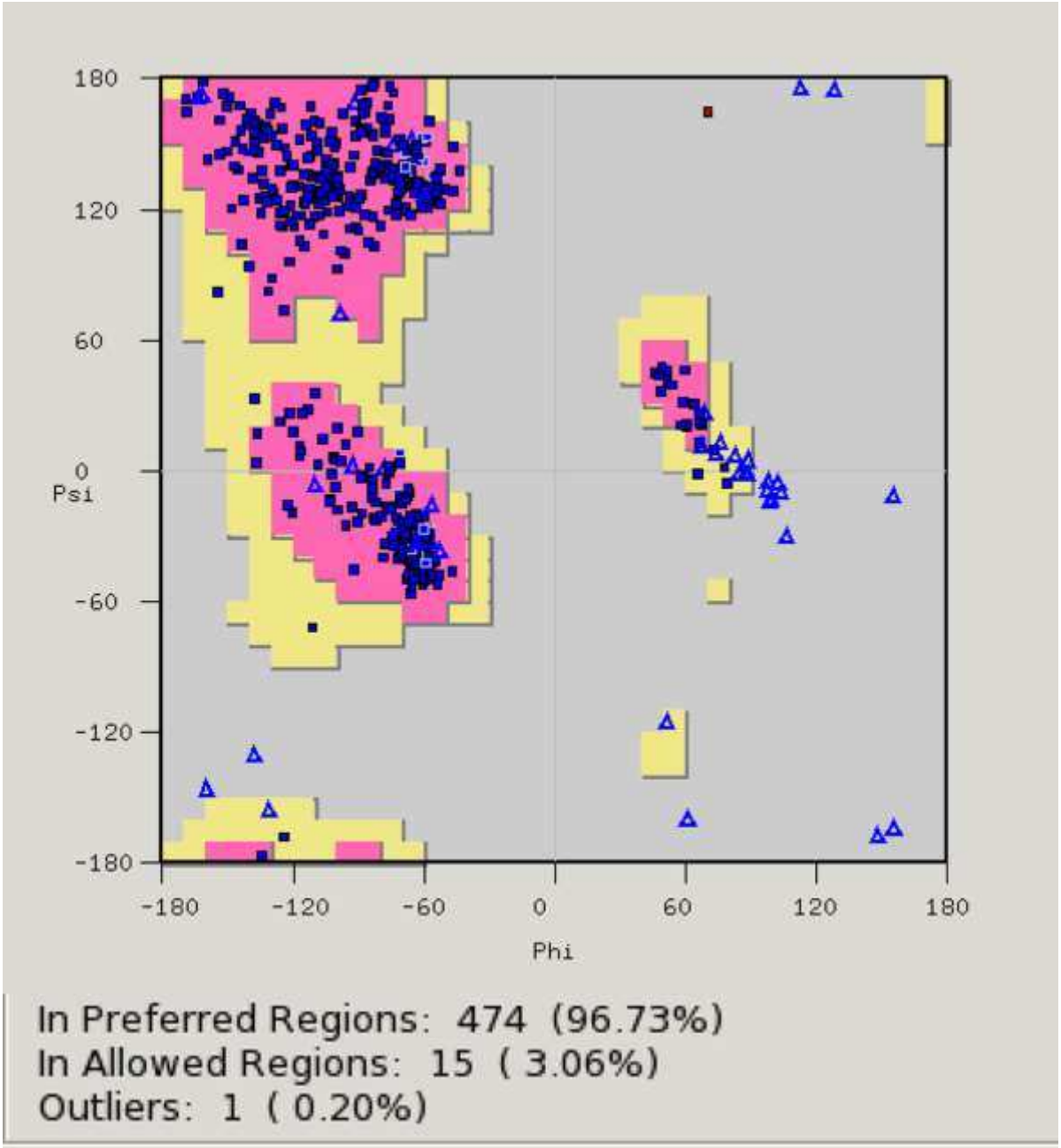


Figure 4.6 Ramachandran plot of the EHL model generated with COOT (Elmsley *et al.* 2010). This shows the energetically favoured protein backbone torsion angles of phi and psi (the angles between N-Cα and Cα-C respectively) for a given amino acid residue sequence and whether the model is agreement with these.

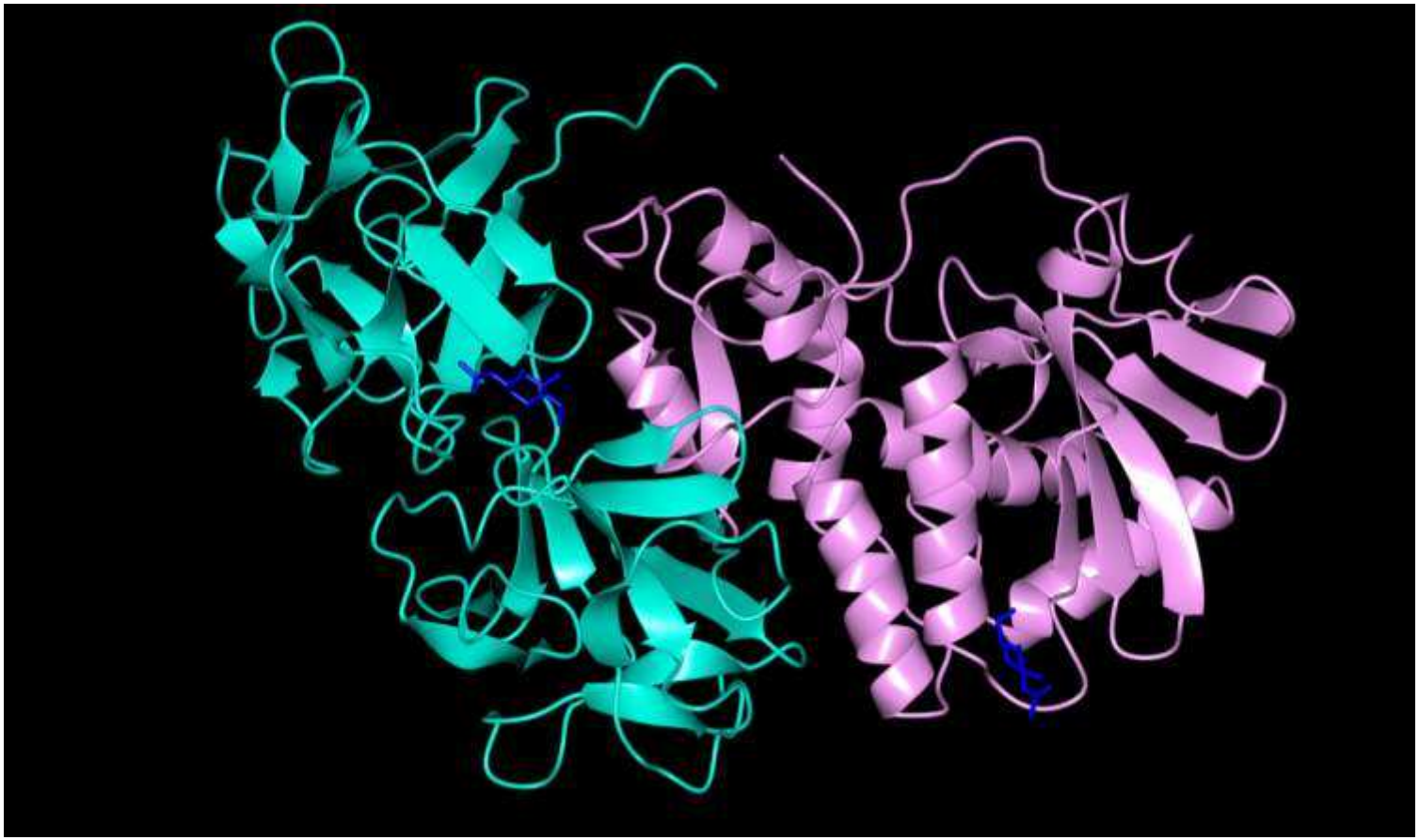


Figure 4.7. A model of EHL built into electron density and poly-alanine backbone trace made in SHELXE4 (Sheldrick, 2010). Two N-linked glycosylation sites were observed in the electron density maps and have been assigned as NAG on chain A at Asn 11 and on chain B at Asn 355.

4.4 Discussion

The current strategies for building model structures of a target protein often rely on the amino acid sequence of the protein being known and used in conjunction with the reflection data collected at the X-ray source to autobuild a structure using a software pipeline. In the event of the sequence being unavailable as with EHL, various techniques may be employed to build a model which actually then informs the sequence. Whilst complex techniques such as cross crystal averaged maps are useful, for the purposes of this thesis, a more simplified approach was taken. Cross crystal averaging takes advantage of the two crystal systems found in the diffraction data sets and can provide extremely clear maps to work with for the experienced crystallographer. The phase information is strengthened by the increase in signal over noise ratio due to solvent flattening. The decision not to use this method was taken in an effort to reduce the possibility of introducing systematic error in the complexity of making and working with cross crystal averaged maps for the beginner. From the highest resolution data, a poly-alanine backbone trace was built using SHELXE4 (Sheldrick, 2010). This method does not build a model structure, it attempts to build peptide main chain in any place it can reasonably do so within the electron density, effectively minimising model bias from the molecular replacement solution. Once the protein sequence for EHL becomes available, it will be a small step to read in the sequence to the maps already created to complete the structure. The limitations of the SHELXE4 maps include having multiple fragments to work with, areas of density which have not been built into the backbone, and areas that are difficult to build into.. However the maps produced are extremely good and allow *ab initio* building with the caveat that ambiguities will ultimately remain for residues which cannot be reasonably

assigned with total confidence such as Asparagine/Aspartic acid, Glutamine/Glutamic acid, Valine/Threonine. These side chains require atomic resolution of at least 1 Å to definitively assign them unless the local environment strongly suggests a particular choice, such as in a hydrophobic pocket or strongly hydrophilic region or as part of a hydrogen bonding network (Hou *et al.,* 2007). Using prior knowledge of RIP structures, a molecular replacement model was also built. The peptide fragments which were known for EHL were used to find homologs within the RIP family with at least 40% similarity (Schwarzenbacher *et al.,* 2004) for phasing. The scaled and merged data (SCALA) (Evans 2011, Evans 2006) from the $P12_11$ EHL crystal were used and phases extracted from the mistletoe lectin IM2T.pdb (Krauspenhaar *et al.* 2002) file using PHASERMR (McCoy *et al.*, 2007), and a single solution was found. The next step was to take the output .pdb from PHASERMR and run it into CHAINSAW (Stein, 2008) in CCP4. This was in order create a poly-alanine model which could be superimposed using SSM in COOT (Elmsley *et al.,* 2010) onto the SHELXE4 back bone trace structure to effectively create a visual feedback between the two structures. The benefit of this was to highlight the conserved residues which fit the unbiased trace build, join peptide chains that were fragmented and to cut out regions of the molecular replacement structure where there was no electron density shown in the SHELXE4 (Sheldrick, 2010) maps. After multiple rounds of refinement using REFMAC5 (Murshudov, 2011; Murshudov, Vagin and Dodson, 1997) the structure was then opened in CCP4 MG (Winn *et al.,* 2011) and the four peptide fragments from the early published work on EHL (Kumar *et al.,* 1993) were located at potentially viable coordinates within the sequence viewer. Using the refined maps the density was matched to the published small fragments,

and with the exception of cysteine 309, which was published as threonine (Kumar *et al.,* 1993), the structure was found to match in 70/71 residues.

As shown in Figures 4.8 and 4.9.1,2 and 3, the superposed active site residues of the ricin A-chain, the catalytic region where depurination of the rRNA A_{4234} is bound and hydrolyzed (Monzingo and Robertus, 1992), are conserved in EHL as the model stands with one exception, tyrosine 80. The active site of 3rti includes residues Trp 211 (187), Arg 180 (156), Glu 177 (153), Tyr 123 (111), Tyr 80 (Ala 73) and Val 81 (74), these all closely matched in the EHL model (coordinates in parantheses) except for the Tyr 80 which appears as an alanine (73) in the EHL density maps and is proposed to be critical to the binding cleft for the adenine substrate in ricin (Yan *et al.*, 1997).

The aforementioned residues are invariant in the family of ribosome inactivating proteins and have been subjected to site directed mutagenesis experiments (Kim and Robertus 1992) in the ricin A protein. In these experiments a one hundred fold reduction in enzymatic activity was observed when Tyr 80 was mutated to serine (Kim and Robertus 1992) and whilst the EHL model structure will certainly undergo considerable refinement in the future, the variant residue at the EHL A-chain site 73 could be hypothesised to show an indication as to the reason for the reduced toxicity of the protein compared to Ricin and other Type II RIPs.

Figure 4.8 EHL model structure with Ricin (*Ricinus communis*) 3rti.pdb structure superposed. The 3rti structure is represented in a purple colour and the EHL model in cyan.

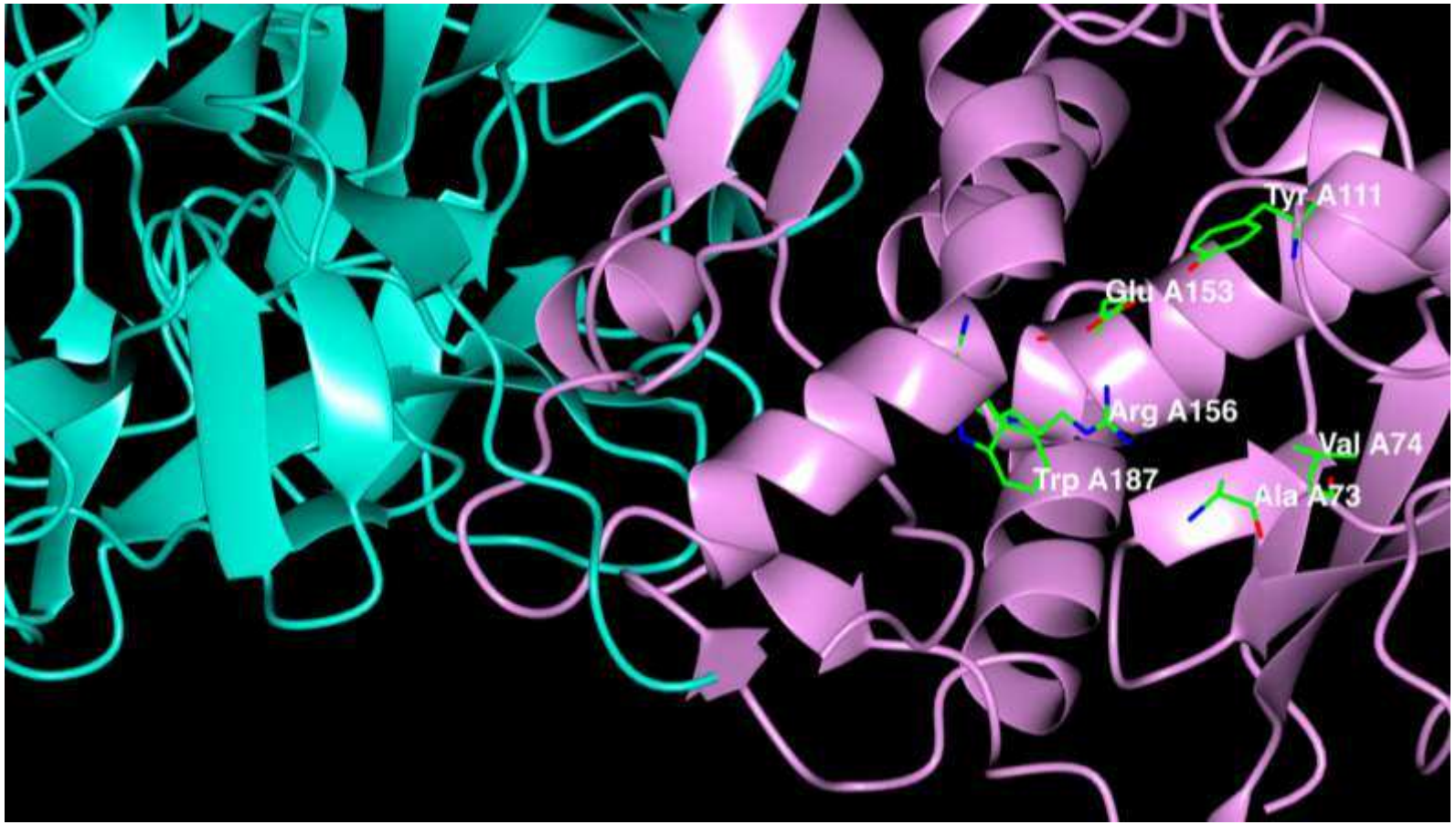


Figure 4.9.1 EHL model with suggested A chain (lilac) active site residues highlighted in green. These are based on the corresponding site and residues from the SSM alignment of the Ricin structure 3rti.pdb seen in figure 4.8.

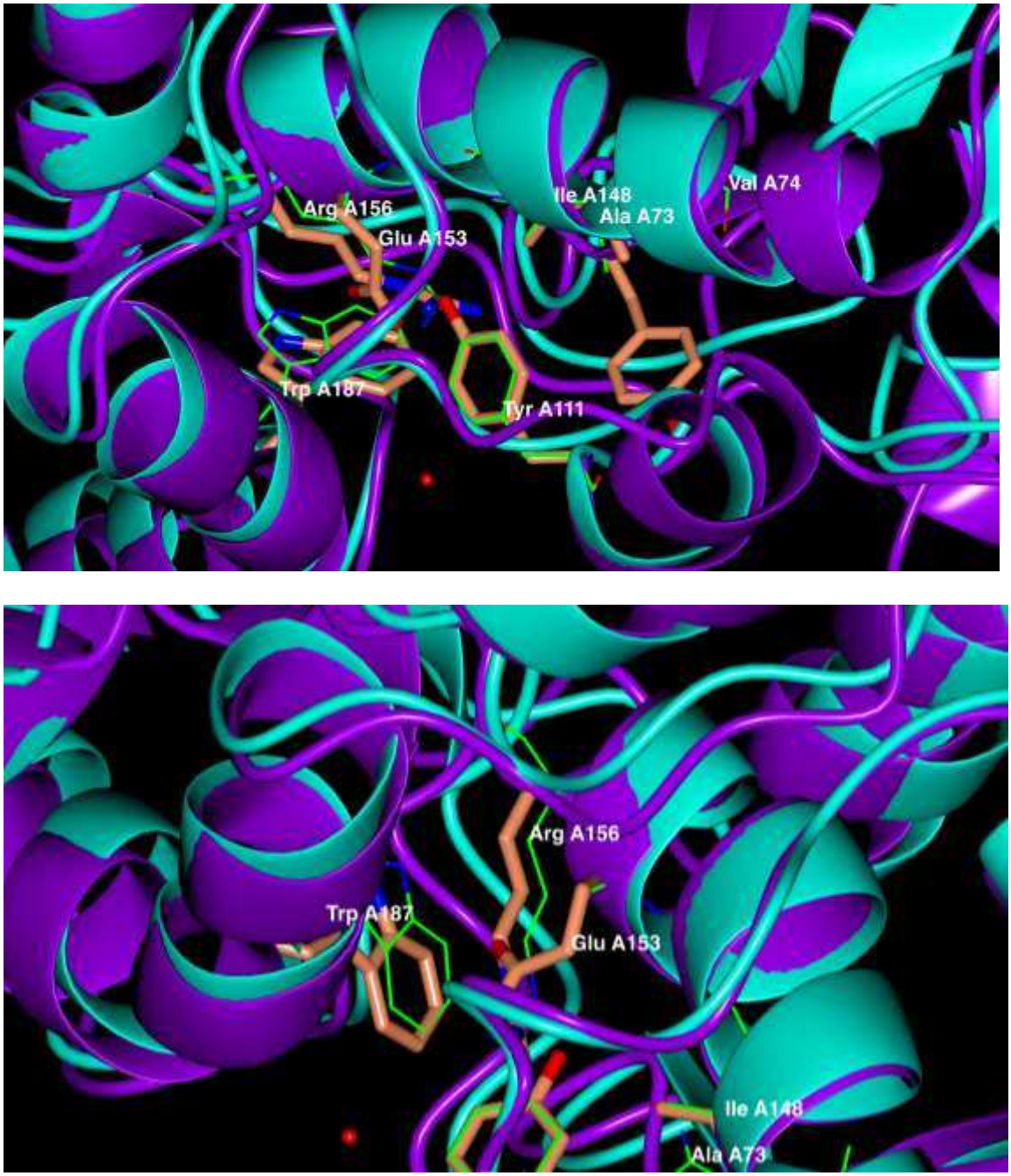


Figure 4.9.2 Superposition of EHL and 3rti A chain active site residues in CCP4MG, EHL residues in bright green, 3rti in sand colour, both orientations +show conservation of tryptophan 211(187), arginine 180 (156), glutamic acid 177 (153), isoleucine 172 (148), tyrosine 123 (111) and C)valine 81 (74) tyrosine 80 is not conserved in EHL in this model and shows alanine at position A73 of EHL. Values in parentheses are for the EHL model.

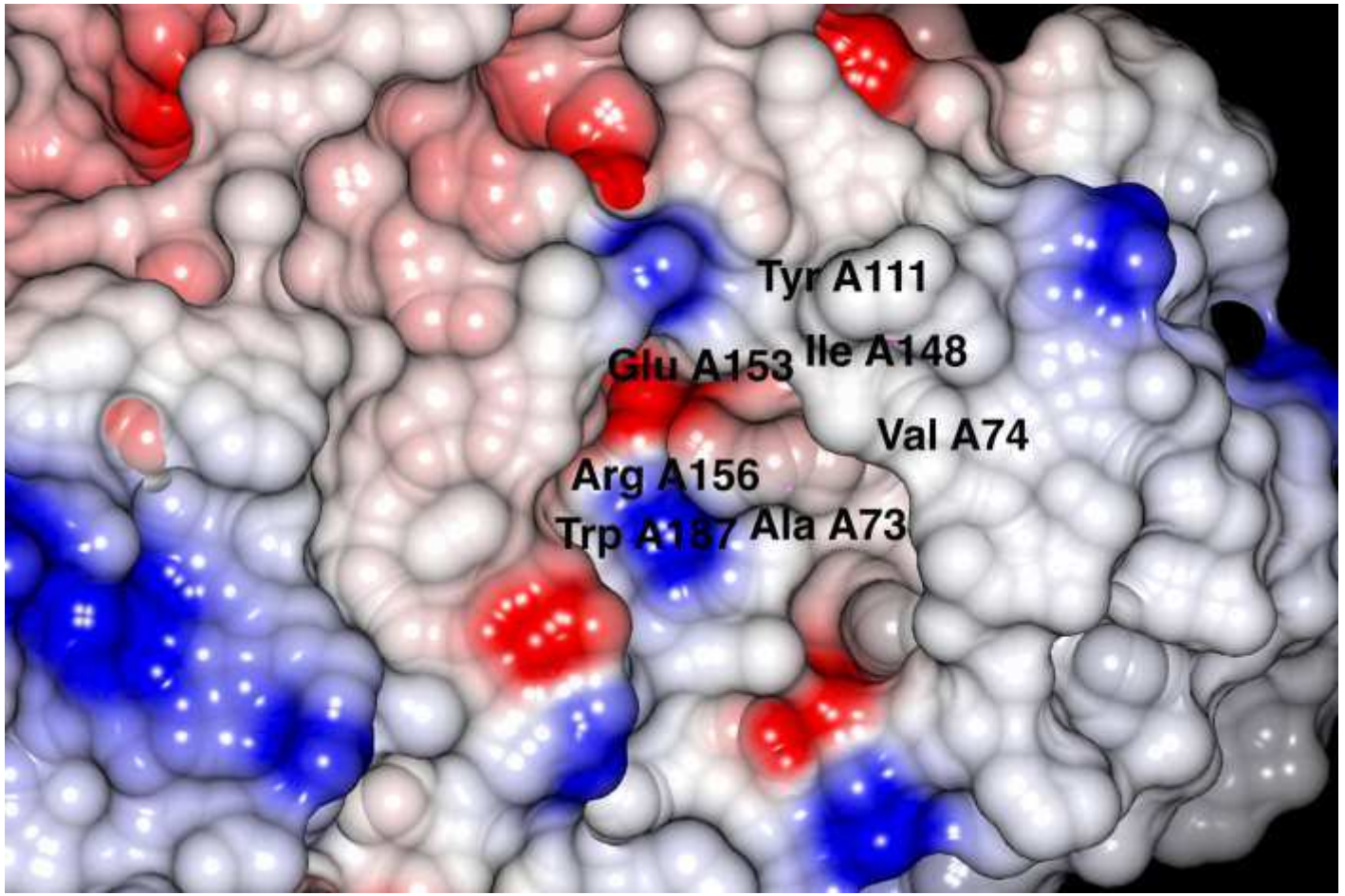


Figure 4.9.3 Theoretical proposed EHL active site surface pocket model of electrochemical interactions (red indicates a negative charge, blue a positive charge and white areas neutral, using conserved region from 3rti as a template(Monzingo and Robertus, 1992).

A number of other conserved features can be identified, such as the $(QxW)_3$ motif on the B-chain found in hugely variant lectin sequences and hypothesized to be involved in carbohydrate binding (Hazes, 1996). BLASTp identified fourteen putative sugar binding sites in the model, as seen in Chapter Three (figure 3.6). A number of disulphide bonds were present in the model; Cys 251 with Cys 270, Cys 381 with Cys 396. However, of note is the lack of inter chain disulphide bonding, which will be resolved when the genomic studies of Winter Aconite become available.

A representative electron density map showing the overall map quality is shown in Figure 4.10. Two N-linked glycosylation sites were added to the structure as observed in the electron density maps (Figures 4.11 and 4.12), namely Asn 11 and Asn 355 of the A and B chains respectively.

In conclusion, the model herein presented has demonstrated the possibility of building a structure, which is plausible both geometrically and biologically in the absence of sequence information. Though R-factors are global indicators so will not highlight incorrect modelling on a local scale and are on the limits of acceptability for this model, the <2% difference between the R_{merge} (23.8%) and R_{free} (25%) would suggest a sensible model (Wlodawer *et al.,* 2008).

The final addition of the peptide fragments published in Kumar *et al.* (1993) into the refined model has strengthened the proposed structure in 4 regions as discussed in Chapter Three. Future work will include the building of a model using the same techniques on the remaining 2 data sets, at 1.8 and 1.9 Å respectively and conducting a comparison to this higher resolution structure. Ultimately, crystallising EHL with a substrate analog such as adenosine mono/triphosphate would show the relationship between the Y80 and Y123

binding of the adenine ring from Adenosine$_{4234}$ of the rRNA substrate seen in ricin (Yan *et al.,* 1997) and the differential effect of the variant A73 in EHL on that substrate binding. Finally, to follow this work to conclusion and deposition of the protein coordinates in the World Wide Protein Data Bank, the addition of the genomic sequence data to the model will allow a high resolution structure to be built into the electron density map.

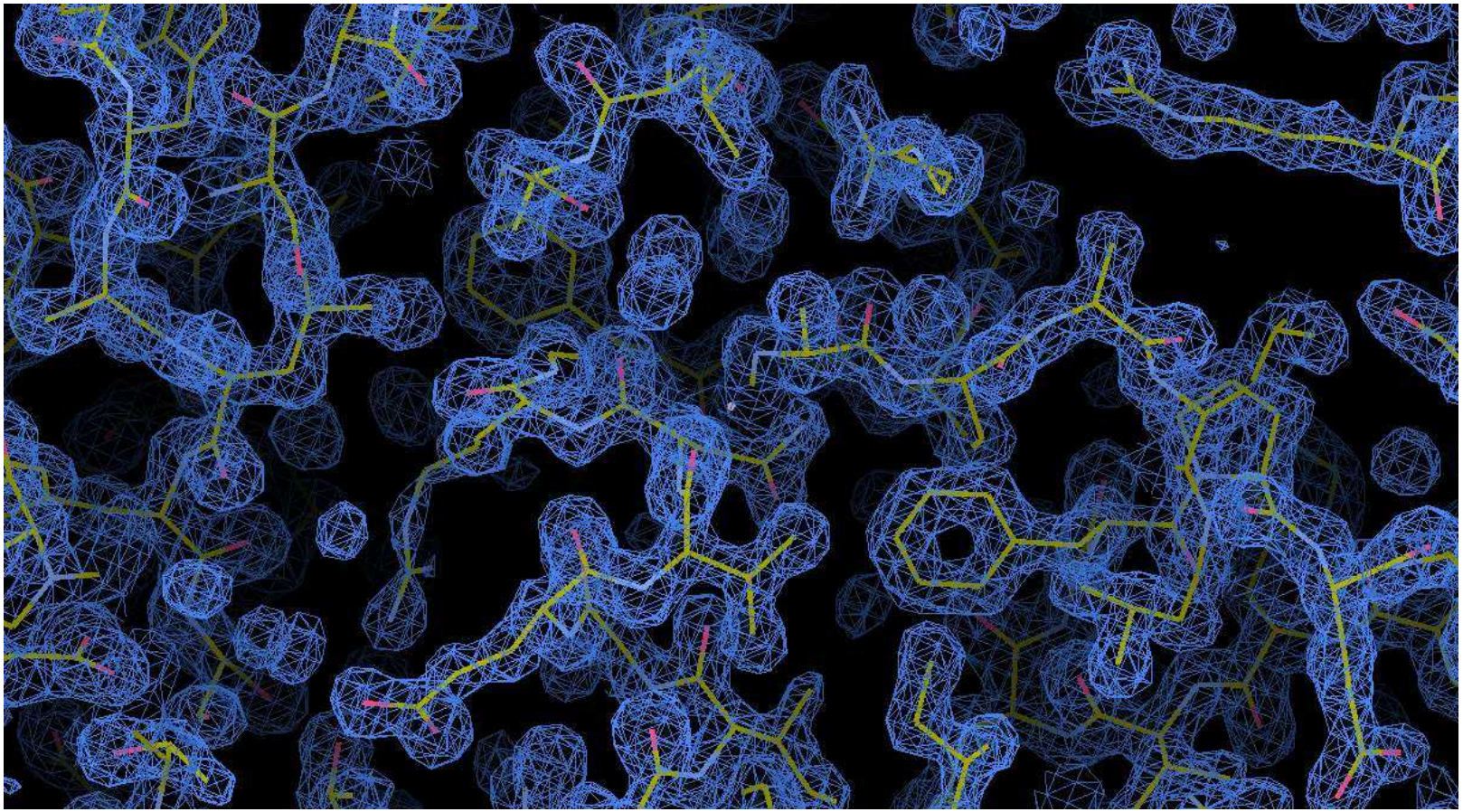


Figure 4.10 Representative regions of density from the SHELXE4 back bone trace job. Untraced regions impact on the completion of building the model structure.

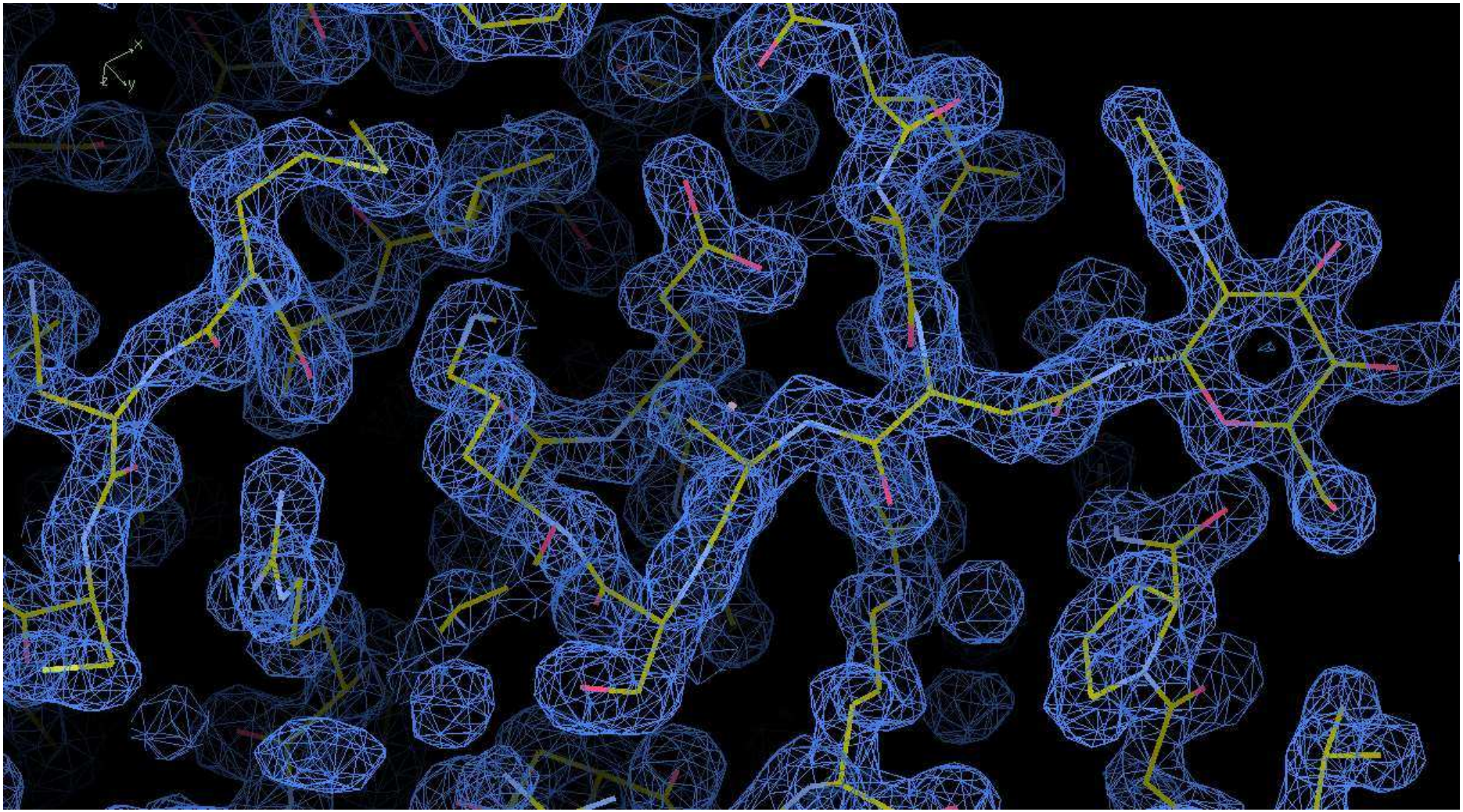


Figure 4.11 Electron density indicates glycosylation site at A: Asn 11.

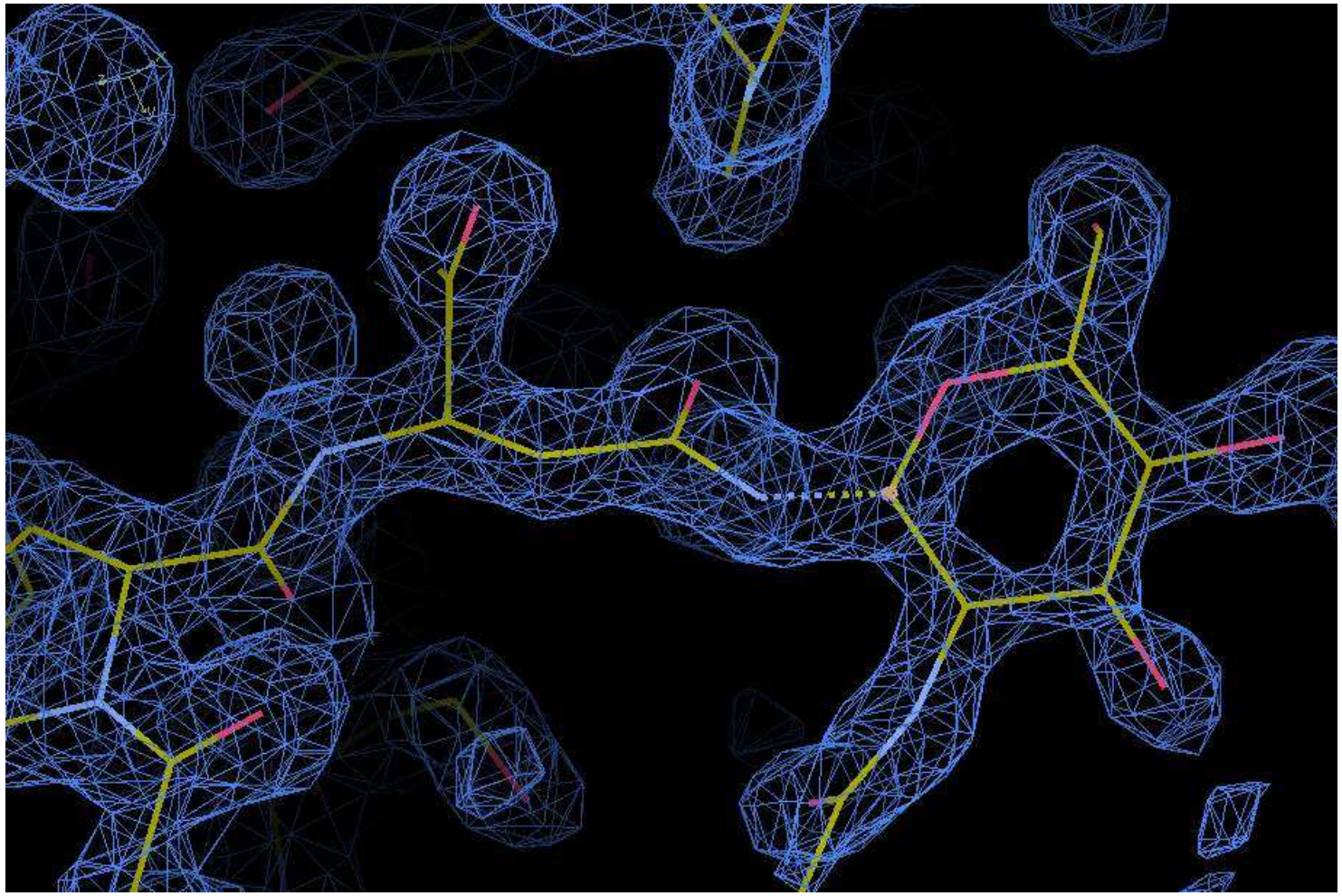


Figure 4.12 Electron density supports glycosylation of B: Asn 355.

Chapter 5: Biological activity against *C. elegans*

Parts of the work reported in this chapter have been published as:

McConnell *et al.* (2015), Winter Aconite (*Eranthis hyemalis*) Lectin as a cytotoxic effector in the lifecycle of *Caenorhabditis elegans*. *PeerJ* 3:e1206; DOI 10.7717/peerj.1206

All statistical analyses were undertaken in 'R' (R Core Team 2013) by Dr Simon Harvey.

5.1 Introduction

In this experimental chapter, the effect of EHL on the free living nematode *Caenorhabditis elegans* (*C. elegans*) was investigated. *C. elegans* is a well-established model organism for initial toxicological studies due to the conserved nature of its biological and biochemical processes, including the stress response and disease pathways (Boyd, Smith and Freedman, 2012). *C. elegans* is an excellent experimental system to use in toxicity studies due to the wide range of mutant phenotypes available, the short life cycle of the worms, their well documented reproductive and lifespan schedule, and their largely transparent body type which makes it possible to observe unusual effects easily (Corsi, Wightman and Chalfie, 2015). As with many other free living nematodes, *C. elegans* passes through a well-defined set of life stages (Figure 5.1) where they ultimately reach the adult stage at up to 1mm in length, having grown from the newly hatched L1 larval stage of approximately 0.25mm in 3 days. Of particular note is the ability of the nematode to enter an alternate life cycle stage

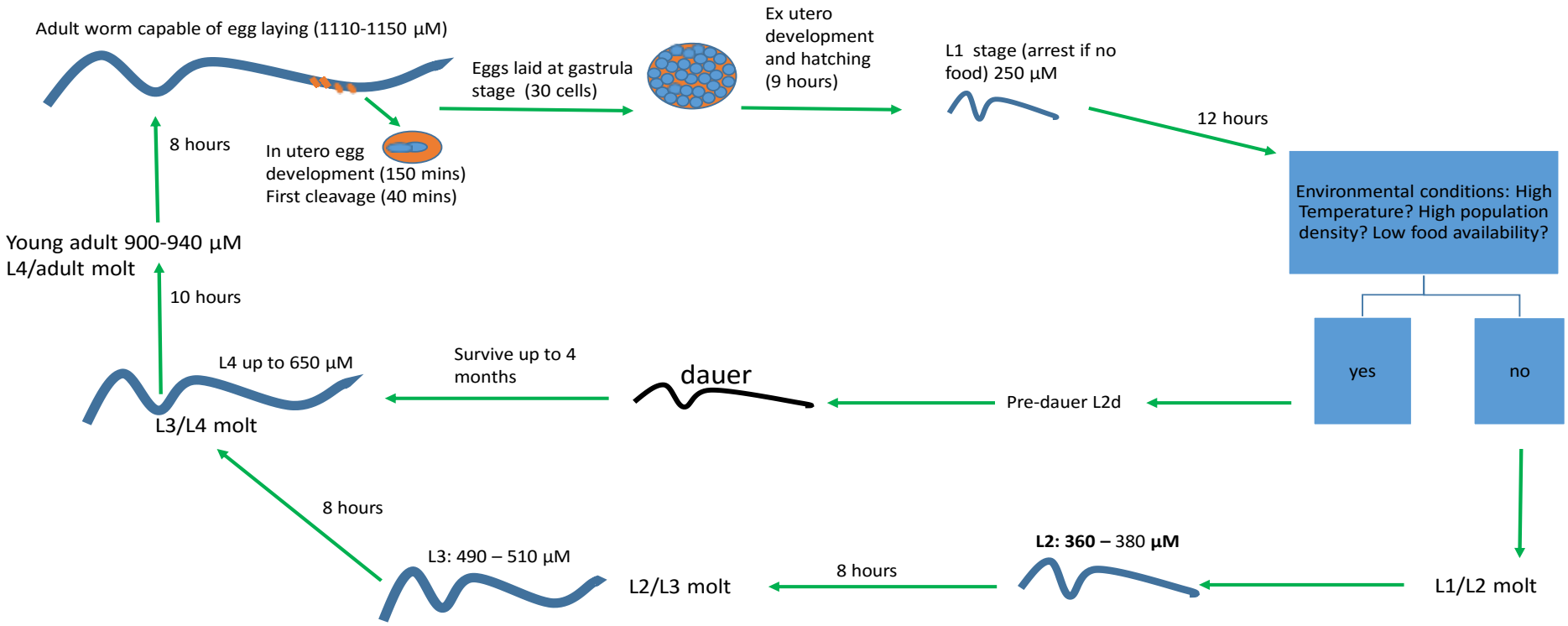


Figure 5.1 The life cycle of the free living nematode *C.elegans*. based on an original diagram from: Altun, Z.F. and Hall, D.H. 2005. Handbook of C. elegans Anatomy. In WormAtlas.http://www.wormatlas.org/hermaphrodite/hermaphroditehomepage.htm

known as the dauer larvae. Naturally induced dauer larval arrest occurs as a result of an L2d larva being exposed to environmental stresses such as a depleted food source or population overcrowding, though the chemosensory signals are detected by the L1 larvae. The worms then develop a cuticle which effectively seals them off from the external environment with an enhanced resistance to chemicals, with the mouth parts being 'plugged' so the worm cannot feed. This resulting dauer larvae is in an arrested developmental state, and can survive for months until conditions improve, when the cuticle is shed and normal development resumes (Corsi, Wightman and Chalfie, 2015).

The research reported here has revealed a range of phenotypes including direct mortality and a constitutive dauer formation phenotype that is consistent with neuronal ablation (Bargmann and Horvitz, 1991) when *C. elegans* is exposed to purified EHL.

5.1.2 Activity against *C. elegans* N2 strain

In order to establish if there was any toxic or biological effects of EHL, *C. elegans* was treated with purified protein. Treating larvae in the first life stage (L1) allows later life effects to be characterised as part of the treatment. A range of phenotypic effects were observed when L1 stage larvae were incubated in EHL consistent with disruption to the chemosensory network. An assay to observe the effect of treatment on life time fecundity, mortality, and constitutive dauer formation was conducted.

5.1.3 Nematode Nanoparticle assays

Two furthers experiments investigating the effect of lectin conjugated gold nanoparticles (EHL+AuNP) were conducted. The first of these (NP1) was a replication of the other experimental work presented and published in McConnell *et al.* (2015), which was carried

out and reported the induction of the constitutive dauer phenotype in lectin treated L1 *C. elegans*. The second experiment (NP2) was carried out using L4 stage (pre-adult) worms and had not previously been conducted with EHL. An observation that the EHL+AuNP sample did not agglutinate erythrocytes; a characteristic of the native protein, was recorded prior to commencement of the experiments.

A calibration curve was established during this work examining the effect of gradient %v/v EHL solutions on N2 L1 larvae and the daf-c phenotype. The results showed that at a concentration of 2.3 mg/ml, an 80% v/v solution is enough to induce approximately 60% dauer formation. Above this lethal effects tend to take over, and below this the dauer formation is reduced and more delayed developmental effects are apparent.

To investigate the mechanism of action of EHL on *C. elegans*, a number of mutant isolates were chosen on the basis of the mutations that affect neuronal function. Strains which carry the dyf-d mutation (dye filling defective) were treated in order to investigate whether blockage of the proposed amphid neuron uptake of EHL would block the daf-C phenotype previously observed in N2. In dyf-d mutants the dauer effect was indeed found to be blocked. It would also be of interest to determine if those individuals that remained as arrested L1s in these experiments were doing so as a consequence of an inability to perceive the food, or if an additional mechanism is at work.

5.2 Materials and methods

5.2.1 Nematode assay

Worms were obtained from the *Caenorhabditis* Genetics Center and maintained using standard methods (Stiernagle, 2006) on nematode growth media plates (NGM) using an

Escherichia coli OP50 strain food source. *C. elegans* strain N2 was used for initial testing and as a control in other assays. To assess the effect of various mutations on EHL-induced dauer larvae formation, the following strains were used: CX2065, *odr-1(n1936)*; CX2205 *odr-3(n2150)*; PR671, *tax-2(p671)*; PR672, *che-1(p672)*; PR813, *osm-5(p813)*; SP1205, *dyf-1(mn335)*; and SP1709, *dyf-10(e1383)*. In all experiments, treatments and genotypes were blind coded, the position of plates within experimental blocks was randomised, and any contaminated plates displaying evidence of fungal growth excluded from all analysis.

For all assays, arrested and synchronised *C. elegans* first stage larvae (L1s) were obtained by allowing eggs, isolated from gravid hermaphrodites by hypochlorite treatment (Stiernagle, 2006), to hatch on NGM plates in the absence of food for 24 hours at 20°C. For experiment 1, arrested N2 L1s were washed from plates, resuspended in sterile M9 nutrient media (3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl, 1 ml of 1M $MgSO_4$, made up to to 1 litre with dH_20). with a series of EHL concentrations from 3.92 to 0 mg/ml and incubated at 20°C for 6 hours. Fifteen worms per treatment picked for analysis. For experiment 2, arrested N2 L1s were treated as above except treatments were 2.94 mg/ml and 0 mg/ml EHL and a greater number of worms per treatment were analysed (n = 55 and n = 33 for the 2.94 and 0 mg/ml treatments, respectively). After incubation, worms were washed 3 times in sterile water. For the analysis of development and fecundity (experiments 1 and 2), worms were transferred in a small volume of liquid to NGM plates without food, then individually transferred from this plate to NGM plates with an *E. coli* OP50 strain food source and maintained at 20°C. Standard methods were then used to analyse the reproductive schedule and lifetime fecundity (Hodgkin and Doniach, 1997). These data were then used to assess the effect of EHL on reproduction as assessed by lifetime reproductive success

(LRS), the total number of progeny produced (experiments 1 and 2), and the intrinsic rate of increase (*r*) (experiment 2). This was calculated by iteration from $\Sigma e^{-rx} l_x m_x = 1$, where l_x represents the age specific survivorship to day *x* and m_x represents the fecundity on day *x* (Vassilieva and Lynch, 1999).

Based on phenotypes observed in the initial screen, the ability of EHL to induce constitutive dauer larvae formation - a dauer-constitutive, or Daf-c, phenotype was investigated in greater detail. Here, worms were treated with 0.98, 1.96 or 2.94 mg/ml of EHL as described above, except that after washing, worms were transferred *en masse* to plates with food. After four days at 20°C, plates were visually scored to assess the proportion of worms that had developed as dauer larvae (number of dauer larvae/total number of worms). After counting, worms were washed from plates and incubated in 1 % SDS for one hour, a treatment that kills all *C. elegans* stages except dauer larvae (Cassada and Russell, 1975), worms were washed once in M9, transferred to fresh NGM plates with food and the number of dauer larvae again counted. These dauer larvae were then transferred individually to NGM plates with food at 20°C and monitored for the next 14 days to determine if they were capable of resuming development. Under these conditions 100% of dauer larvae that form due to lack of food and over-crowding would be expected to recover and to have molted to adulthood after approximately two days (Green and Harvey 2012).

To further analyse EHL-induced dauer larvae formation, N2 and mutant worms (see above for line details) were treated, as described above, with 0, or 1.54 mg/ml EHL, washed, and transferred *en masse* to plates (n = 3 per combination of treatment and genotype) with

food. Plates were then incubated at 20°C for four days, at which time the proportion of worms that had developed as dauer larvae was scored.

Dauer larvae formation in PR672, *che-1(p672)*, was further analysed both in standard dauer larvae formation assays and in assays of growing populations. For assays of dauer larvae formation in response to defined amounts of pheromone, assays were performed as previously described (Golden and Riddle 1984; Green *et al.*, 2014), with worms allowed to lay eggs on assay plates containing dauer pheromone extract and limited amounts of food, and progeny were scored after two days at 25°C. Dauer larvae formation in growing populations was assessed as previously described (Green *et al.*, 2013), with populations initiated with single worms and a defined amount of food being allowed to grow to food exhaustion, except that assays were performed at 25°C.

5.2.2 Nematode Nanoparticle assays

Experiment NP1: A population of N2 *C. elegans* were synchronized using a sodium hyperchlorate solution and the resulting eggs were placed on NGM agar plates without food for 18 hours. The resulting arrested L1 larvae were incubated in 15 ml eppendorf tubes at 20ºC in a solution of one of four treatments (see below) for 6 hours. 3 replicates of each treatment were made.

After incubation, all treatments were subjected to a cycle of 3 washes with dH_2O with a 2 minute centrifugal spin at 2000G. Worms were then added to seeded (*Escherichia coli* OP50) NGM plates and incubated at 20ºC. Plates were then scored on day 3 for survival, arrested development and for dauer larvae formation.

Experiment NP2: Using the worms from the experiment 1 population, eggs were placed on seeded NGM plates and incubated at 20ºC until L4 stage was reached. Treatment was

then carried out in the same set of liquid conditions as experiment 1 for 18 hours. Tubes were placed in a shaking incubator at 20ºC overnight. Worms were washed 3 times and moved onto a single plate. L4 stage worms were then placed individually onto seeded NGM plates, 50 per treatment. Fecundity was measured. Treatments were as follows:

1. M9 liquid nutrient media and EHL @ [1.51 mg/ml]
2. M9 liquid nutrient media and EHL+AuNP
3. M9 liquid nutrient media and AuNP
4. M9 liquid nutrient media

5.3 Results

5.3.1 Activity agains C.elegans N2 and mutant strains

All statistical analyses were conducted in 'R'. Acute treatment of arrested *C. elegans* L1s for 6 hours with EHL at different concentrations resulted in a subsequent range of developmental, fertility and survival defects (Figure. 5.2), with all EHL concentrations reducing LRS (Figure 5.2A, pairwise, Bonferroni corrected, Mann–Whitney U tests against N2 showing reduced fecundity in all EHL treatments). Much of this decrease in fecundity is however a consequence of worms not reproducing in the EHL treatments (comparison of Figure 5.2A and B), although LRS does still decrease over the range of EHL concentrations tested when only those worms that reproduced are considered (Figure 5.2B, Pearson product-moment correlation of LRS against EHL concentration: including the 0 mg/ml group, $r = -0.52$, $p < 0.001$; excluding the 0 mg/ml group, $r = -0.38$, $p = 0.013$). EHL-treated worms that did reproduce showed a delay in development, with many treated worms starting reproduction a day or more after the control worms. Of the worms that did not reproduce, some showed no movement from the point at which they were placed on the

plate, and no response to stimulus after 24 hours, they therefore died as a consequence of the EHL treatment. Other EHL-treated worms were observed to remain as arrested L1s or to develop as dauer larvae, a non-feeding developmentally arrested stage. Many of the non-reproducing worms were found not on a food source, and more than would be expected under these conditions were found to have climbed the sides of the plate; both behaviours are indicative of a disruption to chemosensory ability and an inability to detect the bacterial food.

To investigate these reproductive effects and the survival in more detail, a larger number of worms were assayed (Figure 5.3). Here, EHL treatment resulted in immediate mortality of 41% of EHL treated individuals and again many EHL-treated worms were observed to remain as L1s and to arrest as dauer larvae.

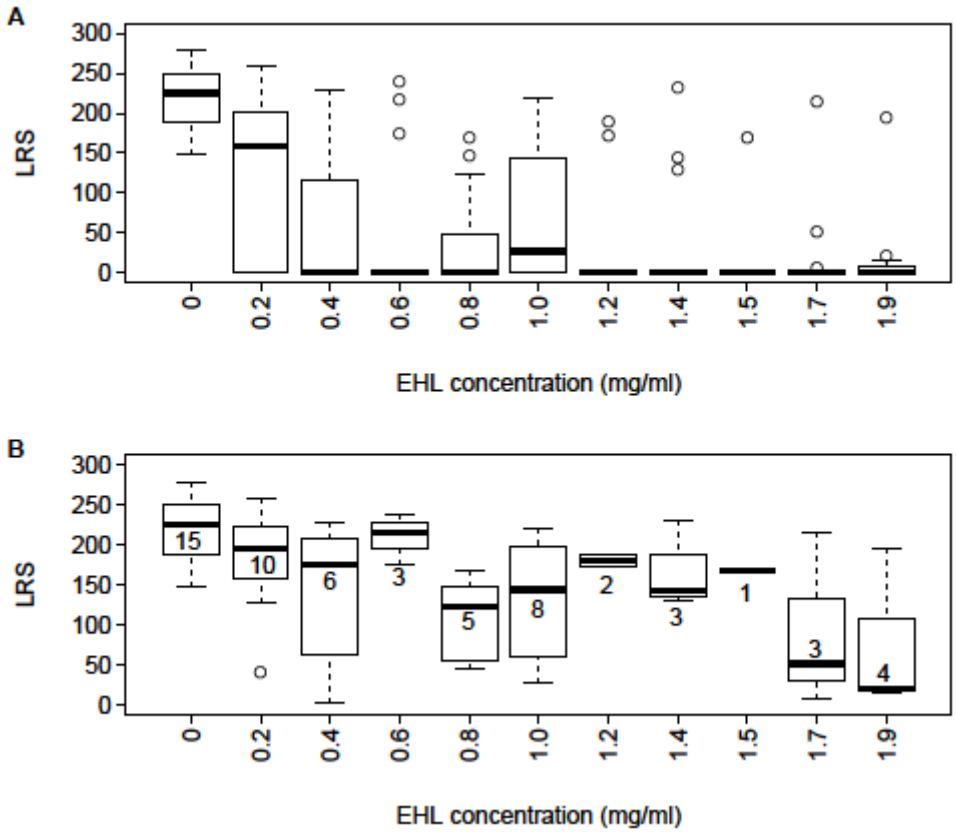


Figure 5.2 EHL reduces fecundity in *C. elegans*. Box plots of lifetime reproductive success (LRS) of (A) EHL treated worms and (B) the subset of EHL treated worms that reproduced for a range of lectin concentrations. For (A), n = 15 for all treatments, for (B) numbers associated with the boxes denote the sample sizes.

These results suggest that EHL treatment affects the sensory neurons. Overall, EHL treatment reduced subsequent LRS (Figure 5.3A, control vs all EHL treated worms, W = 2376.0, p <0.001), with the subset of EHL treated worms that did reproduce producing a greatly reduced number of progeny (control vs reproducing EHL treated worms, W = 792.0, p <0.001). A similar pattern was observed in the analysis of the effects of EHL treatment on

the estimated rate of increase (Figure 5.3B, control vs all EHL treated worms, W = 2376.0, p <0.001; control vs reproducing EHL treated worms, W = 792.0, p<0.001).

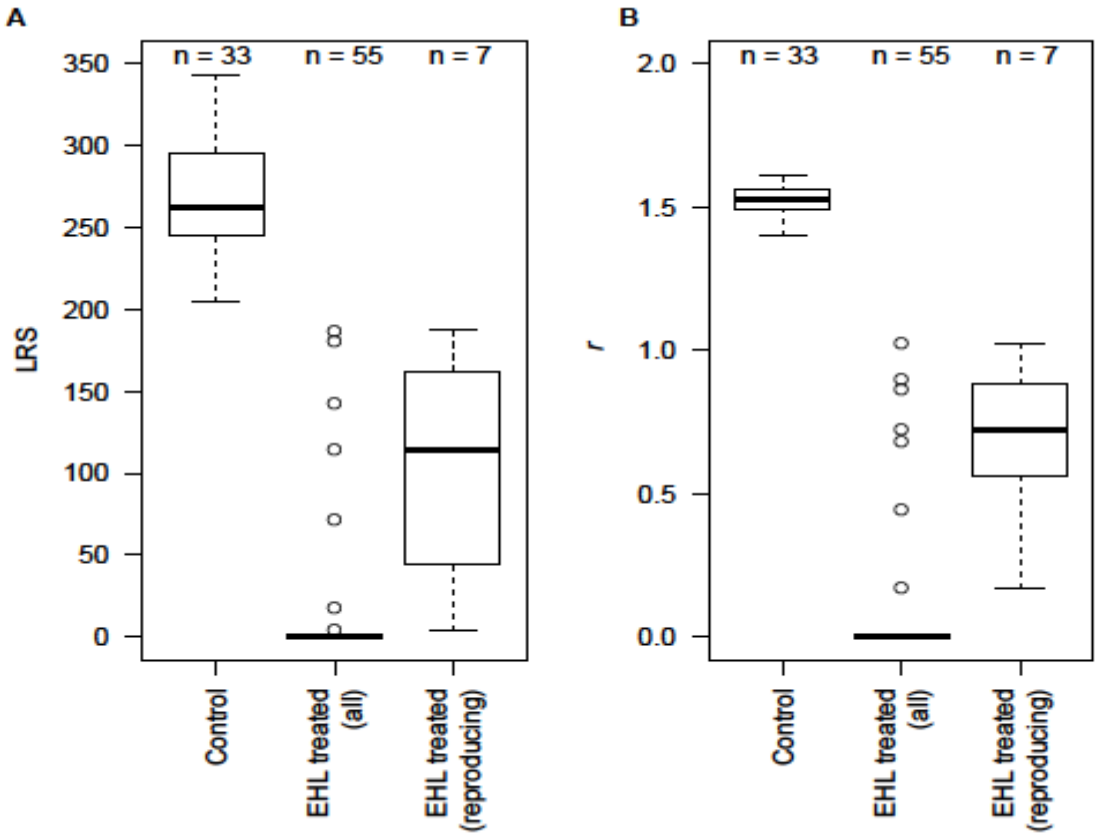


Figure 5.3 EHL reduces fecundity and slows development in *C. elegans*. Boxplots of (A) lifetime reproductive success, LRS, and (B) intrinsic rate of increase, *r*, of control worms, all EHL treated worms and of the subset of EHL treated worms that reproduced. *n* denotes the sample sizes.

To characterise further the development of EHL-treated worms as dauer larvae, an additional population of arrested L1s were treated and analysed. As in the assay for reproductive effects (above), some (50) of the population of these EHL treated worms developed as dauer larvae (Figure 5.4). These worms were then SDS treated, with survival confirming that they were indeed dauer larvae. When transferred to food and

subsequently maintained at 20°C with food, only one worm out of the fifty in the assay had recovered and completed development as a reproductive adult after four days, with another worm recovering after a total of fourteen days. Under these conditions dauer larvae would normally recover rapidly and have commenced reproduction approximately 2 days after transfer to food (Green and Harvey, 2012).

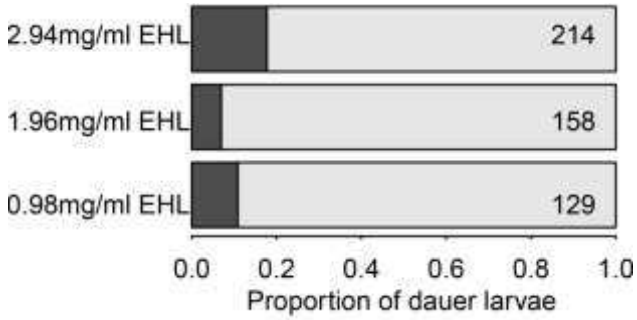


Figure 5.4 EHL treatment induces dauer larvae formation in *C. elegans.* The proportion of larvae that developed into dauer larvae observed in populations in three different EHL treatments (dauer is shaded). In the absence of EHL treatment no dauer formation is observed under these experimental conditions in N2.

Analysis of mutant strains indicates that the ability of EHL treatment to induce dauer larvae formation varied across the genotypes (Table 5.1) However, all genotypes were observed to show a developmental delay in response to EHL treatment, with reproduction of most EHL-treated worms not commencing until day 5 after treatment. Analysis of dauer larvae formation in PR672, *che-1(p672)* showed that this mutation does not block dauer larvae formation in response to defined amounts of pheromone (data not shown) and that similar numbers of dauer larvae are formed in N2 and PR672 in growing populations ($F_{1,33} = 0.05$, $p = 0.82$), *i.e.* this line is not dauer defective.

		Mean % dauer ± S.E.		
Isolate	Genotype	Control	Lectin	Phenotype
N2	Wild-type	0	49 ± 2.5	-
CX2065	*odr-1(n1936) X*	0	53 ± 6.1	Defective chemotaxis[1]
CX2205	*odr-3(n2150) V*	0	35 ± 3.1	Defective chemotaxis[1] Defective osmotic avoidance[1]
PR671	*tax-2(p671) I*	0	8 ± 1.9	Defective chemotaxis[2] Defective thermotaxis[2]
PR672	*che-1(p672) I*	0	0	Defective chemotaxis[2]
PR813	*osm-5(p813) X*	0	0	Defective dye filling[3] Defective osmotic avoidance[3]
SP1205	*dyf-1(mn335) I*	0	0	Defective dye filling[4] Defective chemotaxis[4]
SP1709	*dyf-10(e1383) I*	0	0	Defective dye filling[4] Defective chemotaxis[4]

Table 5.1 Mutations affecting amphid structure can block EHL-induced dauer larvae formation. The mean dauer larvae formation found on three plates of control and EHL treated worms of differing genotypes of *C. elegans*. [1] Bargmann, Hartweig and Horvitz (1993). [2] Dusenbery (1976). [3] Dusenbery (1980). [4] Starich *et al.* (1995).

5.3.2 Nematode nanoparticle assay

In experiment NP1, the previously reported lectin effect (McConnell *et al.*, 2015) was demonstrated in the EHL treated worms, however dauer formation was not observed in the EHL+AuNP treatment. No other significant effects were reported in any of the other treatments. In conjunction with the empirical observation that agglutination properties were absent from the EHL+AuNP sample, this would indicate that a conformational change

has occurred potentially induced by conjugation to AuNP, which blocks the EHL neuronal binding effects.

In experiment NP2, testing with the L4 stage worms, there was a very small effect on lifetime fecundity in the EHL treated worms, and a delay in reproduction in both the EHL treated and the EHL+AuNP worms. From this the conclusion is that naked NPs do not obviously affect the L4s and that the EHL+AuNPs do have a biological activity.

For the treated L1s, there was no dauer larvae formation in response to nanoparticles (NP) or to EHL conjugated nanoparticles and none in the control, but there is dauer larvae and developmental arrest in response to EHL treatment (Table 5.2). EHL also kills L1s, whilst the other treatments do not differ in L1 survival (Table 5.2). From this, the conclusion is that the assay shows that naked NPs do not obviously affect the L1s and that the EHL conjugated nanoparticles do not replicate the EHL effect.

For the treated L4s, there was a difference between treatments in lifetime fecundity (TREATMENT: $F_{3,68} = 2.76$, $p = 0.049$) that is a consequence of a small reduction in lifetime fecundity in the EHL treated worms (EHL treatment significantly different to the control and nanoparticle-treated worms by Fisher's post-hoc testing) (Figure 5.5, panel A). There is also variation between treatments in early reproduction ($H = 14.22$, $df = 3$, $p = 0.003$), with both the EHL treated and the EHL conjugated nanoparticle treated worms showing a reduced early fecundity ($p < 0.05$ in comparison to control worms via Mann-Whitney test) (Figure 5.5, panel B). From this it is concluded that the nanoparticles do not obviously affect

the L4s and that the EHL conjugated nanoparticles do produce a similar biological activity to that produced by EHL.

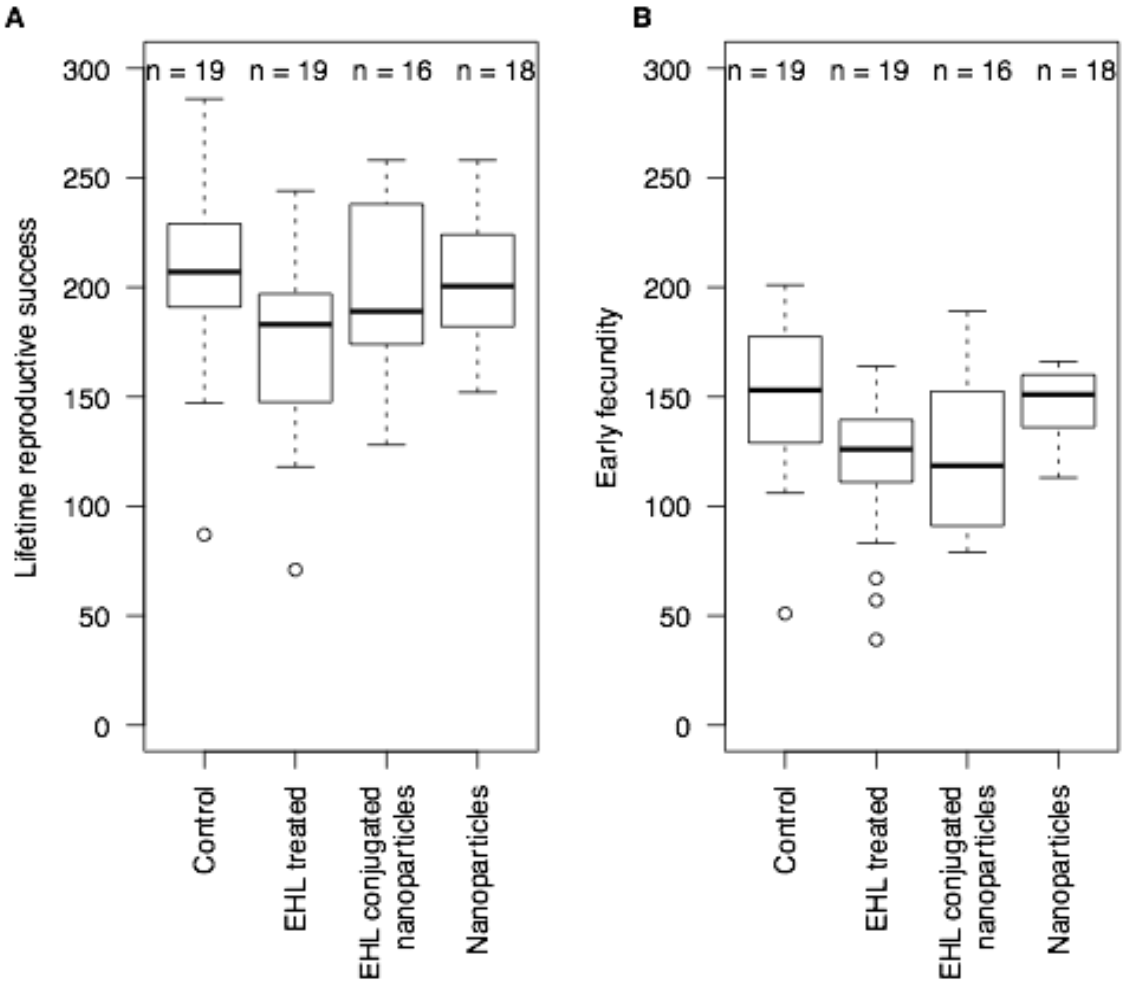


Figure 5.5 EHL conjugated nanoparticles affect early reproduction, but not total reproduction, of *C. elegans* L4s.

Treatment	Number of plates	L1s per plate	Mean % survival (min and max)	Mean % dauer formation (min and max)
Control	11	54.6 ± 3.8	68 (53-81)	0
EHL	12	64.2 ± 7.0	23 (11-40)	24% (0-45)
nanoparticle	11	58.6 ± 3.2	73 (54-84)	0
EHL conjugated nanoparticle	12	48.7 ± 6.6	68 (49-83)	0

Table 5.2 EHL treatment affects survival and development of treated *C. elegans* L1s. When scored, adult worms were only observed on 3 of the 12 EHL plates, with worms showing varied degrees of developmental delay.

5.4 Discussion

EHL was used to study potential lectin-mediated toxicity against *C. elegans.* The bioassays performed indicate that EHL has biocidal properties against *C. elegans*. Four phenotypic effects were identified: reduced fecundity (Figure 5.2 and 5.3), developmental delay, chemosensory disruption and constitutive dauer formation (Figure 5.4 and Table 5.1).

C. elegans physiology is such that at the arrested L1 larval stage, the only cells which are not enclosed by a largely impermeable cuticle are the amphids and phasmids. These are bilaterally symmetrical sensory organs that contain the sensory neurons: each amphid containing twelve neurons and each phasmid containing two (Ward *et al.*, 1975). Of the twelve amphid neurons, the ciliated nerve endings of eight are exposed to the external environment via the amphid pore (Ward *et al.*, 1975). These neurons control a range of

phenotypes, including egg-laying and the decision to develop as a dauer larvae (Albert, Brown & Riddle, 1981). Laser ablation of the ASI, ADF and ASG cells is sufficient to result in constitutive dauer larvae development, with ablation of the ASJ cell resulting in an inability to recover from dauer arrest (Bargmann & Horvitz, 1991). The observations of inappropriate dauer larvae formation and the failure of most such dauer larvae to resume development suggest that EHL is interacting with these neurons and are consistent with EHL treatment resulting in neuron death.

In general, biocidal assays with lectins involve ingestion of the lectin by the target organism. For example, EHL had previously been tested against the coleopteran pest *Diabrotica undecimpunctata howardii,* resulting in a high mortality rate and an 80% reduction in body size of survivors; there were however no previous data on reproductive effects (Kumar *et al.,* 1993). A study of the toxic effects of the CCL2 lectin from *Coprnopsis cinerea* (Ink Cap mushroom) on *C. elegans* reported a phenotype of severe developmental delay when the lectin was transgenically produced by bacteria fed to the worms; the lectin was adsorbed in the epithelial cells of the intestine, potentially degrading the membrane and preventing growth (Schubert *et al.*, 2012).

The absence of a food source in the assays reported here has enabled the observation of entirely new lectin-mediated *C. elegans* phenotypes induced by EHL, including a Daf-c phenotype which has not been previously reported.

The cause of the developmental delay in EHL-treated worms is not clear. Possibilities would include damage to the pharynx and a subsequent reduction in pumping (feeding) ability, or, if some feeding is initiated, damage to the epithelial cells of the intestine as observed in response to the *C. cinerea* CCL2 lectin (Schubert *et al.*, 2012). A further possibility is that

differences in body size and development are also a consequence of damaged neurons (Fujiwara, Sengupta and McIntire, 2002).

Wild-type *C. elegans* take up dyes such as DiI and FITC into the amphid neurons AWB, ASH, ASJ, ASK, ADL and ASI (Hedgecock *et al.*, 1985). Given the likely mode of action of EHL, it is reasoned that mutations that disrupt the normal formation of sensory amphids would block EHL-induced dauer larvae formation. Consistent with this, disruption of *osm-5*, *dyf-1* and *dyf-10*, all mutations in which the amphid neurons cannot take up dyes, result in no EHL-induced dauer larvae formation (Table 5.1). In both *odr-1* and *odr-3* mutants, where dye filling is not affected and EHL would be expected to be able to access the neurons normally, there is no reduction in dauer larvae formation in response to EHL treatment (Table 5.1). In contrast, dauer larvae formation is reduced in PR671, but some are still formed (Table 5.1), indicating that disruption of *tax-2* only partially blocks the effect. TAX-2 forms, with TAX-4, a cyclic nucleotide-gated cation channel that is required for chemotaxis in response to AWC-sensed odorants (Coburn and Bargman, 1996). Axon outgrowth defects have however been noted in *tax-2* mutants, with c. 80% of *tax-2(p671)* animals observed to have abnormal ASJ axons (Coburn and Bargman, 1996). Given that the reduction in the EHL-induced dauer larvae formation observed in the *tax-2* mutants appears to mirror this, therefor it can be inferred that the reduction in dauer larvae formation is a consequence of the axon guidance defects rather than the channel disruption.

That EHL-induced dauer larvae formation is also blocked in *che-1* mutants further supports the hypothesis that EHL is disrupting neurons. CHE-1 is a C2H2-type zinc-finger transcription factor that is required for the identity of ASE neurons (Uchida *et al.*, 2003).

Loss of CHE-1 expression eliminates the function of ASE neurons and *che-1* mutations have previously been shown to suppress Daf-c phenotypes (Reiner *et al.* 2008). No significant structural defects have been observed in *che-1* mutants (Lewis and Hodgkin, 1977) and the results indicate that dauer larvae formation does not appear altered in either standard dauer larvae assays or in growing populations. This implies that the mutation is specifically blocking EHL-induced dauer larvae formation.

It is well established that lectins bind to glycoconjugates on cell surfaces and that toxicity in Type II RIPs is initially due to lectin mediated entry to the cell; this mode of action is consistent with the results presented here. In the case of *C. elegans* the only cells exposed are the amphid neurons. As a Type II RIP EHL can be subject to retrograde transport from the cell surface along the neuronal processes, at which point the ribosomes are inactivated, causing translation to cease (Wiley, Blessing and Reis, 1982). As no post-embryonic somatic division occurs in mature individuals, and multiple chemoreceptors are expressed in a single neuron, ribosome inactivation of the neurons within the amphids would affect many functions derived from chemosensation (Sulston and Horvitz, 1977). Toxicity variables can be attributed to differing carbohydrate specificities but there is also evidence of the role of individual cell types in how they interact with lectins, indicating that any effects are characteristic of both variables (Battelli *et al*., 1997).

Successful extraction using affinity chromatography has enabled assays to be conducted for biocidal properties against *C. elegans.* The results obtained demonstrate a significant reduction in fecundity, development, growth and a high incidence of abnormal dauer development when arrested L1 larvae were treated in the absence of food. The occurrence of dauer formation and a failure to recover in the presence of food would support the

hypothesis that EHL is binding specifically to amphid neurons. Here, mutant screening has demonstrated that EHL can act as a neuronally specific cytotoxin against sensory neurons. Further studies will aim to determine if those individuals that remained as arrested L1s were doing so as consequence of an inability to perceive food or if an additional mechanism is at work. This research shows that EHL has biocidal and potential cytotoxic activity. The successful conjugation of EHL to gold nanoparticles opens many avenues for further research, as a preliminary investigation, this work has shown that the activity of EHL is altered by this conjugation and as such resulted in a lessened biological effect. As lectin binding activity is essential for the natural protein to bind and allow entry to cells, conformational changes may have affected this binding affinity. The results show that EHL conjugated gold nanoparticles do still retain some activity in L4 assays as seen in Figure 5.5, suggesting that ingestion in the absence of glyconjugate binding (which is absent in the L1 assay) of the molecule may present a low level of toxicity. The observation that EH+AuNPs do not agglutinate erythrocytes (cross linkage of carbohydrate surface antigen) would suggest that this is a factor. Of course the inverse hypothesis may also be worthy of investigation, that the toxic A-chain activity may have been altered and that binding still occurs without the cytotoxic effects seen in the intact molecule. As non-RIP lectins have been shown to bind to epithelial cells in the gut causing reduced fitness (Stutz *et al.,* 2015) this requires further study to establish the exact reason for the reduction in toxicity. Since EHL shows specificity for N-acetyl-galactosamine (Cammue, Peeters and Peuman, 1985; Kumar *et al.,* 1993), an overexpressed and incompletely glycosylated sugar in the Tn antigen which characterizes cancer linked O-glycans (Ju, Otto and Cummings, 2011) and other GalNAc specific RIPS, such as Mistletoe lectin and Riproximin, have demonstrated

promising therapeutic relevance as anticancer agents (Voss *et al.*, 2006; Bayer *et al.*, 2012; Adwan *et al.*, 2014), EHL is viable as a candidate for further onwards study of antineoplastic characteristics.

Chapter 6: Development of a methodology for investigation of the effect of EHL on mammalian cell lines.

6.1 Introduction

Due to the cytotoxic nature and selective binding properties of ribosome inactivating proteins, they have been considered as potential agents in cancer therapies. Immunotoxins and other types of conjugates have been prepared, targeting a variety of different tumour specific characteristics such as neoplasia related overexpression of cell surface receptors and angiogenic markers (Hanahan and Weinberg, 2011). Of particular note is work involving the non-toxic RIPs such as Ebulin and Nigrin. The narrow glycan binding profile of these RIPs lessens the occurrence of non-specific binding, as found with the broad range highly toxic Ricin and Abrin (Ferreras *et al.* 2011), despite sharing the identical functional enzymatic mechanism of depurinating A_{4234} in the 28S ribosomal rRNA of the 60S sub-unit of eukaryotic cells (Van Damme *et al.*, 2001).

Gold Nanoparticles (AuNP) have also been shown to have potential therapeutic applications in oncology research (Kodiha *et al.* 2015). Studies shows that cancer cells appear more vulnerable to AuNP uptake than healthy cells (Ahmad *et al.*, 2013) even without functionalised therapeutic agents attached (Ahmad *et al.* 2013). AuNPs can be functionalised with conjugated antibodies, drug preparations or imaging agents and dependent on the cell type and NP morphology, show an increase in permeability across the plasma membrane of cancer cells especially to the nuclear compartment (Kodiha *et al.*, 2015). This allows tumor cells to be targeted, with little effect on local healthy cells. The main advantage to using AuNPs is that they can simultaneously carry both functional

therapeutics and imaging agents to the cell. Due to the specific targeting, possible toxic systemic levels of chemotherapeutics could be lowered, efficacy of drugs improved and localised concentration in the tumour cells increased (Patra *et al.*, 2010). As they themselves have a low inherent toxicity, and designable surface chemistry, AuNPs hold vast potential for targeting specific tumour antigens and infiltrating angiogenic regions (Ahmad *et al.* 2013). To this aim a collaboration was set up with the Bioscope group who functionalised AuNP to EHL, in the first instance to investigate the viability of synthesising such a large conjugate. When successful functionalisation was complete, the AuNP+EHLs were returned for exploratory work with *C. elegans* (see Chapter Five) and mammalian cell lines.

The changes which occur in neoplastic diseases whereby a healthy cell becomes tumorogenic are often marked by epigenetic amendments by the normal cellular machinery, such as glycosylation patterns (Stowell, Ju and Cummings, 2015). Changes to glycosylation can arise as a result of changes in expression levels of enzymes involved in post translational glycan modification to peptides in normal biosynthesis. This can be as a result of genetic mutations, for instance the deletion of a gene involved in enzymatic activity such as *cosmc* which regulates the activity of T-synthase (Hofmann *et al.*, 2015). T-synthase is responsible for converting the incomplete core GalNAc residue Tn antigen to the terminal GlcNAc capped T antigen which is found on healthy cells (Stowell, Ju and Cummings, 2015). In contrast the Tn antigen is perceived to be a cancer biomarker whereby a positive correlation between antigenic abundance and the aggressiveness of the carcinoma is indicated (Madariaga *et al.,* 2014). Rarely occurring in healthy tissue, Tn antigen consists of a clustered sequence of GalNAc residues (Ju, Otto and Cummings, 2011)

and therefore could provide a target for a GalNAc specific carbohydrate binding protein such as EHL to differentiate between healthy and cancerous cells. These antigenic changes occur in disease progression and are indicators of metastatic potential. The major types of protein glycosylation associated with oncogenic changes to the cell are O-linked and N-linked glycosylation. The Tn antigen is an O-linked glycan with the full structure containing several N-acetyl-galactosomine units linked to a serine or threonine on a glycoprotein. It is synthesized as an incomplete product of the O-linked pathway. Lectins have been used to detect these biomarkers, with GalNAc specific lectins which share specificity with EHL such as Soya Bean Agglutinin and Vicia Villosa agglutinin being used in glycan recognition (Madariaga *et al.*, 2014).

Riproximin is an example of a novel type II RIP with antineoplastic properties which demonstrated binding to clustered GalNAc in carbohydrate microarrays. It was used to analyse binding to MDA-MB-231; notably desialylated for the investigation with an unspecified commercial neuraminidase (Roche Scientific) to expose the Tn core residues. (Bayer *et al.,* 2012). This is another avenue for further exploration with EHL which raises the question whether EHL will do the same. EHL, like riproximin, shares higher overall homology with toxic RIPs such as Mistletoe lectin 1, *Polygonatum multiflora* RIPm, ricin and abrin than to Ebulin or Nigrin (the non-toxic RIPs from *Sambucas* species). As it would appear that binding to glycans is influenced by the surrounding underlying peptide sequence and how that interacts with individual carbohydrate binding proteins (Stowell, Ju and Cummings, 2015), and that individual cell lines show diverse interactions, many areas for study present themselves.

Carbohydrates are one of the three main biological macromolecules essential for eukaryotic cell function and structure along with lipids and protein. Sugars are linked to a wide suite of functions in the cell, and are one of the biggest components involved in post translational peptide modifications. Constituent elements of carbon, hydrogen and oxygen combine to form a hydrated carbon compound. Mono saccharides are the smallest component which cannot be further broken down into a hydrated unit. Despite the simple general molecular formula $(CH_2O)_n$, (where n=3,5 or 6) for a mono saccharide, a vast array of structures can be formed through the combining of simple mono and disaccharide units to form poly or oligosaccharide structures. Further modifications such as methylation increase the number of potential carbohydrate structures to an almost infinite number (Stowell, Ju and Cummings, 2015). In certain cases, the clustering of saccharide units on a cell surface in close proximity can also give rise to unique glycan environments which act as a ligand for an equally vast array of substrates. In some cases these will be a lectin or Type II RIP. All cells display carbohydrate groups or glycans on their surface as with the blood group antigens previously discussed, and this is also the basis of the cross linkage due to carbohydrate binding seen in agglutination of erythrocytes (see figure 2.4) and other cells by lectins.

6.2 Materials and Methods

For the initial development of an EHL treatment protocol and analysis of diseased cells, the first experimental lines were chosen on the basis of ease of culture for a novice and with well defined characteristics (see Table 6.1).

Cell line	**A431**	**MDA-MB-468**	**MDA-MB-231**
Pathology	Epidermoid carcinoma	Adenocarcinoma	Adenocarcinoma
Tissue	skin	Breast	breast
Site	Solid tumour	Metastatic site	Metastatic site
Species	*Homo sapiens* 85 yr old Female	*Homo sapiens*	*Homo sapiens* 51 yr old Female
Notable features	p53 TSG missense single point residue mutation at R273 Highly mitogen sensitive	p53 TSG missense single point residue mutation at R273 mitogen sensitive	WNT7B oncogene EGF + TGFα +

Table 6.1 Summary of cell lines used and properties and characteristics of each.

The aim of the experiment was to establish a working protocol for future work. All cell lines were cultured in a 5% CO_2 environment in Dubelcco's modified Eagle media (DMEM) supplemented with 10% fetal calf serum, 1% pen-strep (Penicillin & Streptomycin) and 1% L-glutamine.

Two treatment methods were used in the first set of experiments with A431 cells. As this was a preliminary investigation, it was necessary to consider if the carbohydrate components of the media could competitively inhibit the lectin binding of EHL. Taking this into consideration a first assay was designed to remove excess culture media from the cells and after trypsinisation, the four treatments (AuNP, AuNP+EHL, EHL [2.5 mg/ml], and PBS (137mM NaCl, 2mM KCl, 10mM Na_2HPO_4, 1.7 mM KH_2PO_4) were applied directly to the

cells in 24 well plates. A short incubation followed, and after one hour the treated cells were split into two fractions. One fraction was used for direct visualisation of the cells on a haemocytometer and the other was subject to a cell lysis protocol and the total protein content analysed using SDS-PAGE (See Figure 6.1).

For both parts of the experiment, media was removed from culture flasks, and cells were washed twice with PBS prior to the addition of 1000 µl of trypsin. Cells were placed back in the incubator and checked regularly until they began to lift. The trypsin was then removed and replaced with 200 µl of fresh trypsin and left to completely lift from the flask. Further to this, 5ml of DMEM was added to the flask and 500 µl aliquots were placed in a 24 well plate (12 wells used), media was then added to a total volume of 2000 µl and cells were incubated overnight. All lectin treatments were prepared from a stock concentration of 2.50 mg/ml and PBS which consisted of 137mM NaCl, 2mM KCl, 10mM Na_2HPO_4, 1.7 mM KH_2PO_4.

For Experiment 1.1, when the appearance of an adherent monolayer of approximately 80% confluence was observed in the wells, media was removed and treatments of 250 µl were added. Treatments were added in a ratio of 1:1 with PBS and left to incubate for one hour. Treatment was then removed and wells were washed and trypsinised. For each treatment 50% was removed and treated with Trypan blue to show a contrast between living and dead cells. These samples were placed on a haemocytometer for quantification. The other 50% of the sample was subject to a lysis protocol using 1000 µl RIPA buffer (ThermoScientific) supplemented with 10 µl each of Halt Protease inhibitor cocktail (100x), 0.5 M EDTA (100x) (both from ThermoScientific) and Phosphatase inhibitors (Sigma Aldrich).

For experiment 1.2, standard culture methods were used to remove A431 cells from the culture flask and passage cells to a 24 well plate for overnight incubation as reported above. Excess media was removed and replaced with 1:1 treatment/media to a total volume of 500 µl. Cells were returned to the incubator overnight and checked 18 hours later. Treatments were then removed, and all wells washed with PBS and trypsinised for downstream analysis. The treatments were split as in experiment 1 with 250 µl used for lysis protocol (see Figure 6.1) and 250 µl used for direct imaging.

A second series of experiments with A431 and MDA-MB-468 cell lines were conducted. For experiment 2, the culture method, reagents and consumables were as previously described. In Experiment 2, A431, MDA-MB-468, MDA-MB-231 were cultured using standard methods. Due to technical limitations only A431 and MDA-MB-468 were subject to treatment. Treatment was applied for 2 hours and populations were used for direct imaging (see Figure 6.2) in order to examine the morphological changes observed in experiment 1.1 in more detail. This confirmed that imaging in this way was not an efficient method under lectin treatment.

As a final approach to optimize results, an Alamar Blue assay was added to the method. Cell density was first counted using the haemocytometer to make an approximation of cell number. Since the Alamar Blue standard protocol states that a wide range of cell densities can be used with a linear response seen between 50 and 50000 cells, dependent on cell type and treatment applied, a median range was selected in the region of 5000 cells. Cells were trypsinised and suspended in 1000 µl of media. Subsequently 10µl of cells was removed and mixed with 10µl of trypan blue and pipetted in equal volumes of 10µl to each side of the haemocytometer. Each of the 8 outer quadrants were counted and an average

value calculated; this was multiplied by 100 and used to approximate the number per µl of cells in the sample.

As a result of the cell count, a final volume of 10 µl each was added to a 96 well plate, which equated to approximately 5000 cells per well. Treatments were then added to the wells.

Measurements were taken using the FLUOstar Omega filter-based multi-mode microplate reader. Initially absorbance was read as for figure 6.3A, this was found to be too low resolution to quantify the changes. Subsequent cell lines (Figure 6.3B and 6.3C) were analysed by fluorescence measurement due to increased sensitivity and higher resolution. All analyses of absorbance and fluorescence measurements were conducted in Minitab.

6.3 Results

For technical reasons, the Trypan blue staining could not be imaged in either of the initial experiments (1.1 and 1.2), all effects can be reported but not evidenced at this stage. The different effects between treatments was of sufficient magnitude that visual inspection gave clear outcomes. In Experiment 1.1, the EHL treatment appeared to have killed A431 cells outright at the concentration of 1.25 mg/ml. The AuNP+EHL treatment showed an approximate 1:1 living to dead cell ratio. However of note was that the EHL treated cells had clumped together, whereas the AuNP+EHL cells had not. The PBS and naked NP treatments showed no visually observable effect.

Analysis of the lysis product in lanes 1-4, as seen in figure 6.1 was initially inconclusive and inconsistent. However with later results in the Alamar Blue assay hinting towards mitogenic stimulus of A431 by EHL, the outcome is more complex than first perceived. A second gel produced a comparable outcome. Experiment 1.2, which was conducted in the presence of culture media in the sample and was incubated overnight, showed a much more

distinctive suite of effects in the visual inspection. An increase in size of the AuNP+EHL cells with a distinctive internal granulated appearance could be observed, as well as severe clumping and disruption to the cell membranes of the EHL treatments. This was supported by the results observed on the SDS-PAGE analysis, with a successive reduction in overall protein content in the AuNP+EHL and EHL treatments (lane 7 & 8) compared to the naked NP and PBS (lane 6 &9) (Figure 6.1).The second set of experiments resulted in similar issues of agglutination of cells, as can be seen in Figure 6.2 which shows the different concentrations of EHL applied and the AuNP treatments. Whilst clumping can be a characteristic of certain cell lines, it would appear that the effect in this experiment was absent from control treatments and can therefore reasonably be assumed to be as a consequence of the treatment.

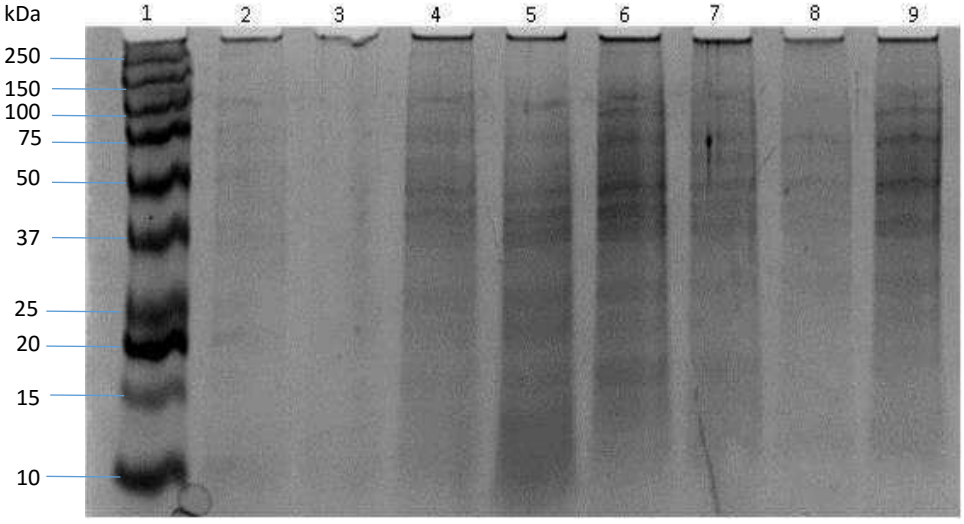


Figure 6.1: 12% SDS-PAGE analysis of whole cell lysate under reducing conditions on cell line A431: Lanes: 1. MM, 2. AuNP Direct treatment(DT), 3.AuNP+EHL DT, 4.EHL DT, 5.PBS DT, 6.AuNP 24 hr 7.AuNP+EHL 24 hr 8.EHL

A431 and PBS treatment

AuNP

EHL 1:10000

AuNP + EHL

EHL 1:1000

EHL 1:1

Figure 6.2 Experiment 2 EHL and EHL+AuNP treatment of A431. Images show the limitation of using the haemocytometer, as EHL treatment appears to cause clumping of the cells (consistent with glycan binding).

Figure 6.3. Individual data plots for Alamar Blue assay with MDA-MB-468, MDA-MB-231 and A431 cell lines. **A**) First trial using Alamar Blue & MDA-MB-468, *t*=18 hours, cell density approx. 5000 per well *n*= 12, Absorbance at λ570nm. **B**) Second assay using MDA-MB-231, *t*=26 hours *n=8* approx. 5000 per well. Fluorescence measured at Exc λ560nm & Em λ590nm. **C**)Third assay using A431 *t*=26 hours *n=8* approx. 5000 per well. Fluorescence measured at Exc λ560nm & Em λ590nm.

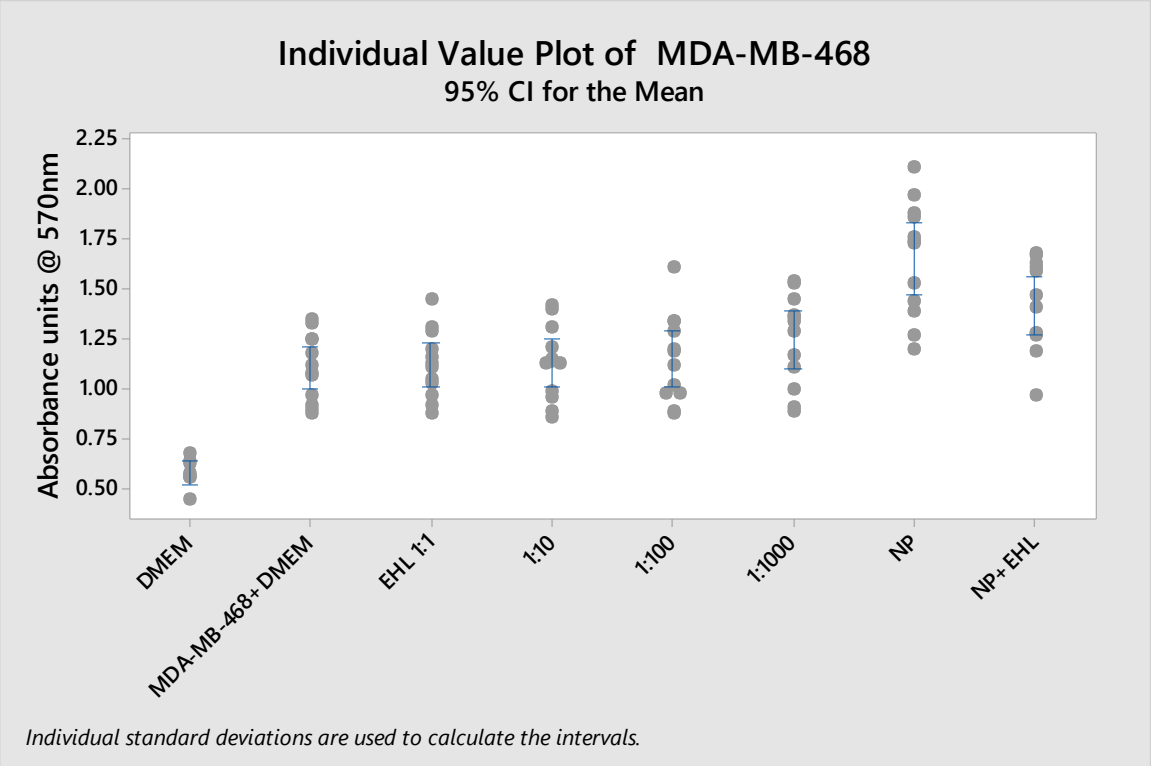


Figure 6.3 A) MDA-MB-468 cells absorbance readings demonstrate that the reagents are being reduced within quantifiable range as seen the positive control vs negative. Kruskal Wallis testing indicates that there is a difference between treatments (Treatment: H = 51.42, DF = 8, P = 0.000) in the amount of cell proliferation indicated by metabolic reduction of Rezasurin at t=18 n=95. Post hoc testing indicates that EHL treatments vs + control did not significantly differ judged from the p-value (1:1 p = 0.9081; 1:10 p = 0.7508, 1:100 p = 0.6236 1:1000 p = 0.0999). AuNP vs AuNp+EHL post hoc (Mann Whitney) testing indicates no significant difference between treatments at p = 0.0531.

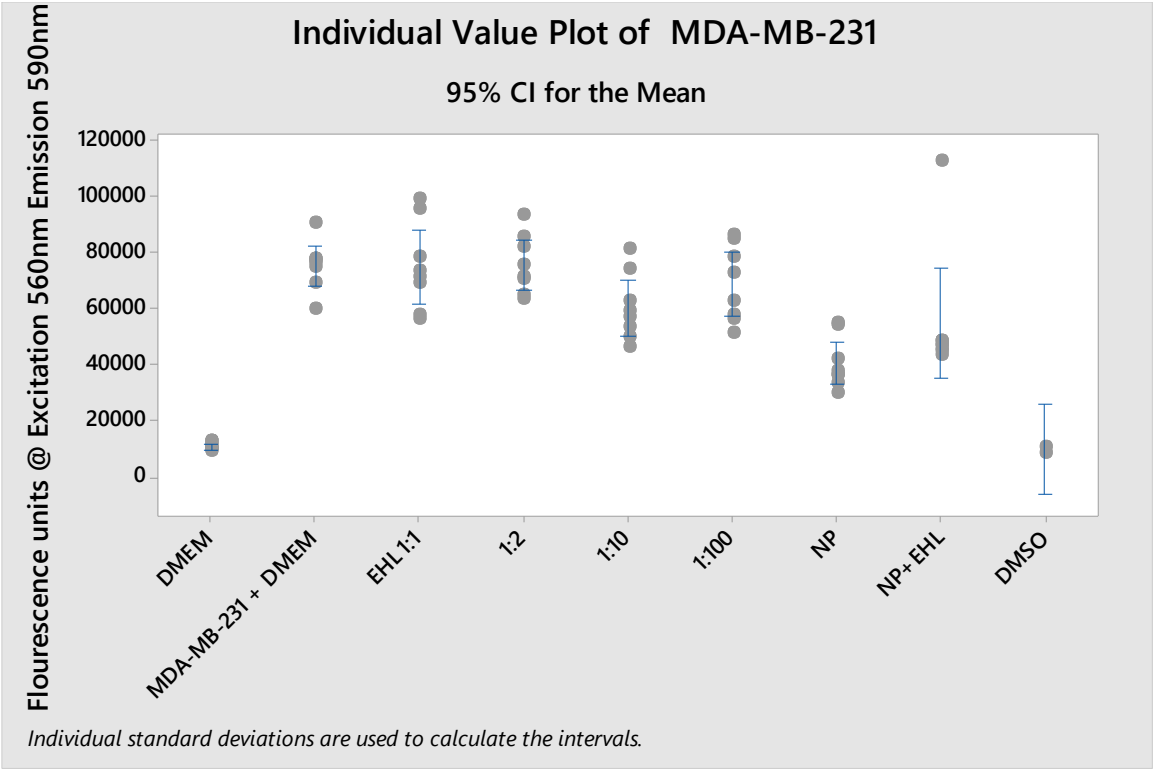


Figure 6.3 B) MDA-MB-231. Results of changing to an assay measuring fluorescence rather than absorbance. For the MDA-MB-231 cells, Kruskal Wallis non parametric testing indicates that there is a no difference between treatments except +control vs –control. (Treatment: H = 46.36, DF = 8, P = 0.000) in the amount of cell proliferation indicated by metabolic reduction of Rezasurin at t=26 n=66. Post hoc testing indicates that EHL treatments vs + control were not significantly different, judging by the p-Value (1:1 p = 0.7132; 1:2 p = 0.9581, 1:100 p = 0.4309) with the 1:10 treatment significant at p = 0.0239. AuNP and AuNP+EHL post hoc (Mann Whitney) testing indicating no difference between treatments at p = 0.0661.

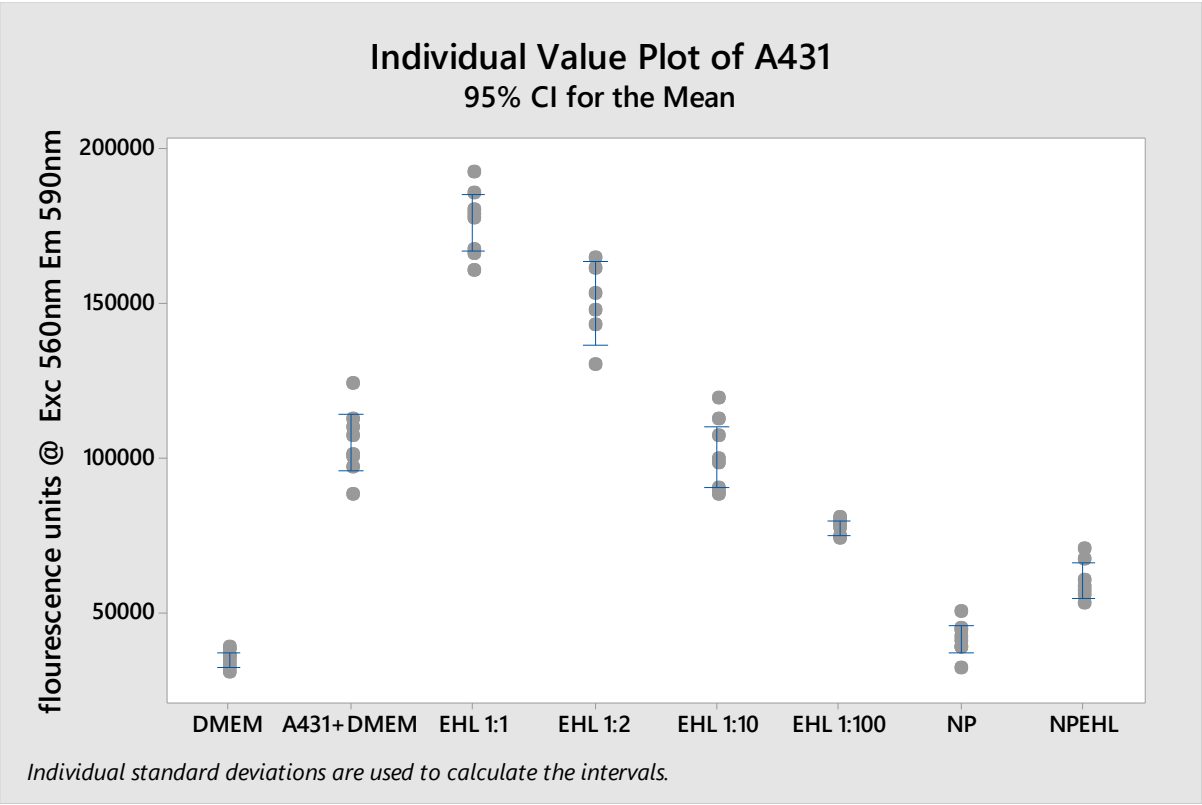


Figure 6.3 C) A431. With the assay working reliably, the results demonstrate increased proliferation of EHL treated sample at higher concentrations and a reduced effect as dose lowers. For the A431 cells, Kruskal Wallis testing indicates that there is a difference between treatments (Treatment: H = 58.67, DF = 8, P = 0.000) in the amount of cell proliferation indicated by metabolic reduction of Rezasurin at t=26 n=65. *Post hoc* testing indicates that EHL treatments vs + control differed, judging by the p-Values (1:1 p = 0.0009; 1:2 p = 0.0024, 1:100 p = 0.0015) with the 1:10 treatment not significant at p = 0.4948. AuNP and AuNp+EHL post hoc (Mann Whitney) testing indicates a difference between treatments at p = 0.0015.

6.4 Discussion

In the initial development of the toxicity assays, achieving an experimental design which would ultimately provide a successful and quantitative analysis of cytotoxicity was the overarching aim. With experimental limitations imposed in both temporal and resource availability, the main objective of the cell experiments was to gain as much information as possible on how to proceed with future research and develop a methodology, rather than define the effects of EHL completely and quantitatively. Initial work clearly highlighted some of the areas for basic method improvement using the apparatus available. On visual inspection, EHL seems to have an agglutinating or clumping effect on A431 which prevents samples from evenly distributing across the haemocytometer, preventing accurate representation of the total cell population. Removal of the culture media in the initial direct treatment experiments also means that degradation of the population will occur as a function of their environment, potentially influencing the number of dead cells in the sample, and therefore was disregarded in the next set of work. The SDS-PAGE analysis of the whole cell lysis products from this method (Figure 6.1) was inconclusive, with potentially anomalous results. The attempts to image cells using trypan blue to differentiate between living and dead cells using the haemocytometer proved non-viable for a number of reasons, mainly due to the clumping of cells giving falsely low cell density readings in the quadrants. Initially, the methodological obstacle to be resolved focused on imaging the cells directly. To this end a second set of experiments was set up, where the aim was to find out if there was a difference in the effect seen on other cell lines. A plan to utilise a Dinolite microscope camera to record the outcome of the experiment was made. A431, MDA-MB-468, and MDA-MB-231 were originally to be included. The decision to

exclude MDA-MB-231 was taken as the cells were not growing well. A two hour treatment was applied to A431 and MDA-MB-468, however only the A431 line allowed for any reasonable imaging. As a result of the observations made, namely the extensive aggregation of cells and different culturing outcomes, a more robust methodology measuring intrinsic enzymatic or metabolic linked outcomes was sought. For the final experimental opportunity, this would allow the death rate to be measured by comparative absorbance or fluorescence and a densitometric value to be calculated. The Alamar Blue (Invitrogen) reagent provided a foundation to achieve that outcome. During this preliminary work it was noted that lectin treated cells did not disperse evenly due to aggregation. The result was that cell counts were not representative of the total sample. In order to overcome this characteristic of the treatment and to allow for a more representative quantification of the observed effects, a new approach was pursued which directly measured the effect on a whole sample's metabolic output. Alamar Blue (Invitrogen) is a reagent used to measure cell proliferation and is used to differentiate between metabolically active cells and those which are non-functioning using a redox indicator, Resazurin (Rampersad, 2012). Typically, the oxidation state is dark blue in colour and, if cells actively proliferate, the reduced state of the reagent will progress through a colorimetric gradient to a bright pink colour and fluorescence will increase over time (Figure 6.4). Alamar Blue and its active reagent resazurin show a colometric and flourimetric linear response to proliferation of a broad range of cell densities. However this can vary widely between cell lines and development times can also be subject to a large gradient in rate of change and the range of output values .

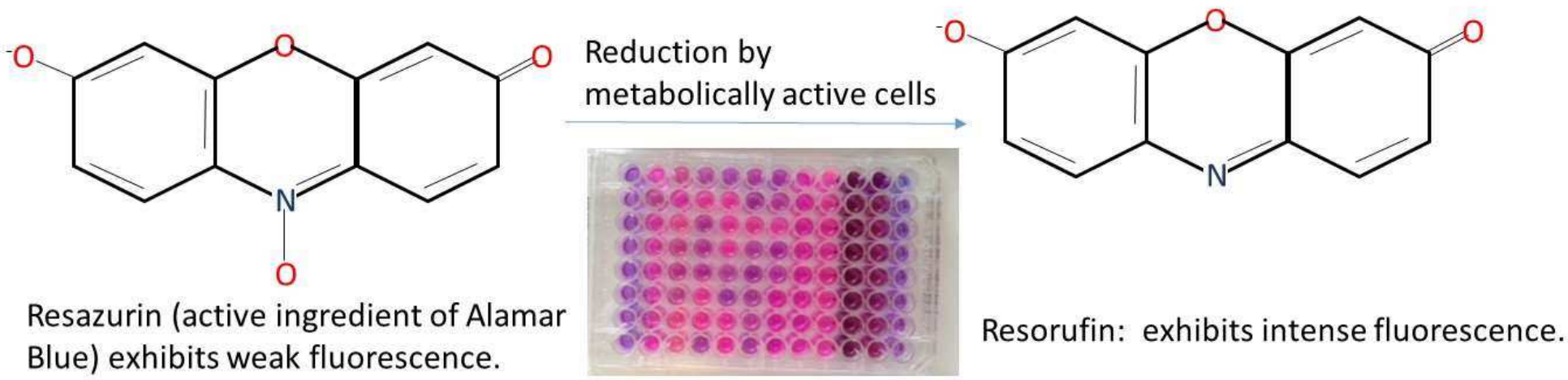


Figure 6.4 Alamar Blue assay uses the reducing properties of metabolically active cells to convert the active ingredient resazurin to resorufin. The fluorescence measured is proportional to the number of living cells. Diagram based on structures published in O'brien *et al.* (2000).

Therefore this assay also required optimisation as part of a new method. Optimisation of the Alamar Blue assay is a timely and complex process in the first instance. However with a single opportunity to conduct the experiment for this work, a standard protocol was used firstly on cell line MDA-MB-468. Initial exploratory results with MDA-MB-468 indicated that a minimum treatment period of 24 hours was required with cell densities in the range of approximately 5000 cells per well on a 96 well plate for the lines used (MDA-MB468, A431 & MDA-MB-231) to begin to observe the predicted colour changes. The initial trial with the MDA-MB-468 was ultimately an assessment of what kind of response would be returned under these parameters, and absorbance was read over an 18 hour period (Figure 6.3A). Absorbance readings did not show as much fine detail as required. Consequently for the next and final two trials, with A431 and MDA-MB-231, fluorescence readings were used as a more sensitive indicator (Figure 6.3B and C).

In future research, cell densities 10 fold higher would be likely to display a faster result with less anomalous data. The method of application employed was to add Alamar Blue to the treatment so as to enable observations over time. The reagent can also be added as an end point marker and future work would include an optimisation of this method. Observations demonstrate a difference in the lethality of protein which had been purified in the presence of protease inhibitor cocktail (Sigma Aldrich) and that which had not, though at this stage of the research that difference could be attributable to other factors too. This could be related to the observations of Virgilio *et al.*, (2010), who suggested that proteolytic cleavage is required for activation of a homologous mature toxin (saporin and ricin), and that the recent trial addition of a protease inhibitor to the extraction method may suppress this outcome despite a positive agglutination test being observed. The findings of this work

are not conclusive, nor are they intended to be so, due to the continued method improvement and optimisation. However, of particular note is the observation that in the individual data plot of A431 it would appear that EHL treated cells at higher concentrations have proliferated more than those cells in lower concentrations or with no treatment applied. Whilst this may seem to be in contradiction to the hypothesis that EHL acts as a cytotoxic agent, work with adding Gal/GalNAc specific lectins such as Jacalin, Amaranth and Peanut lectin to certain cell lines has shown an increase in the proliferation of epithelial cancer cells (Yu *et* al., 2001) with a tentative link to the presence or absence of the T antigen. Whilst this is most certainly only a first step towards a full mammalian cell screen, it is widely recognised that lectins can act as mitogenic agents (Sharon and Lis, 2007) and further work would explore this preliminary finding more deeply. When viewed from this perspective, the initial anomalous results observed in the whole cell lysate SDS-PAGE become more of a point for discussion. The work presented here in no way attempts to report a refined and conclusive experiment, rather the focus is on the development of an optimal technique for doing so.

The statistics presented are a very basic analysis of the findings and have limitations therein. However, it cannot go without note for future research that the most striking outcome of the A431 assay is that there appears to be a dose dependent correlation between increased proliferation and increased concentration of EHL. This mirrors the results seen in Yu *et al.* (2001) in their work on colon cancer cells with Galβ1 GalNAc specific lectins. As A431 is a p53 tumour suppressor gene mutant (p53-R273H) (Reiss *et al.*, 1991), it is reported as therefore being sensitive to mitogenic stimulus (Lamb *et al.*, 1991). With other lectins reported as being mitogenic agents (Sharon and Lis, 2007) this would certainly

be an avenue for further study. MDA-MB-468 also showed an increase in proliferation; since the first data were produced from an absorbance rather than fluorescence reading (with absorbance being less sensitive and furthermore performed under sub optimal conditions); the results at this stage cannot be directly compared. Notably, however, MDA-MB-468 is also a p53-R273H mutant (Muller and Vousden, 2014; Tan *et al.*, 2015) which is a missense single amino acid substitution at arginine 273 of the p53 protein. (Freed-Pastor and Prives, 2012). The p53 tumour suppressor gene product, p53 protein, with more than half of all tumours exhibiting mutation at this locus, is the most frequent target in human cancer mutations (Vogelstein, Lane and Levine, 2000).

Future work would focus on developing the protocol to include longer treatment times, as observations herein suggest that a period of 24 hours would be the minimum appropriate for treatment uptake. General points to improve the method would include the following;

- Using a higher density of cells per well.
- Using only protein extracted in the absence of protease inhibitors.
- Using black culture plates to allow fluorescence to be measured maximum accuracy.
- Measuring fluorescence from the bottom of plates containing adherent cell lines to minimise buffer effects or correcting for Phosphate buffered saline auto-fluorescence.
- Ensuring well volumes exceed 5000 cells by a factor of 10 where possible.
- Correcting for nanoparticle buffer composition and surface plasmon resonance.
- Designing a cell panel to contain both Tn + and Tn - cell lines,

- Investigating mitogenic effects and relationship to p53 mutations.

These adjustments would hopefully allow any significant differences between lines and treatments to be better quantified. The protocol should include all of the above points and would allow more accurate and consistent outcomes to be measured in future studies.

Chapter 7: General discussion and future directions

7.1 Discussion

The aim of this thesis was to conduct an in depth study of the structure and biological effects of the Winter Aconite (*Eranthis Hyemalis*) lectin (EHL). In order for EHL to be used in any future biomedical and biotechnological applications, the underlying structure and biological activity of EHL needed to be elucidated. The current extraction and purification protocol was refined, with the objective of obtaining consistently high protein yields for the additional studies to take place. An investigation of the structure of the EHL protein was also carried out. One aspect of this structural work was using molecular biology techniques to find the underlying gene sequence of the protein. Moreover, EHL was crystallised, and diffraction data were used to establish the three dimensional structure of the protein. The potential toxic effects of EHL in living organisms were explored by conducting several biological assays. These assays initially focused on the model organism *Caenorhabditis elegans*, and both wild type N2 strains and genetically mutated strains were used on this part of the work. The assays were performed using the natural protein, as well as EHL conjugated gold nanoparticles. Finally, preliminary studies of the potential cytotoxic effect of EHL on mammalian cancer cell lines was carried out, in order to develop a viable method for future research on this area.

The previously published extraction and purification protocols were refined as described in detail in Chapter Two. The method improvement has enabled higher yields of protein to be extracted and with techniques such as diafiltration devices being employed, concentrations up to 4 mg/ml have been achieved. With the addition of new steps to the

protocol such as the ammonium sulfate fractionation, an increased purity has been achieved; this has also contributed to the viability of the crystallisation studies performed. Onwards improvement for this purification method would potentially seek to remove the freeze/thaw step in favour of more a efficient chemical method to clarify the product. Whilst for this work the priority was to increase yield, purity and consistency of the output, with those objectives achieved, the parameters of the method could even be further improved given the opportunity to do so in the future, both in economical and temporal aspects. However the overarching aim of this research for the future is to reach an outcome whereby EHL protein can be recombinantly expressed, allowing for poly histidine tags to be utilised and a supply of EHL to be available regardless of season or bulb quality.

Molecular biology techniques were used in an attempt to establish the underlying gene sequence of EHL, as described in more detail in Chapter Three. Initial work using degenerate primers designed from previously published peptide fragments (Kumar *et al*., 1993) was of limited success, and subsequently research funding was granted at the final stage of this work to sequence the genome of Winter Aconite. Genomic DNA was extracted from Winter Aconite and has now been passed by quality standards for genomic sequencing. The primary sequence which will be available as a result of this will require extensive bioinformatics input to assemble the raw data or 'reads'. Whilst this work progresses, the EHL amino acid sequence gathered during this research can be used to probe the genomic data even prior to assembly, and will be a priority for future research in this area. The structure of EHL can be completed as soon as these data become available. The taxonomic comparison of EHL against the RIP family of proteins can then be best informed by examining closely the apparent divergence of the Type II ribosome inactivating

protein from its homologs. The in depth study of the fully complete and refined crystallographic structure will identify all of the glycosylation sites, disulphide bonds and active site residues, providing an opportunity for docking and molecular dynamics studies to begin. A confirmation of these characteristics amongst other structural motifs and domains can provide a means of understanding the regions of evolutionary divergence in the protein which are apparent, particularly the potential differences in the active site regions.

Chapter Four described the work carried out to determine the crystal structure of EHL. Crystallisation screening was carried out using the sitting drop vapour diffusion method and a number of cation/PEG based conditions yielded crystals of diffraction quality. Four crystals yielded diffraction data between 3.2 and 1.6 Å. The highest resolution diffraction data showed a monoclinic $P12_11$ space group, with unit cell dimensions a = 46.84 Å, b = 94.39 Å, c = 71.12 Å, and $\alpha = 90°$ $\beta = 96.47°$ $\gamma = 90°$. Electron density maps generated from the data allowed an initial structure model to be built. From the highest resolution data, a poly-alanine backbone trace was built using SHELXE4; once the protein sequence data are available, it will be possible to read in the sequence to the maps already created to complete the structure. Future work would optimise the crystallisation process by isolating successful conditions and exploring the crystallographic space around them in terms of pH and salt/ precipitant parameters. Whilst the aim of any crystallographic experiment is to yield high resolution data which has already been the outcome herein, there is scope for future work on this aspect of the research. The length of time taken for crystal growth and the consistent output of large single crystals as opposed to many optimally sized crystals limits the work that can be done with them. In order for substrate binding experiments to

take place in the future, it would be ideal to have many crystals with a shorter growth period in the same condition with which to work. Crystallisation studies with a substrate bound with an analog such as adenine monophosphate would provide more structural information about any differences in active site regions as is currently proposed from the model, thus contributing new knowledge to the field of RIP research.

A set of biological assays was conducted to explore the potential toxicity of the protein in living organisms. The initial biological assays focused on the model organism *Caenorhabditis elegans* and included using wild type N2 strains and also an exploration of the effects EHL had on genetically mutated strains. Experiments were carried out with the natural protein and with EHL conjugated gold nanoparticles. Winter Aconite Lectin EHL bound specifically to the amphid neurons of *Caenorhabditis elegans* when treated as an L1 larvae, causing impairment of chemosensory function, development and fecundity, and disruption of the normal dauer response pathway. A different effect was observed in assays conducted on later life stages when the dauer decision is no longer an option. Given the opportunity to pursue this work further, an avenue to explore would be with worms expressing pan-neuronal fluorescence to identify whether the differing phenotypic effects already observed correlate to the outright death, or merely to the damage of particular sets of neurons. Mutant screens could be carried out to potentially identify lectin resistant strains; ultimately this would contribute to our understanding of lectin binding and the relationship with particular receptor groups and cell types. Also of benefit would be to further categorise the effects of treatment during various developmental stages. This would be in order to map changes as differing cell types are exposed, such as epithelial

cells (which are due to cuticular changes)and the onset of pharyngeal pumping (feeding), especially if fluorescently labelled EHL could be used to map localisation of the toxin.

The opportunity to conduct mammalian cancer cell investigations arose towards the conclusion of this study, and whilst the main scope of the experiments was to achieve a working protocol for a toxicity screen for future work, one of the outcomes of that work potentially demonstrates that EHL may have mitogenic properties in combination with p53 mutant lines. It would be interesting to establish whether this as a result of an additional mechanism such as a higher binding affinity in those cells or if it is a purely tumour suppressor gene dependent effect. With this observation in mind using a panel of cells would seek to explore further and to more clearly define the relationship between EHL effects on a number of cell types, both healthy and cancerous.

One of the objectives of this work was to establish the viability of a functionalised EHL nanoparticle conjugate and its potential effect on the cytotoxicity of EHL; and the results from this research can be found in Chapters Three and Five. The aim of this was first to provide some insights into the viability of the conjugate, and then perform some preliminary assays to establish if the EHL biological effect is altered, providing the ground for future research. Gold nanoparticles functionalised with EHL were successfully synthesized by Dr Javier Lodeiro and Ms Jamila Djafari at the BIOSCOPE-PROTEOMASS Laboratory (FCT-UNL, Portugal), and assays using *C. elegans* and mammalian cancer cells performed. With the successful conjugation of EHL to gold nanoparticles in place, countless opportunities have opened up. Further work would investigate the structural impact on EHL from the nanoparticles and whether this consistently renders the protein less active overall as seen in the *C. elegans* experiments. The nematology experiments indicate that

the protein is affected, potentially due to inhibition of binding caused by the presence of the nanoparticle or perhaps because of an overall conformational change has rendered the b-chain, which is of course responsible for agglutination, inactive. The alternative hypothesis is that both chains are affected and that the toxic effect is also reduced. These questions would form the foundations of any further work with the nanoparticle conjugate. This has great potential in cytology applications since nanoparticles themselves show a greater permeability across cancer cell membranes than healthy cells, as discussed in Chapter Six. The idea that they can be used as delivery systems for drugs and imaging agents lends itself to the notion of targeting cells with the A chain of EHL. EHL would seem to show lower overall toxicity as part of the native protein than its more toxic counterparts Ricin and Abrin, but potentially when delivered via a nanoparticle conjugate to the region of interest could deliver the same N-glycosidase activity.

Throughout any discussion of lectins and RIPs, the overall function of the protein remains under debate and whilst many hypotheses have been tested and shown promising explanations, a definitive answer is still elusive. A future study on the fully refined and sequenced structure of EHL could be aimed towards establishing that function.

7.2 Concluding remarks

This study has contributed new knowledge to the field of lectin research. Winter aconite is the first plant in its family to have lectin and ribosome inactivating activity described. This adds to our understanding of the evolutionary relationships in this important class of molecules. Through X-ray crystallography diffraction experiments, a new protein structure has been defined and whilst not yet fully complete, already contributes useful structural information to the field. On completion of the genome assembly and accompanying

structure refinement, it will ultimately result in the deposition of a new structure in the PDB for other researchers to use. The dauer phenotype described has not been reported before, and whilst *C. elegans* is a free living nematode, it can act as a model organism for understanding parasitic nematodes such as *Steinernema* and *Heterorhabditis* (which parasitise insects), and *Strongyloides stercoralis*, which is a human parasitic species (Riddle and Albert, 1997).

The development of a cellular toxicity assay with EHL is in its beginnings. However early indications are that there are interesting and previously unreported effects of this important plant toxin against mammalian cells.

In conclusion, this work has attempted to tell a complete story on which to build further studies of EHL. That story was to experimentally find a way to gather enough protein to conduct further studies, firstly in order that crystals could be grown, to find out the molecular structure of the toxin. The question then to be answered was: how does it affect living organisms and is it a useful tool in biological research? The research conducted shows that EHL displays blood type specific agglutination, mitogenic activity, can be successfully conjugated to gold nanoparticles, displays neuron specific cytotoxic binding properties, and also shows an interesting divergence from ancestral RIPs. The results from this study have unlocked huge potential for EHL to undergo considerable onwards investigation and testing in a variety of directions, and has contributed new knowledge to this expanding area of research.

References

Adams, M.J., Blundell, T.L., Dodson, E.J., Dodson, G.G., Vijayan, M., Baker, E.N., Harding, M.M., Hodgkin, D.C., Rimmer, B. and Sheat, S. (1969) 'Structure of rhombohedral 2 zinc insulin crystals' *Nature*, *224*(5218), pp.491-495.

Adwan, H., Bayer, H., Pervaiz, A., Sagini, M. and Berger, M.R. (2014) 'Riproximin is a recently discovered type II ribosome inactivating protein with potential for treating cancer'. *Biotechnology Advances*, *32*(6), pp.1077-1090.

Ahmad, M.Z., Akhter, S., Rahman, Z., Akhter, S., Anwar, M., Mallik, N. and Ahmad, F.J. (2013) 'Nanometric gold in cancer nanotechnology: current status and future prospect'. *Journal of Pharmacy and Pharmacology*, *65*(5), pp.634-651.

Albert, P.S., Brown, S.J. and Riddle, D.L. (1981) 'Sensory control of dauer larva formation in Caenorhabditis elegans'. *Journal of Comparative Neurology* 198:435-451.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) 'Basic local alignment search tool'. *Journal of Molecular Biology*, *215*(3), pp.403-410.

Altun, Z.F. and Hall, D.H. (2005) 'Handbook of C. elegans Anatomy.WormAtlas'. http://www. wormatlas. org/handbook/contents. htm.

Arnau, J., Lauritzen, C., Petersen, G.E. and Pedersen, J. (2006) 'Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins'. *Protein Expression and Purification*, *48*(1), pp.1-13**.**

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

Boyd, W.A., Smith, M.V. and Freedman, J.H. (2012) '*Caenorhabditis elegans* as a model in developmental toxicology' *Methods in Molecular Biology* 889:15-24.

Bragg, W.H. and Bragg, W.L. (1913) 'The reflection of X-rays by crystals'*Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *88*(605), pp.428-438.

Brooks-Bartlett, J. and Garman, E.F. (2015) 'The Nobel Science: One Hundred Years of Crystallography'. *Interdisciplinary Science Reviews*. Available at: http://dx.doi.org/10.1179/0308018815Z.000000000116 (Accessed 05/09/16)

Cammue, B., Peeters, B., and Peumans, W. (1985) 'Isolation and partial characterization of an N-acetylgalactosamine-specific lectin from winter-aconite (Eranthis hyemalis) root tubers'. *Biochemical Journal* 227:949.

Cassada, R.C. and Russell, R.L. (1975) 'The dauer larva, a post-embryonic developmental variant of the nematode *Caenorhabditis elegans'*. *Developmental Biology* 46:326-342.

Chayen, N.E. and Saridakis, E. (2008) 'Protein crystallization: from purified protein to diffraction-quality crystal'. *Nature methods*, *5*(2), pp.147-153.

Coburn, C.M. and Bargmann, C.I. (1996) 'A putative cyclic nucleotide-gated channel is required for sensory development and function in *C. elegans*' *Neuron* 17:695-706.

Corsi, A.K., Wightman, B. and Chalfie, M. (2015) 'A Transparent window into biology: A primer on *Caenorhabditis elegans*' *Genetics* 200(2), pp.387-407.

Crowfoot, D., Bunn, C. W., Rogers-Low, B. W., and Turner-Jones, A. (1949) *The X-ray crystallographic investigation of the structure of penicillin* (pp. 310-366). Princeton, New Jersey: Princeton University Press,

Dan, X., Liu, W. and Ng, T.B. (2016) 'Development and applications of lectins as biological tools in biomedical research' *Medicinal research reviews*, *36*(2), pp.221-247.

Delatorre, P., Rocha, B.A., Souza, E.P., Oliveira, T.M., Bezerra, G.A., Moreno, F.B., Azevedo, W.F. (2007) 'Structure of a lectin from *Canavalia gladiata* seeds: new structural insights for old molecules' *BMC Structural Biology* 7:52.

Dusenbery, D.B. (1976) 'Chemotactic behavior of mutants of the nematode *C. elegans* that are defective in their attraction to NaCl' *Journal of Experimental Zoology* 198:343-352.

Dusenbery, D.B. (1980) 'Chemotactic behavior of mutants of the nematode *C. elegans* that are defective in osmotic avoidance' *Journal of Comparative Physiology* 137:93-96.

Edwards, M.G and Gatehouse, A.M. (2007) 'Biotechnology in crop protection: Towards sustainable insect control'. In Vurro, M. and Gressel, J. (eds.) *Novel Biotechnologies for Biocontrol Aagent Enhancement and Management*. Springer Science & Business Media. Netherlands: Springer.

Elmsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) 'Features and development of Coot' *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), pp.486-501.

Evans, P.R. (2006) 'Scaling and assessment of data quality' *Acta Crystallographica Section D: Biological Crystallography*, *62*(1), pp.72-82.

Evans, P.R. (2011) 'An introduction to data reduction: space-group determination, scaling and intensity statistics'. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), pp.282-292.

Fernández-Lodeiro, A., Fernández-Lodeiro, J., Núñez, C., Bastida, R., Capelo, J.L. and Lodeiro, C. (2013) 'Polyamine Ligand-Mediated Self-Assembly of Gold and Silver Nanoparticles into Chainlike Structures in Aqueous Solution: Towards New Nanostructured Chemosensors' *ChemistryOpen*, *2*(5-6), pp.200-207.

Ferreras, J.M., Citores, L., Iglesias, R., Jiménez, P. and Girbés, T., (2011) 'Use of ribosome-inactivating proteins from Sambucus for the construction of immunotoxins and conjugates for cancer therapy' *Toxins*, *3*(5), pp.420-441.p.307-310.

Freed-Pastor, W.A. and Prives, C. (2012) 'Mutant p53: one name, many proteins' *Genes & Development*, *26*(12), pp.1268-1286.

Frens, G. (1973) 'Controlled nucleation for the regulation of the particle size in monodisperse gold suspensions' *Nature*, 241(105), pp.20-22.

Fujiwara, M., Sengupta, P. and McIntire, S.L. (2002) 'Regulation of body size and behavioral state of *C. elegans* by sensory perception and the EGL-4 cGMP-dependent protein kinase' *Neuron* 36:1091-1102.

Garman, E., (2003) ''Cool'crystals: macromolecular cryocrystallography and radiation damage' *Current Opinion in Structural Biology*, *13*(5), pp.545-551.

George, O., Solscheid, C., Bertolo, E. and Lisgarten, D., (2011) 'Extraction and purification of the lectin found in the tubers of Eranthis hyemalis (winter aconite)' *Journal of Integrated OMICS*, 1(2), pp.268-272.

Ghazarian, H., Idoni, B. and Oppenheimer, S.B., (2011) 'A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics' *Acta histochemica*, *113*(3), pp.236-247.

Golden, J.W. and Riddle, D.L. (1984) 'The *Caenorhabditis elegans* dauer larva: developmental effects of pheromone, food, and temperature' *Developmental Biology* 102: 368–378

Gorelik, E., Galili, U. and Raz, A. (2001) 'On the role of cell surface carbohydrates and their binding proteins (lectins) in tumor metastasis' *Cancer and Metastasis Reviews*, *20*(3-4), pp.245-277.

Gorrec, F. (2009) 'The MORPHEUS protein crystallization screen'. *Journal of Applied Crystallography*, *42*(6), pp.1035-1042.

Green, E.D., Adelt, G., Baenziger, J.U., Wilson, S. and Van Halbeek, H., (1988) 'The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy' *Journal of Biological Chemistry,* 263(34), pp.18253-18268.

Green, J.W.M. and Harvey S.C. (2012) 'Development of *Caenorhabditis elegans* dauer larvae in growing populations' *Nematology*, 14:165-173.

Green J.W.M., Snoek L.B., Kammenga J.E. and Harvey S.C. (2013) 'Genetic mapping of variation in dauer larvae development in growing populations of *Caenorhabditis elegans' Heredity* 111:306-313.

Green J.W.M., Stastna J.J., Orbidans H.E. and Harvey S.C. (2014) 'Highly polygenic variation in environmental perception determines dauer larvae formation in growing populations of *Caenorhabditis elegans' PLoS One*. 9(11): e112830.

Grimm, C., Chari, A., Reuter, K. and Fischer, U. (2010) 'A crystallization screen based on alternative polymeric precipitants'. *Acta Crystallographica Section D: Biological Crystallography*, *66*(6), pp.685-697.

Hammond, C. and Hammond, C. (2009) *The Basics of Crystallography and Diffraction* (Vol. 12). Oxford: Oxford University Press.

Hanahan, D. and Weinberg, R.A. (2011) 'Hallmarks of cancer: the next generation' *cell*, *144*(5), pp.646-674.

Hartley, M. and Lord, J. (2004) 'Cytotoxic ribosome-inactivating lectins from plants' *Biochimica Et Biophysica Acta - Proteins and Proteomics* 1701:1-14.

Hazes, B. (1996) 'The (QxW) 3 domain: a flexible lectin scaffold' *Protein Science*, *5*(8), pp.1490-1501.

Hedgecock, E.M., Culotti J.G., Thomson J.N. and Perkins L.A. (1985) 'Axonal guidance mutants of *Caenorhabditis elegans* identified by filling sensory neurons with fluorescein dyes' *Developmental Biology* 111:158-170.

Heinrich, E.L., Welty, L.A.Y., Banner, L.R. and Oppenheimer, S.B., (2005) 'Direct targeting of cancer cells: a multiparameter approach' *Acta histochemica*, *107*(5), pp.335-344.

Hodgkin, D. C., Kamper, J., Mackay, M., Pickworth, J., Trueblood, K. N., and White, J. G. (1956) 'Structure of vitamin B12'. *Nature*, *178*(4524), 64-66.

Hodgkin, J. and Doniach, T. (1997) 'Natural variation and copulatory plug formation in *Caenorhabditis elegans' Genetics*, 146:149-164.

Hofmann, B.T., Schlüter, L., Lange, P., Mercanoglu, B., Ewald, F., Fölster, A., Picksak, A.S., Harder, S., El Gammal, A.T., Grupp, K. and Güngör, C. (2015) 'COSMC knockdown mediated aberrant O-glycosylation promotes oncogenic properties in pancreatic cancer' *Molecular Cancer*, 14(1), p.1.

Horneber, M.A., Bueschel, G., Huber, R., Linde, K., Rostock, M. (2008) 'Mistletoe therapy in oncology'*Cochrane Database of Systematic Reviews* 2:CD003297

Hou, X., Chen, M., Chen, L., Meehan, E.J., Xie, J. and Huang, M. (2007) 'X-ray sequence and crystal structure of luffaculin 1, a novel type 1 ribosome-inactivating protein' *BMC Structural Biology*, *7*(1), p.1.

Hutter, E. and Fendler, J.H. (2004) 'Exploitation of localized surface plasmon resonance' *Advanced Materials*, *16*(19), pp.1685-1706.

Hutter, E. and Maysinger, D. (2011) 'Gold nanoparticles and quantum dots for bioimaging' *Microscopy Research and Technique*, *74*(7), pp.592-604.

Iserte, J.A., Stephan, B.I., Goñi, S.E., Borio, C.S., Ghiringhelli, P.D. and Lozano, M.E., (2013) 'Family-specific degenerate primer design: a tool to design consensus degenerated oligonucleotides' *Biotechnology Research International 2013* available at: http://dx.doi.org/10.1155/2013/383646 Accessed 05/09/2016.

Jankarik, J. and Kim, S.H. (1991) 'Sparse matrix sampling: a screening method for crystallisation of proteins'. *Journal of Applied Crystallography, 24*(4), pp. 409-411.

Ju, T., Otto, V. I. and Cummings, R.D. (2011) 'The Tn antigen—structural simplicity and biological complexity' *Angewandte Chemie International Edition* 50(8):1770-1791.

Kaletta, T. and Hengartner, M.O. (2006) 'Finding function in novel targets: *C. elegans* as a model organism' *Nature Reviews Drug Discovery*, *5*(5), pp.387-399.

Kilpatrick, D.C. (1998) 'Use of lectins as mitogens for lymphocytes' *Lectin Methods and Protocols* in Rhodes, J.M. and Milton, J.D. (eds) Methods in Molecular Medicine (vol. 9) pp.385-392. Totowa, USA: Humana press inc.(Springer).

Kim, Y. and Robertus, J.D. (1992) 'Analysis of several key active site residues of ricin A chain by mutagenesis and X-ray crystallography' *Protein Engineering*, *5*(8), pp.775-779

Klass, M. and Hirsh, D. (1976) 'Non-ageing developmental variant of *Caenorhabditis elegans' Nature,* 260, pp.523-525.

Kodiha, M., Wang, Y.M., Hutter, E., Maysinger, D. and Stochaj, U. (2015) 'Off to the organelles—killing cancer cells with targeted gold nano-particles' *Theranostics*, 5(4), pp.357-370.

Krauspenhaar, R., Rypniewski, W., Kalkura, N., Moore, K., DeLucas, L., Stoeva, S., Mikhailov, A., Voelter, W. and Betzel, C. (2002) 'Crystallisation under microgravity of mistletoe lectin I from Viscum album with adenine monophosphate and the crystal structure at 1.9 Å resolution' *Acta Crystallographica Section D: Biological Crystallography*, *58*(10), pp.1704-1707.

Kumar, M.A., Timm, D., Neet, K., Owen, W., Peumans, W.J., Rao, A.G. (1993) 'Characterization of the lectin from the bulbs of *Eranthis hyemalis* (winter aconite) as an inhibitor of protein synthesis' *Journal of Biological Chemistry* 268:25176-25183.

Lamb, D.J., Spotts, G.S., Shubhada, S. and Baker, K.R., (1991) 'Partial characterization of a unique mitogenic activity secreted by rat Sertoli cells' *Molecular and Cellular Endocrinology*, *79*(1), pp.1-12.

Landsteiner, K. (1907) 'Spezifische Bindung und Antikörper. IV. Hämagglutination und Hämolyse'. *Handbuch der Biochemie Des Menschen und der Tiere*, (vol. 2) Part 1, p. 395.

Lapadula, W.J., Puerta, M.V.S. and Ayub, M.J. (2013) 'Revising the taxonomic distribution, origin and evolution of ribosome inactivating protein genes' *PloS one*, *8*(9), p.e72825.

Lesk, A. (2010) *Introduction to protein science: architecture, function, and genomics*. Oxford: Oxford university press.

Lewis, J.A. and Hodgkin, J.A. (1977) 'Specific neuroanatomical changes in chemosensory mutants of the nematode *Caenorhabditis elegans' Journal of Comparative Neurology* 172(3):489-510.

Liu, X., Atwater, M., Wang, J. and Huo, Q. (2007) 'Extinction coefficient of gold nanoparticles with different sizes and different capping ligands' *Colloids and Surfaces B: Biointerfaces*, *58*(1), pp.3-7.

Lodish, H. (2008) *Molecular Cell Biology*. 6th edn. Basingstoke: W H Freeman and Co.

Madariaga, D., Martínez-Sáez, N., Somovilla, V.J., Coelho, H., Valero-González, J., Castro-López, J., Asensio, J.L., Jiménez-Barbero, J., Busto, J.H., Avenoza, A. and Marcelo, F. (2014) 'Detection of tumor-associated glycopeptides by lectins: The peptide context modulates carbohydrate recognition' *ACS chemical biology*, *10*(3), pp.747-756.

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I. and Lanczycki, C.J. (2014) 'CDD: NCBI's conserved domain database' *Nucleic Acids Research*, p.gku1221.

Matthews, B.W. (1968) 'Solvent content of protein crystals' *Journal of Molecular Biology*, *33*(2), pp.491-497.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J., (2007) 'Phaser crystallographic software' *Journal of Applied Crystallography*, *40*(4), pp.658-674.

Mier, J.W. and Gallo, R.C. (1982) 'The purification and properties of human T cell growth factor' *The Journal of Immunology*, *128*(3), pp.1122-1127.

Missouri Botanical Gardens (2016) available at: http://www.missouribotanicalgarden.org/PlantFinder/PlantFinderDetails.aspx?taxonid=263882

Mody, R., Antaram Joshi, S.H. and Chaney, W. (1995) 'Use of lectins as diagnostic and therapeutic tools for cancer' *Journal of Pharmacological and Toxicological Methods*, *33*(1), pp.1-10.

Monzingo, A.F. and Robertus, J.D. (1992) 'X-ray analysis of substrate analogs in the ricin A-chain active site' *Journal of Molecular Biology*, *227*(4), pp.1136-1145.

Morgan, W.T.J. and Watkins, W.M. (1953) 'The inhibition of the haemagglutinins in plant seeds by human blood group substances and simple sugars' *British Journal of Experimental Pathology*, *34*(1), p.94.

Morgan, D.A., Ruscetti, F.W. and Gallo, R. (1976) 'Selective in vitro growth of T lymphocytes from normal human bone marrows' *Science*, *193*(4257), pp.1007-1008.

Muller, P.A. and Vousden, K.H. (2014) 'Mutant p53 in cancer: new functions and therapeutic opportunities' *Cancer Cell*, *25*(3), pp.304-317.

Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) 'Refinement of macromolecular structures by the maximum-likelihood method' *Acta Crystallographica Section D: Biological Crystallography*, *53*(3), pp.240-255.

Murshudov, G.N. (2011) 'REFMAC5 for the refinement of macromolecular crystal structures' *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), pp.355-367.

Newman, J., Egan, D., Walter, T.S., Meged, R., Berry, I., Ben Jelloul, M., Sussman, J.L., Stuart, D.I. and Perrakis, A. (2005) 'Towards rationalization of crystallization screening for small-to medium-sized academic laboratories: the PACT/JCSG+ strategy'. *Acta Crystallographica Section D: Biological Crystallography*, *61*(10), pp.1426-1431.

Nowell, P.C. (1960) 'Phytohemagglutinin: an initiator of mitosis in cultures of normal human leukocytes' *Cancer Research*, *20*(4), pp.462-466.

O'brien, J., Wilson, I., Orton, T. and Pognan, F., (2000). 'Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity'. *European Journal of Biochemistry*, *267*(17), pp.5421-5426.

Otsuka, H., Gotoh, Y., Komeno, T., Ono, T., Kawasaki, Y., Iida, N., Shibagaki, Y., Hattori, S., Tomatsu, M., Akiyama, H., Tashiro, F. (2014) 'Aralin, a type II ribosome-inactivating protein from Aralia elata, exhibits selective anticancer activity through the processed form of a 110-kDa high-density lipoprotein-binding protein: a promising anticancer drug' *Biochemical and Biophysical Research Communications* 453:117-123.

Papadopoulos, J.S. and Agarwala, R. (2007) 'COBALT: constraint-based alignment tool for multiple protein sequences' *Bioinformatics*, *23*(9), pp.1073-1079.

Patra, C.R., Bhattacharya, R., Mukhopadhyay, D. and Mukherjee, P. (2010) 'Fabrication of gold nanoparticles for targeted therapy in pancreatic cancer' *Advanced Drug Delivery Reviews*, *62*(3), pp.346-361.

Peumans, W.J. and Van Damme, E. (1995) 'Lectins as plant defense proteins' *Plant Physiology* 109:347.

Peumans, W.J., Hao, Q. and Van Damme, E.J. (2001) 'Ribosome-inactivating proteins from plants: more than RNA N-glycosidases?' *The FASEB Journal*,*15*(9), pp.1493-1506.

Peumans, W.J. and Van Damme, E.J. (2010) 'Evolution of plant ribosome-inactivating proteins' In: Lord, J.M. and Hartley, M.R. (eds.)*Toxic Plant Proteins* (Vol. 18) pp. 1-26 Berlin: Springer.

R Core Team 2013 (2013) 'R: A language and environment for statistical computing' R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rampersad, S.N. (2012) 'Multiple applications of Alamar Blue as an indicator of metabolic function and cellular health in cell viability bioassays' *Sensors*,*12*(9), pp.12347-12360.

Rao, K., Rathore, K.S., Hodges, T.K, Fu, X., Stoger, E., Sudhakar, D., Bown, D.P. (1998) 'Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper' *Plant Journal* 15:469-477.

Reiss, M., Brash, D.E., Munoz-Antonia, T., Simon, J.A., Ziegler, A., Vellucci, V.F. and Zhou, Z.L. (1991) 'Status of the p53 tumor suppressor gene in human squamous carcinoma cell lines' *Oncology research*, *4*(8-9), pp.349-357.

Renkonen, K.O. (1948) 'Studies on hemagglutinins present in seeds of some representatives of the family of Leguminoseae' In *Annales Medicinae Experimentalis et Biologiae Fenniae* (Vol. 26, No. 1, pp. 66-72).

Riddle, D.L. and Albert, P.S. (1997) 'Genetic and Environmental Regulation of Dauer Larva Development' *Cold Spring Harbor Monograph Archive*, *33*, pp.739-768.

Riddle, D., Blumenthal, T., Meyer, B., Priess, J. (1997) 'Introduction to *C. elegans*' *Cold Spring Harbor Monograph Archive* Available at: http://cshmonographs.org/csh/index.php/monographs/article/view/5036/4135

Royal Horticultural Society.org (2016) available at: https://www.rhs.org.uk/Plants/6540/i-Eranthis-hyemalis-i/Details

Rupp, B., (2009) *Biomolecular crystallography: principles, practice, and application to structural biology*. New York, Garland Science.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) 'DNA sequencing with chain-terminating inhibitors'. *Proceedings of the National Academy of Sciences*, *74*(12), pp.5463-5467.

Schubert, M., Bleuler-Martinez, S., Butschi, A., Wälti, M.A., Egloff, P., Stutz, K., Yan, S., Wilson, I.B.H., Hengartner, M.O., Aebi, M., Allain, F.H.T. and Künzler, M. (2012) 'Plasticity of the β-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system' *PLoS Pathogens* 8:e1002706.

Schumacher, K., Schneider, B., Reich, G., Stiefel, T., Stoll, G., Bock, P.R., Hanisch, J. and Beuth, J. (2002) 'Influence of postoperative complementary treatment with lectin-standardized mistletoe extract on breast cancer patients. A controlled epidemiological multicentric retrolective cohort study' *Anticancer Research*, *23*(6D), pp.5081-5087.

Schwarz, H.P. and Dorner, F. (2003) 'Karl Landsteiner and his major contributions to haematology' *British Journal of Haematology*, *121*(4), pp.556-565.

Schwarzenbacher, R., Godzik, A., Grzechnik, S.K. and Jaroszewski, L. (2004) 'The importance of alignment accuracy for molecular replacement' *Acta Crystallographica Section D: Biological Crystallography*, *60*(7), pp.1229-1236.

Sela, B.A., Wang, J.L. and Edelman, G.M. (1975) 'Isolation of lectins of different specificities on a single affinity adsorbent' *Journal of Biological Chemistry,* 250(18), pp.7535-7538.

Shang, L., Wang, Y., Jiang, J. and Dong, S. (2007) 'pH-dependent protein conformational changes in albumin:Gold nanoparticle bioconjugates: A spectroscopic study' *Langmuir 23*, 2714–2721.

Sharon, N. and Lis, H. (2004) 'History of lectins: From hemagglutinins to biological recognition molecules' *Glycobiology* 14:53-62.

Sharon, N. and Lis, H. (2007) *Lectins*. 2nd edn. Berlin: Springer Science & Business Media.

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

Shevchenko, A., Sunyaev, S., Loboda, A., Shevchenko, A., Bork, P., Ens, W. and Standing, K.G. (2001) 'Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching' *Analytical Chemistry*, *73*(9), pp.1917-1926.

Starich, T.A., Herman, R.K., Kari, C.K., Schackwitz, W.S., Schuyler, M.W., Collet, J., Thomas, J.H. and Riddle, D.L. (1995) 'Mutations affecting the chemosensory neurons of *Caenorhabditis elegans*' *Genetics* 139:171-188.

Stastna, J.J., Snoek, L.B., Kammenga, J.E. and Harvey, S.C. (2015) 'Genotype-dependent lifespan effects in peptone deprived *Caenorhabditis elegans*' *Scientific reports*, *5*.

Stein, N. (2008) 'CHAINSAW: a program for mutating pdb files used as templates in molecular replacement' *Journal of Applied Crystallography*,*41*(3), pp.641-643.

Stiernagle, T. (2006) 'Maintenance of *C. elegans*' in WormBook ed. The *C. elegans* Research Community, *WormBook*, doi/10.1895/wormbook.1.101.1, http://www.wormbook.org.

Stillmark, H. (1888) *'Ueber ricin: ein giftiges ferment aus den samen von Ricinus comm. 1. und einigen anderen euphorbiaceen'.* Doctoral dissertation, Schnakenburg's buchdruckr.

Stowell, S.R., Ju, T. and Cummings, R.D. (2015) 'Protein glycosylation in cancer'. *Annual Review of Pathology*, *10*, p.473.

Stutz, K., Kaech, A., Aebi, M., Künzler, M. and Hengartner, M.O., (2015) 'Disruption of the *C. elegans* intestinal brush border by the fungal lectin CCL2 phenocopies dietary lectin toxicity in mammals' *PloS One*, 10(6), p.e0129381.

Sulston, J.E. and Horvitz, H. (1977) 'Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*' *Developmental Biology* 56:110-156.

Tan, B.S., Tiong, K.H., Choo, H.L., Chung, F.F.L., Hii, L.W., Tan, S.H., Yap, I.K., Pani, S., Khor, N.T., Wong, S.F. and Rosli, R. (2015) 'Mutant p53-R273H mediates cancer cell survival and anoikis resistance through AKT-dependent suppression of BCL2-modifying factor (BMF)' *Cell Death & Disease*, *6*(7), p.e1826.

Tonevitsky, A., Agapov, I., Shamshiev, A., Temyakov, D., Pohl, P., Kirpichnikov, M. (1996) 'Immunotoxins containing A-chain of Mistletoe lectin I are more active than immunotoxins with ricin A-chain' *FEBS Letters* 392:166-168.

Uchida, O., Nakano, H., Koga, M. and Ohshima, Y. (2003) 'The *C. elegans* che-1 gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons' *Development* 130:1215-1224.

Van Damme, E.J., Hao, Q., Chen, Y., Barre, A., Vandenbussche, F., Desmyter, S., Rougé, P. and Peumans, W.J. (2001) 'Ribosome-inactivating proteins: a family of plant proteins that do more than inactivate ribosomes' *Critical Reviews in Plant Sciences*, *20*(5), pp.395-465.

Vassilieva, L.L. and Lynch, M. (1999) 'The rate of spontaneous mutation for life-history traits in *Caenorhabditis elegans*' Genetics 151:119-129.

Virgilio, M.D., Lombardi, A., Caliandro, R. and Fabbrini, M.S. (2010) 'Ribosome-inactivating proteins: from plant defense to tumor attack' *Toxins*,*2*(11), pp.2699-2737.

Vogelstein, B., Lane, D. and Levine, A.J. (2000) 'Surfing the p53 network' *Nature*, *408*(6810).

Voss, C., Eyol, .E, Frank, M., Von der Lieth, C.W. and Berger, M.R. (2006) 'Identification and characterization of riproximin, a new type II ribosome-inactivating protein with antineoplastic activity from Ximenia americana' *FASEB Journal* 20:1194-1196.

Ward, S., Thomson, N., White, J.G., Brenner, S. (1975) 'Electron microscopical reconstruction of the anterior sensory anatomy of the nematode *Caenorhabditis elegans' Journal of Comparative Neurology*, 160:313-337.

Watson, J.D. and Crick, F.H. (1953) 'Molecular structure of nucleic acids' *Nature*, *171*(4356), pp.737-738.

Weston, S.A., Tucker, A.D., Thatcher, D.R., Derbyshire, D.J. and Pauptit, R.A. (1994) 'X-ray structure of recombinant ricin A-chain at 1.8 Å resolution' *Journal of Molecular Biology*, *244*(4), pp.410-422.

Wiley, R.G., Blessing, W. and Reis, D.J. (1982) 'Suicide transport: Destruction of neurons by retrograde transport of ricin, abrin, and modeccin' *Science* 216:889-890.

Wingfield, P. (2001) 'Protein precipitation using ammonium sulfate' *Current Protocols in Protein Science*, pp.A-3F.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A. and McNicholas, S.J. (2011) 'Overview of the CCP4 suite and current developments' *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), pp.235-242.

Wlodawer, A., Minor, W., Dauter, Z. and Jaskolski, M. (2008) 'Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures' *Febs Journal*, *275*(1), pp.1-21.

Yan, X., Hollis, T., Svinth, M., Day, P., Monzingo, A.F., Milne, G.W. and Robertus, J.D., (1997) 'Structure-based identification of a ricin inhibitor'.*Journal of Molecular Biology*, *266*(5), pp.1043-1049.

Yu, L.G., Milton, J.D., Fernig, D.G. and Rhodes, J.M. (2001) 'Opposite effects on human colon cancer cell proliferation of two dietary Thomsen-Friedenreich antigen-binding lectins' *Journal of cellular physiology*, *186*(2), pp.282-287.

Appendix 1.1 Publication manuscript as submitted to PeerJ.

Winter Aconite (Eranthis hyemalis) Lectin as a cytotoxic effector in the lifecycle of Caenorhabditis elegans

Marie-Therese McConnell, David R Lisgarten, Lee J Byrne, Simon C Harvey and Emilia Bertolo*

Biomolecular Research Group, School of Human and Life Sciences, Canterbury Christ Church University, Canterbury, Kent, England, CT1 1QU

* Address Correspondence to: Emilia Bertolo, School of Human and Life Sciences, Canterbury Christ Church University, Canterbury, Kent, England, CT1 1QU Telephone: 01227 78 2335 Email: emilia.bertolo@canterbury.ac.uk

Abstract

The lectin found in the tubers of the Winter Aconite (Eranthis hyemalis) plant is an N-acetyl-D-galactosamine specific Type II Ribosome Inactivating Protein (RIP); Type II RIPs have shown anti-cancer properties, and hence have potential as therapeutic agents. Here we present a modified protocol for the extraction and purification of the E. hyemalis lectin (EHL) using affinity chromatography. De novo amino acid sequencing of EHL confirms its classification as a Type II Ribosome Inactivating protein. The biocidal properties of EHL have been investigated against the nematode Caenorhabditis elegans. Arrested first stage larvae treated with EHL have shown some direct mortality, with surviving larvae subsequently showing a range of phenotypes including food avoidance, reduced fecundity, developmental delay and constitutive dauer larvae formation. Both inappropriate dauer larvae development and failure to locate to bacterial food source are consistent with the disruption of chemosensory function and the ablation of amphid neurons. Further investigation indicates that mutations that disrupt normal amphid formation can block the EHL-induced dauer larvae formation. In combination, these phenotypes indicate that EHL is cytotoxic and suggest a cell specific activity against the amphid neurons of C. elegans.

Introduction

Lectins are a class of carbohydrate binding proteins ubiquitously expressed in plants, animals, bacteria and viruses, characterised by their ability to agglutinate erythrocytes (Peumans & Van Damme, 1995), a property that enabled the development of the ABO system of blood typing. The second characteristic common to all lectins is the ability to bind carbohydrates selectively based on the individual sugar specifity of the lectin. This also results in lectins binding to the carbohydrate moieties of extracellular glycoconjugates specifically and reversibly without introducing conformational changes to the mono- or oligosaccharides to which they bind (Sharon & Lis, 2004). Plant lectins are involved in a wide range of processes including carbohydrate transport, cell-cell signalling, cell surface binding and recognition, pathogenic defence and in potentially mediating symbiotic relationships (Sharon & Lis, 2004). Plant lectins play a key role in defence, with many specifically binding to epithelial cells of herbivore and nematode guts (Schubert et al., 2012; Delatorre et al., 2007). Insecticidal, antifungal and antiviral qualities have also been widely described (e.g. Kumar et al., 1993; Rao et al., 1998; Peumans, Hao & Van Damme, 2001; Edwards & Gatehouse, 2007). For example, Balsamin, from Momordica balsamina, demonstrates potent anti-HIV activity (Kaur et al., 2013).

In recent years the potential of lectins for use in cancer therapies has become a significant research focus due to their ability to preferentially bind to specific carbohydrates, and differentiate between glycosylation patterns. Moreover, a number of plant derived lectins have been shown to have potent in vitro and in vivo anti-cancer effects (e.g. Voss et al., 2006; Otsuka et al., 2014) inducing autophagous and apoptotic pathways in tumour cells,

and some are already used therapeutically. For instance, the recombinant mistletoe lectin rViscumin has been through phase 1 clinical trials and a number of other native mistletoe lectin preparations such as Lektinol and Iscador are prescribed widely throughout Europe as adjuvant therapies although their efficacy is not readily quantified (Horneber et al., 2008).

Using a modified extraction protocol developed from previously published studies (Cammue, Peeters & Peumans 1985, Kumar et al. 1993, George et al. 2011). This paper focuses on the lectin found in the tubers of Winter Aconite, Eranthis hyemalis, (EHL). To date, EHL is the sole lectin representative described from the Ranunculaceae and has been identified as a Type II Ribosome Inactivating Protein (RIP) (Kumar et al., 1993). EHL preferentially binds N-acetyl-galactosamine, but also binds galactose, galacto-pyransosyl-D-glucose, and to a lesser degree D-ribose (Kumar et al., 1993). Type II RIPs are classified as chimerolectins and cytotoxic N-glycosidases and consist of either one or two heterodimers linked by disulphide bonds. The B chain subunit is a sugar specific lectin containing the highly conserved ricin B domain and will bind to extracellular glyconjugates. This mediates entry to the intracellular environment for the attached cytotoxic A chain by endocytosis (Virgilio et al., 2010). The toxin is then subjected to retrograde transport via the Golgi complex; to the endoplasmic reticulum where the disulphide bonds are reduced and the A chain is free in the cytosol to refold into an enzymatically active form (Hartley & Lord, 2004). The A chain acts as an inhibitor of eukaryotic protein synthesis by cleaving a single adenosine (A_{4234}) in the 28s rRNA subunit preventing Elongation Factor 2 from binding and resulting in immediate and absolute cessation of peptide elongation (Hartley & Lord, 2004).

Type II RIPs are an area of increasing interest due to their antineoplastic properties, and their glycomic binding profile can be used to target specific glycans of biological molecules. For instance, the GalNac specific RIP Ximenia americana (Riproximin) (Voss et al., 2006), Sambucas sp (Ferreras et al., 2011) and ML1 from Viscum album (Tonevitsky et al., 1996) show higher binding affinity for tumour cells than for healthy cells. This can be partly attributed to the expression of particular surface saccharide groups in the changing glycomics of malignant cells (Peumans, Hao & Van Damme, 2001; Bayer et al., 2012). The use of lectins and Type II RIPs is also indicated in some studies for marking metastatic proliferation of tumour cells due to excessive glycosylation in metastatic cell lines (Zhou et al., 2015).

Herein we present a modified extraction and purification protocol for EHL. Protein sequencing of EHL further supports the findings of earlier work that EHL is a Type II RIP. We also investigate the effect of EHL on the free living nematode Caenorhabditis elegans. C elegans is a well-established model organism for initial toxicological studies due to the conserved nature of its biological and biochemical processes including stress response and disease pathways (Boyd, Smith & Freedman, 2012). Our research has revealed a range of phenotypes including direct mortality and a constitutive dauer formation phenotype that is consistent with neuronal ablation.

Material and Methods

Preparation of affinity chromatography column

An Amersham chromatography column was packed at room temperature with a final bed volume of 8 ml of Fetuin-Agarose in solution with 0.5 M NaCl and immobilised on cross-linked 4 % beaded agarose (Sigma-Aldrich Company ltd, UK). Prior to use the column was equilibrated with 8 column volumes of Phosphate buffered solution (PBS).

Extraction of EHL

60 g of E. hyemalis tubers supplied by Eurobulbs Ltd (UK) were prepared using a modified method from those described in previous studies (Cammue, Peeters & Peumans, 1985, Kumar et al. 1993, George et al. 2011) with adaptions as follows. The tubers were finely sliced before being homogenised with 250 ml of ice cold PBS containing 5 mM thiourea and left to settle on ice for 30 minutes. The homogenate was removed and stored and the remaining slurry was mixed with a further 250 ml of PBS. The two fractions were then combined and stirred at 4°C for 4 hours. The homogenised mixture was then centrifuged (Sorvall RC6 plus HSC) at 20,000 g for 30 minutes. The supernatant was retained and frozen at -80°C overnight in order to induce aggregation of any remaining lipid content in sample. The sample was then defrosted and filtered through 3MM Whatman filter paper before undergoing a further centrifuge cycle of 20,000 g for 20 minutes. The clarified supernatant then underwent ammonium sulphate precipitation.

Ammonium sulphate precipitation.

Solid ammonium sulphate was added slowly to the crude extract initially to a saturation point of 40%, and after one hour of stirring at 4°C was centrifuged at 10,000g for 15 minutes. The pellet was re-suspended in 15mls of PBS. Ammonium sulphate was then added to the supernatant to a final saturation point of 80%, with a further hour of stirring at

4°C. The resulting pellet was also re-suspended in PBS. Agglutination activity was found to be retained in the 40% pellet and absent from the 80% pellet. An SDS-PAGE gel confirmed the presence of target protein in the 40% fraction. Samples were dialysed against PBS in 3 buffer changes consisting of 200 x sample volume each including a final overnight exchange.

Purification of EHL

The crude dialysed extract was applied to a Fetuin-agarose affinity chromatography column at a rate of 1 ml per minute using ÄKTA Express protein purification system (GE Healthcare), non-target proteins were then eluted with PBS until absorbance at λ_{280} was restored to base line value circa 40 mAU. The affinity matrix was then equilibrated with PBS and subsequently EHL was eluted with 40 mM 1,3 diaminopropanol (DAP) and peak fractions were collected in 0.5ml aliquots. The oligosaccharide structure of Fetuin has been well defined and shown to have Gal and GlcNAc branched residues present. Fetuin has been purified using the lectin RCA I which has specificity for Galactose and N-acetylgalactosamine (Green et al., 1988). Its use, therefore as an affinity chromatography media for the lectin/type II RIP purification is based on this complementary interaction. The eluant was neutralised with 2-Amino-2-hydroxymethyl-propane-1,3-diol hydrochloride (Tris-HCl) at pH 7.0. Peak fractions were applied to pre-equilibriated PD-10 desalting columns (GE Healthcare) and buffer exchanged into PBS.

Analysis of EHL

Purified EHL was tested for agglutination ability using defibrinated rabbit erythrocytes (TCS Biosciences), with 20 µl of post column eluant, or control, added to a 20 µl sample

of erythrocytes in a welled microscope slide. The purified EHL was also analysed by SDS-PAGE, with both reduced and non-reduced samples electrophoresed on 12 % gels and subsequently stained with Coomassie Brilliant Blue. Concentration was measured using a Qubit flourometric protein assay.

Purified EHL was commercially sequenced. In-gel tryptic digestion was performed after reduction with DTE and S-carbamidomethylation with iodoacetamide. Gel pieces were washed two times with 50% (v:v) aqueous acetonitrile containing 25 mM ammonium bicarbonate, then once with acetonitrile and dried in a vacuum concentrator for 20 min. Sequencing-grade, modified porcine trypsin (Promega) was dissolved in the 50 mM acetic acid supplied by the manufacturer, then diluted 5-fold with 25 mM ammonium bicarbonate to give a final trypsin concentration of 0.02 μg/μL. Gel pieces were rehydrated by adding 10 μL of trypsin solution, and after 10 min enough 25 mM ammonium bicarbonate solution was added to cover the gel pieces. Digests were incubated overnight at 37ºC.

A 1 μL aliquot of each peptide mixture was applied to a ground steel MALDI target plate, followed immediately by an equal volume of a freshly-prepared 5 mg/mL solution of 4-hydroxy-α-cyano-cinnamic acid (Sigma) in 50% aqueous (v:v) acetonitrile containing 0.1% , trifluoroacetic acid (v:v).

Positive-ion MALDI mass spectra were obtained using a Bruker ultraflex III in reflectron mode, equipped with a Nd:YAG smart beam laser. MS spectra were acquired over a range of 800-5000 m/z. Final mass spectra were externally calibrated against an adjacent spot containing 6 peptides (des-Arg^1-Bradykinin, 904.681; Angiotensin I, 1296.685; Glu^1-

Fibrinopeptide B, 1750.677; ACTH (1-17 clip), 2093.086; ACTH (18-39 clip), 2465.198; ACTH (7-38 clip), 3657.929.). Monoisotopic masses were obtained using a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) and a S/N threshold of 2.

For each spot the ten strongest precursors, with a S/N greater than 30, were selected for MS/MS fragmentation. Fragmentation was performed in LIFT mode without the introduction of a collision gas. The default calibration was used for MS/MS spectra, which were baseline-subtracted and smoothed (Savitsky-Golay, width 0.15 m/z, cycles 4); monoisotopic peak detection used a SNAP averagine algorithm (C 4.9384, N 1.3577, O 1.4773, S 0.0417, H 7.7583) with a minimum S/N of 6. Bruker flexAnalysis software (version 3.3) was used to perform spectral processing and peak list generation.

De novo sequencing of tandem mass spectra was performed by hand, with a-, b-, b^0-, y-, y^0- and y*-ions considered as possible fragment ions. De novo derived peptides sequences were matched to homologous protein sequences using the on-line MS-BLAST service provided by Washington University. The results of which were consistent with the in-house homology search results conducted on confidently assigned sequences using the University of Virginia UVa FASTA Server, the FASTS and SSearch algorithms were used for homology searching against the SwissProt (NCBI) and PDB databases.

Activity against C. elegans

Worms were obtained from the Caenorhabditis Genetics Center and maintained using standard methods (Stiernagle, 2006), on nematode growth media plates (NGM) using Escherichia coli OP50 strain food source. N2 was used for initial testing and as a control

in other assays. To assess the effect of various mutations on EHL-induced dauer larvae formation, the following strains were used: CX2065, odr-1(n1936); CX2205 odr-3(n2150);

PR671, tax-2(p671); PR672, che-1(p672); PR813, osm-5(p813); SP1205, dyf-1(mn335); and SP1709, dyf-10(e1383). In all experiments, treatments and genotypes were blind coded, the position of plates within experimental blocks was randomised, and contaminated plates excluded from all analysis.

For all assays, arrested and synchronised C. elegans first stage larvae (L1s) were obtained by allowing eggs, isolated from gravid hermaphrodites by hypochlorite treatment (Stiernagle, 2006), to hatch on NGM plates in the absence of food for 24 hours at 20°C. For experiment 1, arrested N2 L1s were washed from plates, resuspended in M9 with a series of EHL concentrations from 3.92 to 0 mg/ml, incubated at 20°C for 6 hours, and 15 worms per treatment were picked for analysis. For experiment 2, arrested N2 L1s treated as above except treatments were 2.94 mg/ml and 0 mg/ml EHL and a greater number of worms per treatment were analysed (n = 55 and n = 33 for the 2.94 and 0 mg/ml treatments, respectively). After incubation, worms were washed 3 times in water. For the analysis of development and fecundity (experiments 1 and 2), worms were transferred in a small volume of liquid to NGM plates without food, then individually transferred from this plate to NGM plates with E. coli OP50 strain food source and maintained at 20°C. Standard methods were then used to analyse the reproductive schedule and lifetime fecundity (Hodgkin & Doniach, 1997). These data were then used to assess the effect of EHL on reproduction as assessed by lifetime reproductive success (LRS), the total number of progeny produced (experiments 1 and 2), and the intrinsic rate of increase (r)

(experiment 2), calculated by iteration from $\Sigma e^{-rx} l_x m_x = 1$, where l_x represents the age specific survivorship to day x and m_x represents the fecundity on day x (Vassilieva & Lynch, 1999).

Based on phenotypes observed in the initial screen, the ability of EHL to induce constitutive dauer larvae formation (a dauer-constitutive, or Daf-c, phenotype) was investigated in greater detail. Here, worms were treated with 0.98, 1.96 or 2.94 mg/ml EHL, as described above, except that after washing, worms were transferred en masse to plates with food. After four days at 20°C plates were visually scored to assess the proportion of worms that had developed as dauer larvae (number of dauer larvae/total number of worms). After counting, worms were washed from plates and incubated in 1 % SDS for one hour, a treatment that kills all C. elegans stages except dauer larvae (Cassada & Russell, 1975), worms were washed once in M9, transferred to fresh NGM plates with food and the number of dauer larvae again counted. These dauer larvae were then transferred individually to NGM plates with food at 20°C and monitored for the next 14 days to determine if they were capable of resuming development. To further analyse EHL-induced dauer larvae formation, N2 and mutant worms were treated, as described above, with 0, or 1.54 mg/ml EHL, washed, and transferred en masse to plates (n = 3 per combination of treatment and genotype) with food. Plates were then incubated at 20°C for four days at which time the proportion of worms that had developed as dauer larvae was scored.

Dauer larvae formation in PR672, che-1(p672), was further analysed both in standard dauer larvae formation assays and in assays of growing populations. For assays of dauer

larvae formation in response to defined amounts of pheromone, assays were performed as previously described (Golden & Riddle 1984; Green et al., 2014), with worms allowed to lay eggs on assay plates containing dauer pheromone extract and limited amounts of food, and progeny scored after two days at 25°C. Dauer larvae formation in growing populations was assessed as previously described (Green et al., 2013), with populations initiated with single worms and a defined amount of food allowed to grow to food exhaustion, except that assays were performed at 25°C.

Results

Purification and characterisation of EHL

Qubit fluorometric measurement showed a typical concentration of 2.5 mg/ml using our revised and improved purification strategy. This shows an approximately 5 to 6-fold increase in recovery in comparison to the previously reported yield of 380 µg/ml in George et al. (2011). Non-reducing SDS-PAGE analysis produced characteristic reduced protein bands at circa 31 and 28 kDa as well as an unreduced band circa 50kDa (Fig. 1); these values are consistent with those previously reported in the literature (Kumar et al., 1993). An intense agglutination response of rabbit erythrocytes was exhibited and thus confirmed the presence of EHL (Fig. 2).

De novo sequencing analysis.

De novo sequence analysis of fragmentation spectra was carried out and suggested two confidently assigned peptide fragment sequences, with the following tags, QQWA[L/I]YSDST[L/I]R and NWNNNGNP[L/I]Q[L/I]WQCTQQQNQR (Fig. 3). Peptide fragment homology searches produced matches to various Type II Ribosome

Inactivating Proteins (RIPS) all within ricin-b domain (carbohydrate binding) regions. This result confirms the status of EHL as a Type II RIP, as has been reported in previously published sequence data (Kumar et al., 1993). Sequence tags QQWA[L/I]YSDST[L/I]R and NWNNNGNP[L/I]Q[L/I]WQCTQQQNQR were matched to two regions in Nigrin-b (SNA-V; Sambucas nigra agglutinin-V) UniProtKB P33183.2 correlating to residues 470-481 and 325-345 respectively using the FASTS search facility against the SwissProt database. Close homology to other Type II RIPS was also matched within the ricin b domain of Abrin-a (P11140) and Ricin (P02879). PDB structures were searched using SSearch algorithm and also showed matches to SNA-II (3C9Z) (Sambucas nigra agglutinin II), Abrin-a (1ABRB) and Ricin (1RZOB) in the ricin b domain as well as ML1 from Viscum album (1QNKB).

EHL affects survival, development and reproduction in C. elegans

Acute treatment of arrested C. elegans L1s for 6 hours with EHL at different concentrations resulted in a subsequent range of developmental, fertility and survival defects (Fig. 4), with all EHL concentrations reducing lifetime reproductive success (LRS) (Fig. 4A, pairwise, Bonferroni corrected, Mann–Whitney U tests against N2 showing reduced fecundity in all EHL treatments). Much of this decrease in fecundity is however a consequence of worms not reproducing in the EHL treatments (comparison of Fig. 4A and B), although LRS does still decrease over the range of EHL concentrations tested when only those worms that reproduced are considered (Fig. 4B, Pearson product-moment correlation of LRS against EHL concentration: including the 0 mg/ml group, $r = -0.52$, $p < 0.001$; excluding the 0 mg/ml group, $r = -0.38$, $p = 0.013$). EHL-treated worms that did reproduce showed a delay in development, with many treated worms starting

reproduction a day or more after the control worms. Of the worms that did not reproduce, some showed no movement from the point at which they were placed on the plate and no response to stimulus after 24 hours and therefore died as a consequence of the EHL treatment. Other EHL-treated worms were observed to remain as arrested L1s or to develop as dauer larvae, a non-feeding developmentally arrested stage. Many of the non-reproducing worms were found not on a food source, and more than would be expected under these conditions were found to have climbed the sides of the plate; both behaviours are indicative of a disruption to chemosensory ability and an inability to detect the bacterial food.

To investigate these reproductive effects and the survival in more detail, a larger number of worms were assayed (Fig. 5). Here, EHL treatment resulted in immediate mortality of 41% of EHL treated individuals and again EHL-treated worms were observed to remain as L1s and to arrest as dauer larvae. These results suggest that EHL treatment affects the sensory neurons. Overall, EHL treatment reduced subsequent LRS (Fig. 5A, control vs all EHL treated worms, $W = 2376.0$, $p < 0.001$), with the subset of EHL treated worms that did reproduce producing a greatly reduced number of progeny (control vs reproducing EHL treated worms, $W = 792.0$, $p < 0.001$). A similar pattern was observed in the analysis of the effects of EHL treatment on the estimated rate of increase (Fig. 5B, control vs all EHL treated worms, $W = 2376.0$, $p < 0.001$; control vs reproducing EHL treated worms, $W = 792.0$, $p < 0.001$).

To further characterise the development of EHL-treated worms as dauer larvae, an additional set of arrested L1s were analysed. As in the assay for reproductive effects

(above), some of these EHL treated worms developed as dauer larvae (Fig. 6). These worms were then SDS treated, with survival confirming that they were indeed dauer larvae. Fifty of these dauer larvae were then transferred to plates with food and maintained at 20°C, with only one worm out of the fifty recovering and completing development as a reproductive adult after four days, and a second recovering after a total of fourteen days. Under these conditions dauer larvae normally recover rapidly and would be expected to have commenced reproduction approximately 2 days after transfer to food (Green & Harvey, 2012).

Analysis of mutant strains indicates that the ability of EHL treatment to induce dauer larvae formation varied across the genotypes (Table 1). All genotypes were however observed to show a developmental delay in response to EHL treatment, with reproduction of most EHL-treated worms not commencing until day 5 after treatment. Analysis of dauer larvae formation in PR672, che-1(p672) showed that this mutation does not block dauer larvae formation in response to defined amounts of pheromone and that similar numbers of dauer larvae are formed in N2 and PR672 in growing populations ($F_{1,33} = 0.05$, $p = 0.82$).

Discussion

We have successfully isolated the type II RIP found in the tubers of the Winter Aconite, E. hyemalis, by modification of a previously published protocol (Cammue, Peeters & Peumans, 1985; Kumar et al., 1993). Analysis indicates EHL is a heterodimeric protein consisting of two chains of molecular weights of approximately 28 and 31 kDa (Fig. 1).

Protein sequencing confirms that EHL is a Type II RIP with the cytotoxic potential for depurination of eukaryotic ribosomes.

EHL was used to study potential lectin-mediated toxicity against C. elegans. The bioassays performed indicate that EHL has biocidal properties against C. elegans. Four phenotypic effects were identified: reduced fecundity (Fig. 4 and 5), developmental delay, chemosensory disruption and constitutive dauer formation (Fig. 6 and Table 1). C. elegans physiology is such that at the arrested L1 larval stage, the only cells which are not enclosed by a largely impermeable cuticle are the amphids and phasmids. These are bilaterally symmetrical sensory organs that contain the sensory neurons: each amphid containing twelve neurons and each phasmid containing two (Ward et al., 1975). Of the twelve amphid neurons, the ciliated nerve endings of eight are exposed to the external environment via the amphid pore (Ward et al., 1975). These neurons control a range of phenotypes, including egg-laying and the decision to develop as a dauer larvae (Albert, Brown & Riddle, 1981). Laser ablation of the ASI, ADF and ASG cells is sufficient to result in constitutive dauer larvae development, with ablation of the ASJ cell resulting in an inability to recover from dauer arrest (Bargmann & Horvitz, 1991). Our observations of inappropriate dauer larvae formation and the failure of most such dauer larvae to resume development indicate that EHL is interacting with these neurons and are consistent with EHL resulting in neuron death.

In general, biocidal assays with lectins involve ingestion of the lectin by the target organism. For example, EHL had previously been tested against the coleopteran pest Diabrotica undecimpunctata howardii, resulting in a high mortality rate and an 80%

reduction in body size of survivors; there was however no previous data on reproductive effects (Kumar et al., 1993). A study of the toxic effects of the CCL2 lectin from Coprnopsis cinerea (Ink Cap mushroom) on C. elegans reported a phenotype of severe developmental delay; the lectin was adsorbed in the epithelial cells of the intestine, potentially degrading the membrane and preventing growth (Schubert et al., 2012). The absence of a food source in our assay has therefore enabled the observation of entirely new lectin-mediated C. elegans phenotypes induced by EHL, including a Daf-c phenotype which has not been reported before.

The cause of the developmental delay in EHL-treated worms is not clear. Possibilities would include damage to the pharynx and a subsequent reduction in pumping (feeding) ability, or, if some feeding is initiated, damage to the epithelial cells of the intestine as observed in response to the C. cinerea CCL2 lectin (Schubert et al., 2012). A further possibility is that differences in body size and development are also a consequence of damaged neurons (see Fujiwara, Sengupta & McIntire, 2002).

Wild-type C. elegans take up dyes such as DiI and FITC into the amphid neurons AWB, ASH, ASJ, ASK, ADL and ASI (Hedgecock et al., 1985). Given the likely mode of action of EHL, we reasoned that mutations that disrupt the normal formation of sensory amphids would block EHL-induced dauer larvae formation. Consistent with this, disruption of osm-5, dyf-1 and dyf-10, all mutations in which the amphid neurons cannot take up dyes, result in no EHL-induced dauer larvae formation (Table 1). In both odr-1 and odr-3 mutants, where dye filling is not affected and EHL would be expected to be able to access the neurons normally, there is no reduction in dauer larvae formation in response to EHL

treatment (Table 1). In contrast, dauer larvae formation is reduced in PR671, but some are still formed (Table 1), indicating that disruption of tax-2 only partially blocks the effect. TAX-2 forms, with TAX-4, a cyclic nucleotide-gated cation channel that is required for chemotaxis in response to AWC-sensed odorants (Coburn & Bargman, 1996). Axon outgrowth defects have however been noted in tax-2 mutants, with c. 80% of tax-2(p671) animals observed to have abnormal ASJ axons (Coburn & Bargman, 1996). It is not clear if the reduction in the EHL-induced dauer larvae formation observed in the tax-2 mutants is a consequence of the axon guidance defects or the channel disruption.

That EHL-induced dauer larvae formation is also blocked in che-1 mutants further supports the hypothesis that EHL is disrupting neurons. CHE-1 is a C2H2-type zinc-finger transcription factor that is required for the identity of ASE neurons (Uchida et al., 2003). Loss of CHE-1 expression eliminates the function of ASE neurons and che-1 mutations have previously been shown to suppress Daf-c phenotypes (Reiner et al., 2008). No significant structural defects have been observed in che-1 mutants (Lewis & Hodgkin, 1977) and our results indicate that dauer larvae formation does not appear altered in either standard dauer larvae assays or in growing populations. This thereby implies that the mutation is specifically blocking EHL-induced constitutive dauer larvae formation.
It is well established that lectins bind to glyconjugates on cell surfaces and that toxicity in RIPs is due to lectin mediated entry to the cell; this mode of action is consistent with the results presented here. In the case of C. elegans the only cells exposed are the amphid neurons. As a Type II RIP EHL can be subject to retrograde transport from the cell surface along the neuronal processes, at which point the ribosomes are inactivated, causing translation to cease (Wiley, Blessing & Reis, 1982). As no post-embryonic somatic

division occurs in mature individuals, and multiple chemoreceptors are expressed in a single neuron, ribosome inactivation of the neurons within the amphids would affect many functions derived from chemosensation (Sulston & Horvitz, 1977). Toxicity variables can be attributed to differing carbohydrate specificities but there is also evidence of the role of individual cell types in how they interact with lectins, indicating that any effects are characteristic of both variables (Battelli et al., 1997).

In conclusion, successful extraction using affinity chromatography has enabled assays to be conducted for biocidal properties against C. elegans. The results obtained demonstrate a significant reduction in fecundity, development, growth and a high incidence of abnormal dauer development when arrested L1 larvae were treated in the absence of food. The occurrence of dauer formation and a failure to recover in the presence of food supports the hypothesis that EHL is binding specifically to amphid neurons. Mutant screening has demonstrated that EHL can act as a neuronally specific cytotoxin, an effect which has previously been described with Ricin and other RIPs on mammalian sensory neurons (Wiley, Blessing & Reis, 1982; Tong et al., 2012). Further studies will aim to determine if those individuals that remained as arrested L1s were doing so as a consequence of an inability to perceive the food or if an additional mechanism is at work.

Our research shows that EHL has biocidal and potential cytotoxic activity. Moreover, EHL shows specificity for GalNac, an overexpressed sugar in the Tn (GalNac clustered) antigen which characterises cancer linked O-glycans (Ju, Otto & Cummings, 2011). Other GalNac Type II RIPs such as Mistletoe Lectin (ML1) and Riproximin have demonstrated promising therapeutic relevance as anticancer agents (Voss et al., 2006; Bayer et al., 2012;

Adwan et al., 2014. These factors suggest that EHL is a viable candidate for further study in respect of antineoplastic characteristics.

Acknowledgments

With thanks to Adam Dowle at The Proteomics laboratory at the Bioscience Technology Facility, Department of Biology, University of York.

References

Adwan H, Bayer H, Pervaiz A, Sagini M, & Berger M R. (2014). Riproximin is a recently discovered type II ribosome inactivating protein with potential for treating cancer. Biotechnology advances 32(6):1077-1090.

Albert PS, Brown SJ, Riddle DL. 1981. Sensory control of dauer larva formation in Caenorhabditis elegans. Journal of Comparative Neurology 198:435-451.

Bargmann CI, Horvitz HR. 1991. Control of larval development by chemosensory neurons in Caenorhabditis elegans. Science 251:1243-1246.

Bargmann CI, Hartweig EA, Horvitz HR. 1993. Odorant-selective genes and neurons mediate olfaction in C. elegans. Cell 74:515-27.

Battelli MG, Barbieri L, Bolognesi A, Buonamici L, Valbonesi P, Polito L, van Damme E, Peumans W, Stirpe F. 1997. Ribosome-inactivating lectins with polynucleotide: Adenosine glycosidase activity. FEBS Letters 408:355-359.

Bayer H, Essig K, Stanzel S, Frank M, Gildersleeve JC, Berger MR, Voss C. 2012. Evaluation of Riproximin binding properties reveals a novel mechanism for cellular targeting. Journal of Biological Chemistry 287:35873-35886.

Boyd WA, Smith MV, Freedman JH. 2012. Caenorhabditis elegans as a model in developmental toxicology. Methods in Molecular Biology 889:15-24.

Cammue B, Peeters B, Peumans W. 1985. Isolation and partial characterization of an N-acetylgalactosamine-specific lectin from winter-aconite (Eranthis hyemalis) root tubers. Biochemical Journal 227:949.

Cassada RC, Russell RL. 1975. The dauer larva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans. Developmental Biology 46:326-342.

Coburn CM, Bargmann CI. 1996. A putative cyclic nucleotide-gated channel is required for sensory development and function in C. elegans. Neuron 17:695-706.

Delatorre P, Rocha BA, Souza EP, Oliveira TM, Bezerra, GA, Moreno FB, Azevedo WF 2007. Structure of a lectin from Canavalia gladiata seeds: new structural insights for old molecules. BMC Structural Biology 7:52.

Dusenbery DB. 1976. Chemotactic behavior of mutants of the nematode C. elegans that are defective in their attraction to NaCl. Journal of Experimental Zoology 198:343-352.

Dusenbery DB. 1980. Chemotactic behavior of mutants of the nematode C. elegans that are defective in osmotic avoidance. Journal of Comparative Physiology 137:93-96.

Edwards MG, Gatehouse AM. 2007. Biotechnology in crop protection: Towards sustainable insect control. In: Novel biotechnologies for biocontrol agent enhancement and management. Netherlands: Springer.

Ferreras JM, Citores L, Iglesias R, Jiménez P, Girbés T. 2011. Use of ribosome-inactivating proteins from Sambucus for the construction of immunotoxins and conjugates for cancer therapy. Toxins 3:420-441.

Fujiwara M, Sengupta P, McIntire SL. 2002. Regulation of body size and behavioral state of C. elegans by sensory perception and the EGL-4 cGMP-dependent protein kinase. Neuron 36:1091-1102.

George O, Solscheid C, Bertolo E, Lisgarten D. 2011. Extraction and purification of the lectin found in the tubers of Eranthis hyemalis (winter aconite). Journal of Integrated OMICS 1(2): 268-272.

Golden JW, Riddle DL (1984) The Caenorhabditis elegans dauer larva: developmental effects of pheromone, food, and temperature. Developmental Biology 102: 368–378.

Green ED, Adelt G, Baenziger JU, Wilson S, Van Halbeek H. 1988. The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-released oligosaccharides by 500-megahertz 1H NMR spectroscopy. Journal of Biological Chemistry 263(34):18253-18268.

Green JWM, Harvey SC. 2012. Development of Caenorhabditis elegans dauer larvae in growing populations. Nematology 14:165-173.

Green JWM, Snoek LB, Kammenga JE,Harvey SC. 2013. Genetic mapping of variation in dauer larvae development in growing populations of Caenorhabditis elegans. Heredity 111:306-313.

Green JWM, Stastna JJ, Orbidans HE, Harvey SC. 2014. Highly polygenic variation in environmental perception determines dauer larvae formation in growing populations of Caenorhabditis elegans. PLoS One. 9(11): e112830.

Hartley M, Lord J. 2004. Cytotoxic ribosome-inactivating lectins from plants. Biochimica Et Biophysica Acta - Proteins and Proteomics 1701:1-14.

Hedgecock EM, Culotti JG, Thomson JN, Perkins LA. 1985. Axonal guidance mutants of Caenorhabditis elegans identified by filling sensory neurons with fluorescein dyes. Developmental Biology 111:158-170.

Hodgkin J, Doniach T. 1997. Natural variation and copulatory plug formation in Caenorhabditis elegans. Genetics 146:149-164.

Horneber MA, Bueschel G, Huber R, Linde K, Rostock M. 2008. Mistletoe therapy in oncology. Cochrane Database of Systematic Reviews 2:CD003297.

Ju T, Otto V I, Cummings R D. (2011). The Tn antigen—structural simplicity and biological complexity. Angewandte Chemie International Edition 50(8):1770-1791.

Kaur I, Puri M, Ahmed, Z, Blanchet FP, Mangeat B, Piguet V. 2013. Inhibition of HIV-1 Replication by Balsamin, a Ribosome Inactivating Protein of Momordica balsamina. PloS One 8:e73780.

Kumar MA, Timm D, Neet K, Owen W, Peumans WJ, Rao AG. 1993. Characterization of the lectin from the bulbs of Eranthis hyemalis (winter aconite) as an inhibitor of protein synthesis. Journal of Biological Chemistry 268:25176-25183.

Lewis JA, Hodgkin JA. 1977. Specific neuroanatomical changes in chemosensory mutants of the nematode Caenorhabditis elegans. Journal of comparative Neurology 172(3):489-510.

Otsuka H, Gotoh Y, Komeno T, Ono T, Kawasaki Y, Iida N, Shibagaki Y, Hattori S, Tomatsu M, Akiyama H, Tashiro F. 2014. Aralin, a type II ribosome-inactivating protein from Aralia elata, exhibits selective anticancer activity through the processed

form of a 110-kDa high-density lipoprotein-binding protein: a promising anticancer drug. Biochemical and Biophysical Research Communications 453:117-123.

Peumans WJ, Van Damme E. 1995. Lectins as plant defense proteins. Plant Physiology 109:347.

Peumans WJ, Hao Q, Van Damme EJ. 2001. Ribosome-inactivating proteins from plants: more than RNA N-glycosidases? FASEB Journal 15:1493-1506.

Reiner DJ, Ailion M, Thomas JH, Meyer BJ. 2008. C. elegans anaplastic lymphoma kinase ortholog SCD-2 controls dauer formation by modulating TGF-β signaling. Current Biology 18(15):1101-1109.

Rao K, Rathore KS, Hodges TK, Fu X, Stoger E, Sudhakar D, Bown DP. 1998. Expression of snowdrop lectin (GNA) in transgenic rice plants confers resistance to rice brown planthopper. Plant Journal 15:469-477.

Schubert M, Bleuler-Martinez S, Butschi A, Wälti MA, Egloff P, Stutz K, Yan S, Wilson IBH, Hengartner MO, Aebi M, Allain, FHT, Künzler M. 2012. Plasticity of the β-trefoil protein fold in the recognition and control of invertebrate predators and parasites by a fungal defence system. PLoS Pathogens 8:e1002706.

Sharon N, Lis H. 2004. History of lectins: From hemagglutinins to biological recognition molecules. Glycobiology 14:53-62.

Starich TA, Herman RK, Kari CK, Schackwitz WS, Schuyler MW, Collet J, Thomas JH, Riddle DL. 1995. Mutations affecting the chemosensory neurons of Caenorhabditis elegans. Genetics 139:171-188.

Stiernagle T. 2006. Maintenance of C. elegans. WormBook ed. The C. elegans Research Community, WormBook, doi/10.1895/wormbook.1.101.1, http://www.wormbook.org.

Sulston JE, Horvitz H. 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Developmental Biology 56:110-156.

Tonevitsky A, Agapov I, Shamshiev A, Temyakov D, Pohl P, Kirpichnikov M. 1996. Immunotoxins containing A-chain of Mistletoe lectin I are more active than immunotoxins with ricin A-chain. FEBS Letters 392:166-168.

Tong WM, Sha O, Ng TB, Cho EY, Kwong WH. 2012. Different in vitro toxicity of ribosome-inactivating proteins (RIPs) on sensory neurons and Schwann cells. Neuroscience letters 524(2):89-94.

Uchida O, Nakano H, Koga M, Ohshima Y. 2003. The C. elegans che-1 gene encodes a zinc finger transcription factor required for specification of the ASE chemosensory neurons. Development 130:1215-1224.

Vassilieva LL, Lynch M. 1999. The rate of spontaneous mutation for life-history traits in Caenorhabditis elegans. Genetics 151:119-129.

Virgilio MD, Lombardi A, Caliandro R, Fabbrini MS. 2010. Ribosome-inactivating proteins: From plant defense to tumor attack. Toxins 2:2699-2737.

Voss C, Eyol E, Frank M, von der Lieth C W, Berger MR. 2006. Identification and characterization of riproximin, a new type II ribosome-inactivating protein with antineoplastic activity from Ximenia americana. FASEB Journal 20:1194-1196.

Ward S, Thomson N, White JG, Brenner S. 1975. Electron microscopical reconstruction of the anterior sensory anatomy of the nematode Caenorhabditis elegans. Journal of Comparative Neurology 160:313-337.

Wiley RG, Blessing W, Reis DJ. 1982. Suicide transport: Destruction of neurons by retrograde transport of ricin, abrin, and modeccin. Science 216:889-890.

Zhou SM, Cheng L, Guo SJ, Wang Y, Czajkowsky DM, Gao H, Hu XF, Tao SC. 2015. Lectin RCA-I specifically binds to metastasis-associated cell surface glycans in triple-negative breast cancer. Breast Cancer Research 17(1):36.

Appendix 1.2 Published figures and legends

		Mean % dauer ± S.E.		
Isolate	Genotype	Control	Lectin	Phenotype
N2	Wild-type	0	49 ± 2.5	-
CX2065	odr-1(n1936) X	0	53 ± 6.1	Defective chemotaxis[1]
CX2205	odr-3(n2150) V	0	35 ± 3.1	Defective chemotaxis[1] Defective osmotic avoidance[1]
PR671	tax-2(p671) I	0	8 ± 1.9*	Defective chemotaxis[2] Defective thermotaxis[2]
PR672	che-1(p672) I	0	0*	Defective chemotaxis[2]
PR813	osm-5(p813) X	0	0*	Defective dye filling[3] Defective osmotic avoidance[3]
SP1205	dyf-1(mn335) I	0	0*	Defective dye filling[4] Defective chemotaxis[4]
SP1709	dyf-10(e1383) I	0	0*	Defective dye filling[4] Defective chemotaxis[4]

Mutations affecting amphid structure can block EHL-induced dauer larvae formation. The mean dauer larvae formation found on three plates of control and 80% (1.548mg/ml) EHL treated worms of differing genotypes of C. elegans. *Denotes genotypes where the proportion of dauer larvae observed in the EHL treatment differs from that observed in N2 ($p < 0.05$, Fisher's Exact Test with Bonferroni adjustment to correct for multiple testing). [1]Bargmann, Hartweig & Horvitz (1993). [2]Dusenbery (1976). [3]Dusenbery (1980). [4]Starich et al.(1995).

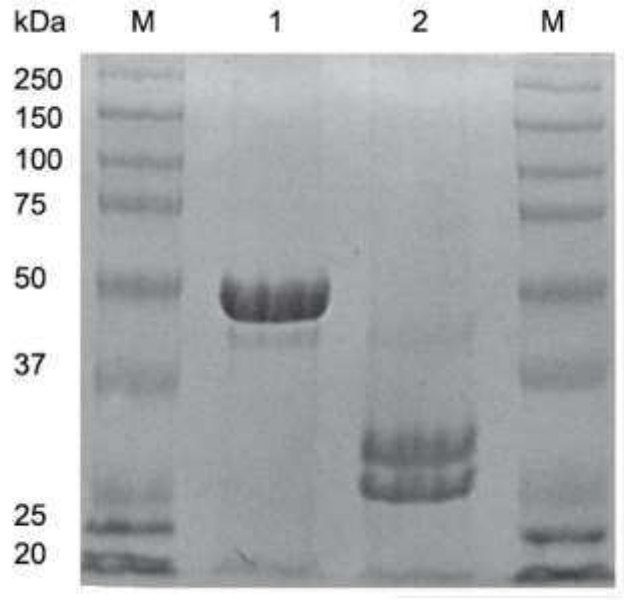


Figure 1: EHL is a heterodimeric Type II Ribosome Inactivating Protein. Coomassie Blue stained 12% SDS-PAGE analysis of EHL showing molecular weight marker (M), non-reduced EHL (1) and β -mercaptoethanol reduced EHL (2).

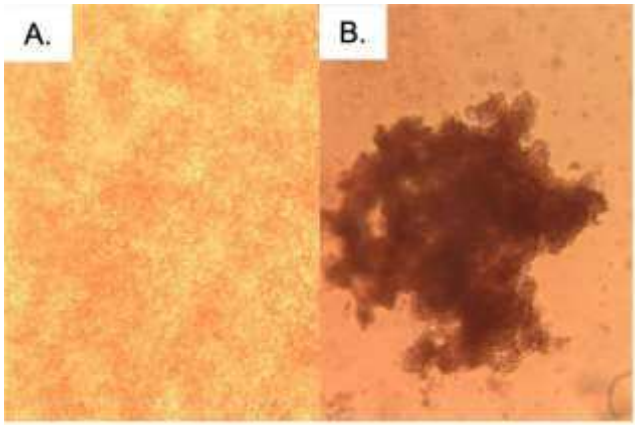


Figure 2: EHL induces agglutination of erythrocytes. (A) Control, with no clumping of erythrocytes observed, and (B) incubation with EHL, which results in agglutination.

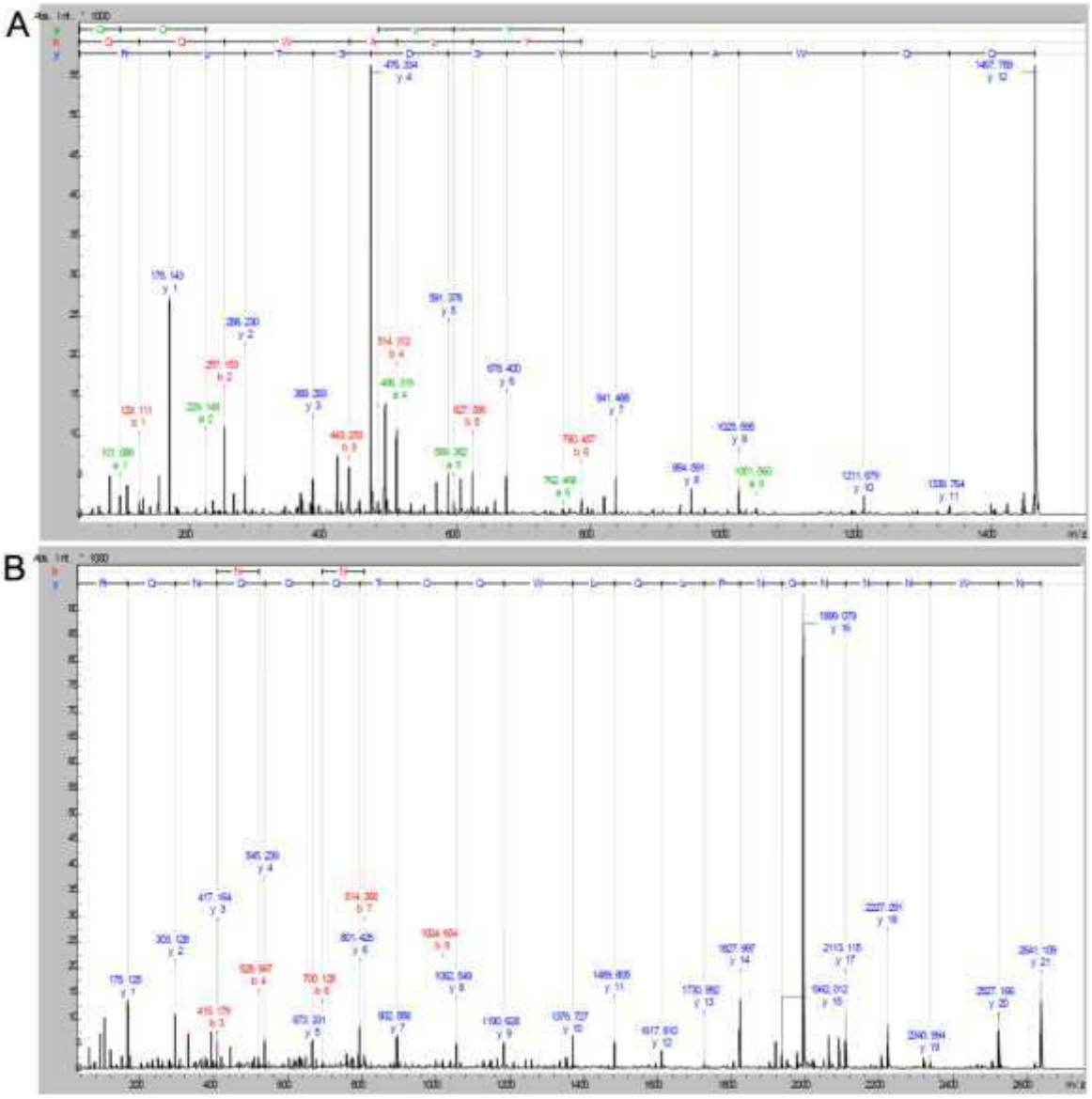


Figure 3: De novo peptide sequencing based on the MALDI-TOF MS/MS fragmentation spectra of EHL in-gel tryptic digestion after reduction with DTE and S-carbamidomethylation with iodocetamide. (A) QQWA(L/I)YSDST(L/I)R, the most confidently assigned peptide with a good level of overlapping y- and b- ion series. (B) NWNNNGNP(L/I)Q(L/I)WQCTQQQNQR, a strong y- ion series is observed throughout the peptide

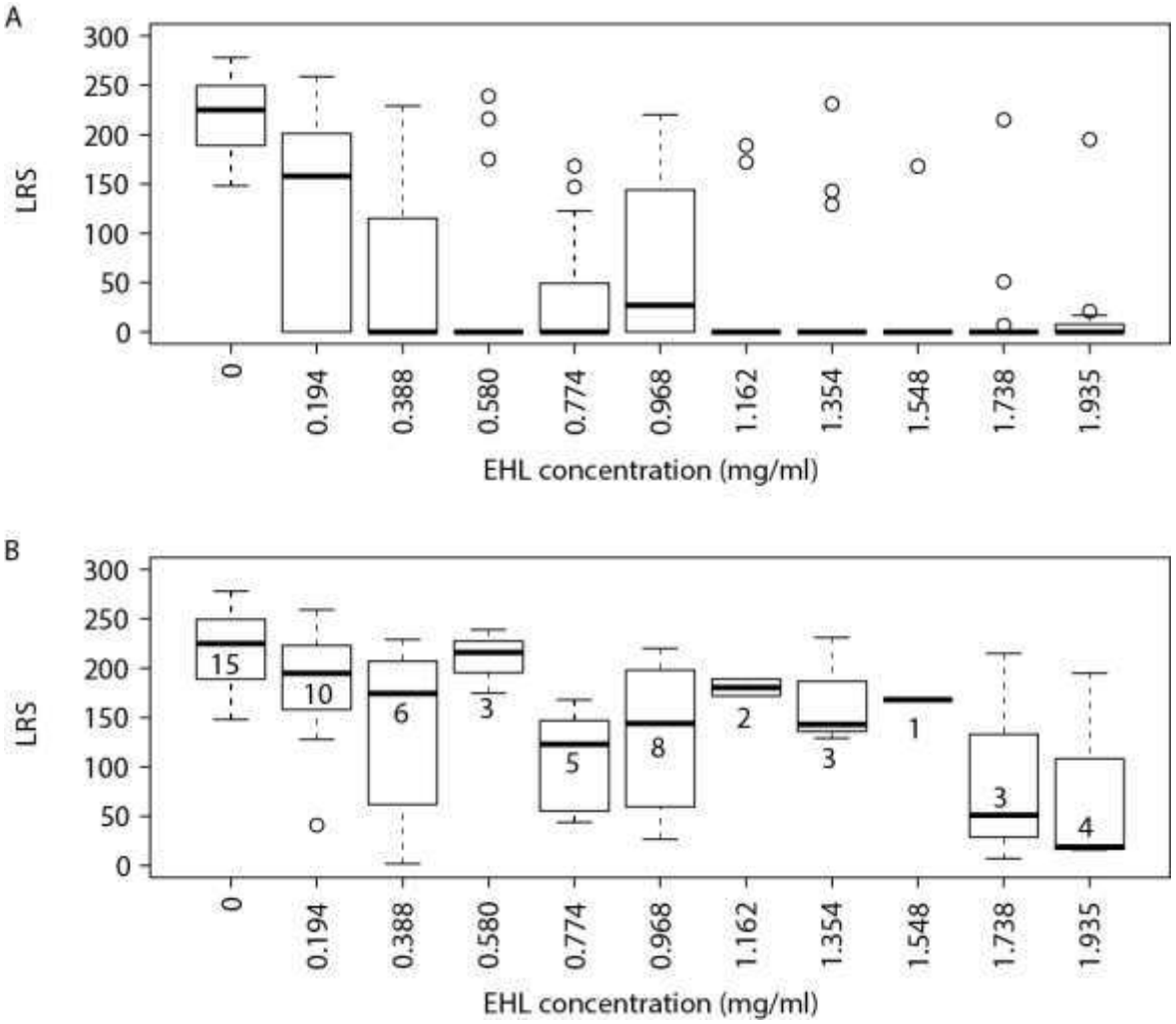


Figure 4: EHL reduces fecundity in C. elegans. Box plots of Lifetime reproductive success (LRS) of (A) EHL treated worms and (B) the subset of EHL treated worms that reproduced for a range of lectin concentrations. For (A) n=15 for all treatments, for (B) numbers associated with the boxes denote the sample sizes

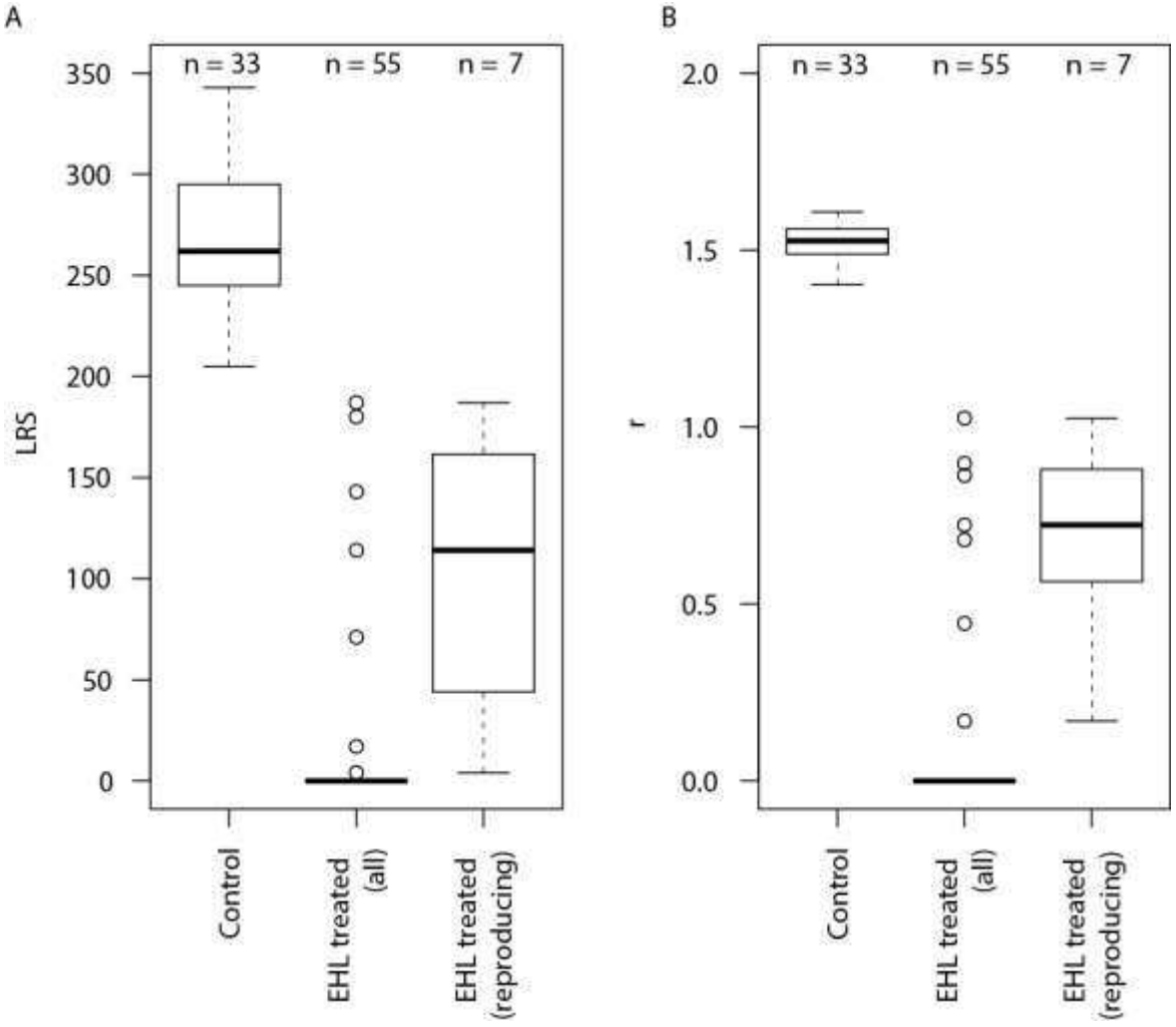


Figure 5: EHL reduces fecundity and slows development in C. elegans. Boxplots of (A) lifetime reproductive success, LRS, and (B) intrinsic rate of increase, r, of control worms, all EHL treated worms and of the subset of EHL treated worms that reproduced. n denotes the sample sizes

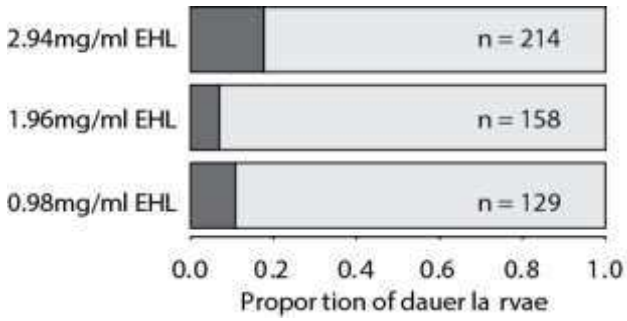


Figure 6: EHL treatment induces dauer larvae formation in C. elegans. The proportion of larvae that developed into dauer larvae observed in populations in three different EHL treatments. In the abscence of EHL treatment no dauer formation is observed under these experimental conditions in N2

Appendix 2 Crystal screen chemical conditions

Morpheus	Conc	Units	Salt	Conc	Units	Buffer2	pH	Conc3	Units3	Precipitant3
A1	0.06	M	Divalents	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
A2	0.06	M	Divalents	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
A3	0.06	M	Divalents	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
A4	0.06	M	Divalents	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
A5	0.06	M	Divalents	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
A6	0.06	M	Divalents	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
A7	0.06	M	Divalents	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
A8	0.06	M	Divalents	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
A9	0.06	M	Divalents	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
A10	0.06	M	Divalents	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
A11	0.06	M	Divalents	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
A12	0.06	M	Divalents	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4
B1	0.09	M	Halogens	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
B2	0.09	M	Halogens	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
B3	0.09	M	Halogens	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
B4	0.09	M	Halogens	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
B5	0.09	M	Halogens	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
B6	0.09	M	Halogens	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
B7	0.09	M	Halogens	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
B8	0.09	M	Halogens	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
B9	0.09	M	Halogens	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
B10	0.09	M	Halogens	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
B11	0.09	M	Halogens	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
B12	0.09	M	Halogens	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4
C1	0.09	M	NPS	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
C2	0.09	M	NPS	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
C3	0.09	M	NPS	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
C4	0.09	M	NPS	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
C5	0.09	M	NPS	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
C6	0.09	M	NPS	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
C7	0.09	M	NPS	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
C8	0.09	M	NPS	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
C9	0.09	M	NPS	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
C10	0.09	M	NPS	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
C11	0.09	M	NPS	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
C12	0.09	M	NPS	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4
D1	0.12	M	Alcohols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
D2	0.12	M	Alcohols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
D3	0.12	M	Alcohols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
D4	0.12	M	Alcohols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
D5	0.12	M	Alcohols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
D6	0.12	M	Alcohols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
D7	0.12	M	Alcohols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
D8	0.12	M	Alcohols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
D9	0.12	M	Alcohols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
D10	0.12	M	Alcohols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
D11	0.12	M	Alcohols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
D12	0.12	M	Alcohols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4

Morpheus continued:

E1	0.12	M	Ethylene glycols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
E2	0.12	M	Ethylene glycols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
E3	0.12	M	Ethylene glycols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
E4	0.12	M	Ethylene glycols	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
E5	0.12	M	Ethylene glycols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
E6	0.12	M	Ethylene glycols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
E7	0.12	M	Ethylene glycols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
E8	0.12	M	Ethylene glycols	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
E9	0.12	M	Ethylene glycols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
E10	0.12	M	Ethylene glycols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
E11	0.12	M	Ethylene glycols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
E12	0.12	M	Ethylene glycols	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4
F1	0.12	M	Monosaccharides	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
F2	0.12	M	Monosaccharides	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
F3	0.12	M	Monosaccharides	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
F4	0.12	M	Monosaccharides	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
F5	0.12	M	Monosaccharides	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
F6	0.12	M	Monosaccharides	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
F7	0.12	M	Monosaccharides	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
F8	0.12	M	Monosaccharides	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
F9	0.12	M	Monosaccharides	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
F10	0.12	M	Monosaccharides	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
F11	0.12	M	Monosaccharides	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
F12	0.12	M	Monosaccharides	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4
G1	0.1	M	Carboxylic acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
G2	0.1	M	Carboxylic acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
G3	0.1	M	Carboxylic acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
G4	0.1	M	Carboxylic acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
G5	0.1	M	Carboxylic acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
G6	0.1	M	Carboxylic acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
G7	0.1	M	Carboxylic acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
G8	0.1	M	Carboxylic acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
G9	0.1	M	Carboxylic acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
G10	0.1	M	Carboxylic acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
G11	0.1	M	Carboxylic acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
G12	0.1	M	Carboxylic acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4
H1	0.1	M	Amino acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 1
H2	0.1	M	Amino acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 2
H3	0.1	M	Amino acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 3
H4	0.1	M	Amino acids	0.1	M	Buffer System 1	6.5	50	% v/v	Precipitant Mix 4
H5	0.1	M	Amino acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 1
H6	0.1	M	Amino acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 2
H7	0.1	M	Amino acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 3
H8	0.1	M	Amino acids	0.1	M	Buffer System 2	7.5	50	% v/v	Precipitant Mix 4
H9	0.1	M	Amino acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 1
H10	0.1	M	Amino acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 2
H11	0.1	M	Amino acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 3
H12	0.1	M	Amino acids	0.1	M	Buffer System 3	8.5	50	% v/v	Precipitant Mix 4

MIDAS	Conc1	Units1	Salt	Conc2	Units2	Buffer2	pH	Conc3	Units3	Precipitant3	Conc4	Units4	Preicpitant4	Conc5	Units5	Precipitant5	
1-1				0.1	M	HEPES	6.0	50	% v/v	Polypropylene glycol 400	5	% v/v	Dimethyl sulfoxide				
1-2				0.1	M	MES	5.5	12	% w/v	Polyvinylpyrrolidone							
1-3				0.1	M	HEPES	6.5	45	% w/v	Poly(acrylic acid sodium salt) 2100							
1-4								14	% v/v	Poly(acrylic acid-co-maleic acid) solution							
1-5	0.5	M	Ammonium phosphate monobasic					12.5	% w/v	Poly(acrylic acid sodium salt) 2100							
1-6				0.1	M	Tris	8.5	19	% v/v	Poly(acrylic acid-co-maleic acid) solution							
1-7								10	% v/v	Polypropylene glycol 400							
1-8								5	% w/v	Poly(acrylic acid sodium salt) 2100							
1-9				0.1	M	MES	6.0	25	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)							
1-10	0.1	M	Sodium sulfate					24	% w/v	Polyvinylpyrrolidone							
1-11	0.2	M	Calcium chloride dihydrate	0.1	M	HEPES	6.5	35	% v/v	Pentaerythritol ethoxylate (15/4 EO/OH)							
1-12				0.1	M	Potassium/sodium phosphate	7.0	35	% v/v	Polypropylene glycol 400							
1-13	0.2	M	Sodium chloride	0.1	M	MES	5.5	20	% v/v	Jeffamine® D-2000	10	% v/v	Jeffamine® M-2005				
1-14	0.2	M	Sodium thiocyanate	0.1	M	HEPES	7.0	15	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)							
1-15	0.2	M	Potassium acetate	0.1	M	HEPES	7.0	5	% w/v	Polyvinyl alcohol	10	% v/v	Jeffamine® T-403				
1-16	0.2	M	Sodium chloride	0.1	M	MES	6.0	45	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)							
1-17				0.1	M	HEPES	7.0	8	% w/v	Polyvinyl alcohol	10	% v/v	1-Propanol				
1-18	0.1	M	Lithium sulfate	0.1	M	HEPES	7.0	30	% w/v	Polyvinylpyrrolidone							
1-19				0.2	M	Imidazole	7.0	40	% v/v	Polypropylene glycol 400							
1-20	0.06	M	Lithium sulfate	0.1	M	HEPES	7.5	8	% v/v	Poly(acrylic acid-co-maleic acid) solution	3	% v/v	Pentaerythritol ethoxylate (3/4 EO/OH)				
1-21	0.1	M	Sodium chloride	0.1	M	Tris	8.0	35	% v/v	Jeffamine® SD-2001							
1-22								30	% v/v	Jeffamine® M-600	10	% v/v	Dimethyl sulfoxide				
1-23								20	% v/v	Polypropylene glycol 400	10	% v/v	1-Propanol				
1-24				0.1	M	HEPES	6.5	28	% v/v	Poly(acrylic acid-co-maleic acid) solution							
1-25								15	% v/v	Jeffamine® ED-2003	10	% v/v	Ethanol				
1-26	0.2	M	Sodium chloride	0.1	M	MES	6.0	30	% v/v	Jeffamine® ED-2003							
1-27	0.1	M	Sodium malonate dibasic monohydrate	0.1	M	MES	5.5	25	% v/v	Jeffamine® SD-2001							
1-28	0.2	M	Sodium chloride	0.1	M	MES	6.0	15	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)							
1-29	0.2	M	Magnesium chloride hexahydrate					35	% v/v	Pentaerythritol ethoxylate (3/4 EO/OH)							
1-30								40	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)	15	% v/v	Ethanol				
1-31				0.1	M	Tris	8.0	50	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)							
1-32	0.2	M	Sodium chloride	0.1	M	Tris	8.0	12.5	% w/v	Polyvinylpyrrolidone	10	% w/v	PEG 4000				
1-33	0.1	M	Sodium chloride					25	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)	10	% v/v	Dimethyl sulfoxide				
1-34	0.2	M	Ammonium sulfate	0.1	M	HEPES	7.5	35	% w/v	Poly(acrylic acid sodium salt) 2100							
1-35	0.1	M	Magnesium formate dihydrate	0.1	M	Tris	8.5	30	% v/v	Pentaerythritol ethoxylate (15/4 EO/OH)							
1-36	0.2	M	Potassium acetate					24	% v/v	Poly(acrylic acid-*co*-maleic acid) solution							
1-37				0.1	M	Tris	8.0	60	% v/v	Polypropylene glycol 400							
1-38				0.1	M	HEPES	7.5	30	% v/v	Pentaerythritol ethoxylate (15/4 EO/OH)	6	% w/v	Polyvinylpyrrolidone				
1-39								45	% v/v	Polypropylene glycol 400	10	% v/v	Ethanol				
1-40								10	% v/v	Pentaerythritol ethoxylate (3/4 EO/OH)	10	% v/v	1-Butanol				
1-41				0.1	M	HEPES	7.0	12.5	% w/v	Poly(acrylic acid sodium salt) 2100	6	% v/v	Jeffamine® SD-2001				
1-42				0.1	M	HEPES	6.5	6	% w/v	Polyvinylpyrrolidone							
1-43				0.1	M	HEPES	6.5	20	% v/v	Jeffamine® ED-2003							
1-44				0.1	M	Tris	8.0	20	% v/v	Glycerol ethoxylate	10	% v/v	Tetrahydrofuran				
1-45				0.2	M	Imidazole	7.0	25	% v/v	Jeffamine® D-2000							
1-46	0.2	M	Potassium chloride	0.1	M	HEPES	6.5	30	% v/v	Jeffamine® SD-2001							
1-47	0.1	M	Sodium chloride					30	% v/v	Polypropylene glycol 400							

MIDAS	Conc1	Units1	Salt	Conc2	Units2	Buffer2	pH	Conc3	Units3	Precipitant3	Conc4	Units4	Preicpitant4	Conc5	Units5	Precipit
1-48								20	% v/v	Jeffamine® SD-2001	15	% v/v	1-Propanol			
2-1	0.2	M	Lithium sulfate	0.1	M	Tris	8.0	25	% v/v	Jeffamine® T-403						
2-2	0.2	M	Potassium acetate					35	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)						
2-3	0.2	M	Potassium chloride	0.1	M	Glycine	9.5	20	% v/v	Pentaerythritol ethoxylate (15/4 EO/OH)						
2-4	0.2	M	Sodium thiocyanate	0.1	M	HEPES	7.0	40	% v/v	Pentaerythritol propoxylate (5/4 PO/OH)						
2-5	0.2	M	Potassium chloride	0.1	M	HEPES	6.5	15	% v/v	Jeffamine® T-403	15	% v/v	Jeffamine® ED-2003			
2-6	0.2	M	Potassium acetate	0.1	M	MES	6.0	15	% v/v	Pentaerythritol ethoxylate (15/4 EO/OH)	3	% v/v	Jeffamine® T-403			
2-7	0.1	M	Sodium malonate dibasic monohydrate	0.1	M	HEPES	7.0	30	% w/v	Poly(acrylic acid sodium salt) 2100						
2-8								10	% v/v	Jeffamine® D-2000	10	% v/v	Jeffamine® M-2005	10	% v/v	Ethanol
2-9	0.1	M	Lithium sulfate	0.1	M	Tris	8.0	25	% v/v	Jeffamine® ED-2003						
2-10				0.1	M	Tris	8.0	10	% v/v	Jeffamine® T-403	10	% v/v	Jeffamine® ED-2003			
2-11	0.1	M	Lithium sulfate	0.1	M	HEPES	6.5	25	% w/v	Poly(acrylic acid sodium salt) 2100						
2-12	0.2	M	Magnesium chloride hexahydrate	0.1	M	HEPES	7.5	15	% w/v	Poly(acrylic acid sodium salt) 2100						
2-13				0.1	M	HEPES	6.5	40	% v/v	Jeffamine® D-2000						
2-14	0.5	M	Sodium chloride	0.1	M	Tris	8.0	10	% w/v	Poly(acrylic acid sodium salt) 2100						
2-15				0.1	M	Potassium/sodium phosphate	7.0	14	% v/v	Jeffamine® ED-900	11	% v/v	Jeffamine® SD-2001			
2-16	0.2	M	Sodium chloride	0.1	M	BICINE	9.0	20	% w/v	Poly(acrylic acid sodium salt) 2100						
2-17	0.2	M	Sodium malonate dibasic monohydrate	0.1	M	MES	5.5	20	% v/v	Jeffamine® D-2000						
2-18	0.2	M	Potassium chloride	0.1	M	Tris	8.0	30	% v/v	Jeffamine® M-2070						
2-19								20	% v/v	Jeffamine® M-2070	20	% v/v	Dimethyl sulfoxide			
2-20	0.2	M	Magnesium chloride hexahydrate	0.1	M	MES	5.5	40	% v/v	Pentaerythritol propoxylate (17/8 PO/OH)						
2-21				0.1	M	Tris	8.0	20	% w/v	Poly(acrylic acid sodium salt) 5100						
2-22				0.1	M	HEPES	7.0	28	% v/v	Polyethyleneimine						
2-23	0.1	M	Ammonium formate	0.1	M	HEPES	7.0	20	% w/v	SOKALAN® CP 7						
2-24	0.2	M	Sodium sulfate	0.1	M	Tris	8.0	20	% w/v	SOKALAN® HP 56						
2-25	0.1	M	Potassium chloride	0.1	M	HEPES	7.0	25	% w/v	SOKALAN® CP 7						
2-26	0.3	M	Ammonium formate	0.1	M	HEPES	7.0	20	% w/v	SOKALAN® CP 5						
2-27								40	% v/v	Glycerol ethoxylate						
2-28				0.1	M	Tris	8.5	30	% v/v	Glycerol ethoxylate						
2-29								55	% v/v	Polypropylene glycol 400						
2-30	0.2	M	Lithium citrate tribasic tetrahydrate					35	% v/v	Glycerol ethoxylate						
2-31	0.2	M	Ammonium acetate	0.1	M	MES	6.5	30	% v/v	Glycerol ethoxylate						
2-32				0.1	M	Tris	8.0	20	% w/v	SOKALAN® CP 42	5	% v/v	Methanol			
2-33				0.1	M	Tris	7.0	25	% w/v	SOKALAN® CP 42	10	% v/v	Tetrahydrofuran			
2-34	0.1	M	Lithium acetate dihydrate	0.1	M	Bis-Tris	6.0	20	% w/v	SOKALAN® CP 42						
2-35				0.2	M	HEPES	6.5	10	% v/v	Jeffamine® M-2005						
2-36				0.1	M	Bis-Tris	6.0	15	% w/v	SOKALAN® CP 5						
2-37				0.1	M	Bis-Tris	6.0	25	% w/v	SOKALAN® CP 42						
2-38								35	% v/v	Jeffamine® D-2000						
2-39				0.1	M	Tris	8.5	20	% v/v	Glycerol ethoxylate	3	% v/v	Polyethyleneimine			
2-40	0.2	M	Ammonium chloride	0.1	M	HEPES	7.5	25	% v/v	Glycerol ethoxylate						
2-41				0.1	M	Tris	8.5	10	% w/v	SOKALAN® CP 42						
2-42				0.1	M	MES	6.0	30	% w/v	Poly(acrylic acid sodium salt) 5100	10	% v/v	Ethanol			
2-43	0.2	M	Potassium citrate tribasic monohydrate					15	% w/v	SOKALAN® CP 42						
2-44				0.1	M	Tris	8.5	30	% w/v	SOKALAN® CP 42						
2-45	0.2	M	Ammonium acetate	0.1	M	HEPES	7.0	25	% w/v	SOKALAN® HP 56						
2-46				0.1	M	Tris	8.5	25	% w/v	SOKALAN® CP 5						
2-47	0.2	M	Ammonium formate					10	% w/v	Polyvinylpyrrolidone	20	% w/v	PEG 4000			
2-48				0.1	M	Tris	8.0	15	% w/v	Polyvinylpyrrolidone	25	% w/v	PEG 5000 MME			

JCSG+														
Units1	Salt1	Conc2	Units2	Salt2	Conc5	Units5	Buffer5	pH	Conc6	Units6	Precipitan	Conc7	Units7	Precipitant7
M	Lithium sulfate				0.1	M	Sodium acetate	4.5	50	% w/v	PEG 400			
	None				0.1	M	Sodium citrate	5.5	20	% w/v	PEG 3000			
M	Ammonium citrate dibasic						None	-	20	% w/v	PEG 3350			
M	Calcium chloride dihydrate				0.1	M	Sodium acetate	4.6	30	% v/v	MPD			
M	Magnesium formate dihydrate						None	-	20	% w/v	PEG 3350			
M	Lithium sulfate				0.1	M	Phosphate/citrate	4.2	20	% w/v	PEG 1000			
	None				0.1	M	CHES	9.5	20	% w/v	PEG 8000			
M	Ammonium formate						None	-	20	% w/v	PEG 3350			
M	Ammonium chloride						None	-	20	% w/v	PEG 3350			
M	Potassium formate						None	-	20	% w/v	PEG 3350			
M	Ammonium phosphate monobasic				0.1	M	Tris	8.5	50	% v/v	MPD			
M	Potassium nitrate						None	-	20	% w/v	PEG 3350			
M	Ammonium sulfate				0.1	M	Citrate	4			None			
M	Sodium thiocyanate						None	-	20	% w/v	PEG 3350			
	None				0.1	M	BICINE	9	20	% w/v	PEG 6000			
	None				0.1	M	HEPES	7.5	10	% w/v	PEG 8000	8	% v/v	Ethylene glycol
	None				0.1	M	Sodium cacodylate	6.5	40	% v/v	MPD	5	% w/v	PEG 8000
	None				0.1	M	Phosphate/citrate	4.2	40	% v/v	Ethanol	5	% w/v	PEG 1000
	None				0.1	M	Sodium acetate	4.6	8	% w/v	PEG 4000			
M	Magnesium chloride hexahydrate				0.1	M	Tris	7	10	% w/v	PEG 8000			
	None				0.1	M	Citrate	5	20	% w/v	PEG 6000			
M	Magnesium chloride hexahydrate				0.1	M	Sodium cacodylate	6.5	50	% v/v	PEG 200			
M	Sodium citrate tribasic dihydrate						None	6.5			None			
M	Potassium citrate tribasic monohydrate						None	-	20	% w/v	PEG 3350			
M	Sodium chloride				0.1	M	Phosphate/citrate	4.2	20	% w/v	PEG 8000			
M	Lithium chloride				0.1	M	Citrate	4	20	% w/v	PEG 6000			
M	Ammonium nitrate						None	-	20	% w/v	PEG 3350			
	None				0.1	M	HEPES	7	10	% w/v	PEG 6000			
M	Sodium pl	0.8	M	Potassium phosphate monobasic	0.1	M	Sodium HEPES	7.5			None			
	None				0.1	M	Phosphate/citrate	4.2	40	% v/v	PEG 300			
M	Zinc acetate dihydrate				0.1	M	Sodium acetate	4.5	10	% w/v	PEG 3000			
	None				0.1	M	Tris	8.5	20	% v/v	Ethanol			
	None				0.1	M	Sodium/potassium phosphate	6.2	25	% v/v	1,2-Propa	10	%v/v	Glycerol
	None				0.1	M	BICINE	9	10	% w/v	PEG 20,00	2	% v/v	1,4-Dioxane
M	Ammonium sulfate				0.1	M	Sodium acetate	4.6			None			
	None						None	-	10	% w/v	PEG 1000	10	% w/v	PEG 8000
	None						None	-	24	% w/v	PEG 1500	20	% v/v	Glycerol
M	Magnesium chloride hexahydrate				0.1	M	Sodium HEPES	7.5	30	% v/v	PEG 400			
M	Sodium chloride				0.1	M	Sodium/potassium phosphate	6.2	50	% v/v	PEG 200			
M	Lithium sulfate				0.1	M	Sodium acetate	4.5	30	% w/v	PEG 8000			
	None				0.1	M	HEPES	7.5	70	% v/v	MPD			
M	Magnesium chloride hexahydrate				0.1	M	Tris	8.5	20	% w/v	PEG 8000			
M	Lithium sulfate				0.1	M	Tris	8.5	40	% v/v	PEG 400			
	None				0.1	M	Tris	8	40	% v/v	MPD			
M	Ammonium sulfate						None	-	25.5	% w/v	PEG 4000	15	% v/v	Glycerol
M	Calcium acetate hydrate				0.1	M	Sodium cacodylate	6.5	40	% v/v	PEG 300			
M	Calcium chloride dihydrate				0.07	M	Sodium acetate	4.6	14	% v/v	2-Propanc	30	% v/v	Glycerol

JCSG+																					
Units1	Salt1	Conc2	Units2	Salt2	Conc3	Units3	Salt3	Conc4	Units4	Salt4	Conc5	Units5	Buffer5	pH	Conc6	Units6	Precipitan	Conc7	Units7	Precipitan
M	Potassium phosphate monobasic												None	-	16	% w/v	PEG 8000	20	% v/v	Glycerol
M	Sodium citrate tribasic dihydrate										0.1	M	Sodium ca	6.5			None			
M	Ammoniu	0.2	M	Sodium chloride							0.1	M	Sodium ca	6.5			None			
M	Sodium chloride										0.1	M	HEPES	7.5	10	% v/v	2-Propanol			
M	Ammoniu	0.2	M	Lithium sulfate							0.1	M	Tris	8.5			None			
	None										0.1	M	CAPS	10.5	40	% v/v	MPD			
M	Zinc acetate dihydrate										0.1	M	Imidazole	8	20	% w/v	PEG 3000			
M	Zinc acetate dihydrate										0.1	M	Sodium ca	6.5	10	% v/v	2-Propanol			
M	Ammonium phosphate dibasic										0.1	M	Sodium ac	4.5			None			
M	Magnesium sulfate heptahydrate										0.1	M	MES	6.5			None			
	None										0.1	M	BICINE	9	10	% w/v	PEG 6000			
M	Calcium acetate hydrate										0.08	M	Sodium ca	6.5	14.4	% w/v	PEG 8000	20	% v/v	Glycerol
	None										0.1	M	Imidazole	8	10	% w/v	PEG 8000			
M	Cesium chloride										0.1	M	MES	6.5	30	% v/v	Jeffamine® M-600			
M	Ammonium sulfate										0.1	M	Citrate	5			None			
	None										0.1	M	Tris	8	20	% v/v	MPD			
	None										0.1	M	HEPES	7.5	20	% v/v	Jeffamine® M-600			
M	Magnesium chloride hexahydrate										0.1	M	Tris	8.5	50	% v/v	Ethylene glycol			
	None										0.1	M	BICINE	9	10	% v/v	MPD			
M	Succinic acid												None	7			None			
M	DL-Malic acid												None	7			None			
M	Sodium malonate dibasic monohydrate												None	7			None			
M	Sodium malonate dibasic monohydrate										0.1	M	HEPES	7	0.5	% v/v	Jeffamine® ED-2003			
M	Succinic acid										0.1	M	HEPES	7	1	% w/v	PEG 2000 MME			
	None										0.1	M	HEPES	7	30	% v/v	Jeffamine® M-600			
	None										0.1	M	HEPES	7	30	% v/v	Jeffamine® ED-2003			
M	Magnesium chloride hexahydrate										0.1	M	HEPES	7.5	22	% w/v	Poly(acrylic acid sodium salt) 5100			
M	Cobalt(II) chloride hexahydrate										0.1	M	Tris	8.5	20	% w/v	Polyvinylpyrolidone			
M	TMAO										0.1	M	Tris	8.5	20	% w/v	PEG 2000 MME			
M	Cobalt(II)	0.005	M	Cadmium chloride hemi(pentahy	0.005	M	Magnesium chloride hexahydra	0.005	M	Nickel(II)	0.1	M	HEPES	7.5	12	% w/v	PEG 3350			
M	Sodium malonate dibasic monohydrate												None	-	20	% w/v	PEG 3350			
M	Succinic acid												None	-	15	% w/v	PEG 3350			
M	DL-Malic acid												None	-	20	% w/v	PEG 3350			
M	Potassium thiocyanate												None	-	30	% w/v	PEG 2000 MME			
M	Potassium bromide												None	-	30	% w/v	PEG 2000 MME			
M	Ammonium sulfate										0.1	M	Bis-Tris	5.5			None			
M	Sodium chloride										0.1	M	Bis-Tris	5.5			None			
M	Magnesium formate dihydrate										0.1	M	Bis-Tris	5.5			None			
M	Ammonium sulfate										0.1	M	Bis-Tris	5.5	1	% w/v	PEG 3350			
	None										0.1	M	Bis-Tris	5.5	25	% w/v	PEG 3350			
M	Calcium chloride dihydrate										0.1	M	Bis-Tris	5.5	45	% v/v	MPD			
M	Ammonium acetate										0.1	M	Bis-Tris	5.5	45	% v/v	MPD			
M	Ammonium acetate										0.1	M	Bis-Tris	5.5	17	% w/v	PEG 10,000			
M	Ammonium sulfate										0.1	M	Bis-Tris	5.5	25	% w/v	PEG 3350			
M	Sodium chloride										0.1	M	Bis-Tris	5.5	25	% w/v	PEG 3350			
M	Lithium sulfate										0.1	M	Bis-Tris	5.5	25	% w/v	PEG 3350			
M	Ammonium acetate										0.1	M	Bis-Tris	5.5	25	% w/v	PEG 3350			
M	Magnesium chloride hexahydrate										0.1	M	Bis-Tris	5.5	25	% w/v	PEG 3350			
M	Ammonium acetate										0.1	M	HEPES	7.5	45	% v/v	MPD			

PACT

Well #	Conc1	Units1	Salt 1	Conc2	Units2	Buffer2	pH	Conc3	Units3	Precipitant3
A1				0.1	M	SPG	4.0	25	% w/v	PEG 1500
A2				0.1	M	SPG	5.0	25	% w/v	PEG 1500
A3				0.1	M	SPG	6.0	25	% w/v	PEG 1500
A4				0.1	M	SPG	7.0	25	% w/v	PEG 1500
A5				0.1	M	SPG	8.0	25	% w/v	PEG 1500
A6				0.1	M	SPG	9.0	25	% w/v	PEG 1500
A7	0.2	M	Sodium chloride	0.1	M	Sodium acetate	5.0	20	% w/v	PEG 6000
A8	0.2	M	Ammonium chloride	0.1	M	Sodium acetate	5.0	20	% w/v	PEG 6000
A9	0.2	M	Lithium chloride	0.1	M	Sodium acetate	5.0	20	% w/v	PEG 6000
A10	0.2	M	Magnesium chloride hexahydrate	0.1	M	Sodium acetate	5.0	20	% w/v	PEG 6000
A11	0.2	M	Calcium chloride dihydrate	0.1	M	Sodium acetate	5.0	20	% w/v	PEG 6000
A12	0.01	M	Zinc chloride	0.1	M	Sodium acetate	5.0	20	% w/v	PEG 6000
B1				0.1	M	MIB	4.0	25	% w/v	PEG 1500
B2				0.1	M	MIB	5.0	25	% w/v	PEG 1500
B3				0.1	M	MIB	6.0	25	% w/v	PEG 1500
B4				0.1	M	MIB	7.0	25	% w/v	PEG 1500
B5				0.1	M	MIB	8.0	25	% w/v	PEG 1500
B6				0.1	M	MIB	9.0	25	% w/v	PEG 1500
B7	0.2	M	Sodium chloride	0.1	M	MES	6.0	20	% w/v	PEG 6000
B8	0.2	M	Ammonium chloride	0.1	M	MES	6.0	20	% w/v	PEG 6000
B9	0.2	M	Lithium chloride	0.1	M	MES	6.0	20	% w/v	PEG 6000
B10	0.2	M	Magnesium chloride hexahydrate	0.1	M	MES	6.0	20	% w/v	PEG 6000
B11	0.2	M	Calcium chloride dihydrate	0.1	M	MES	6.0	20	% w/v	PEG 6000
B12	0.01	M	Zinc chloride	0.1	M	MES	6.0	20	% w/v	PEG 6000
C1				0.1	M	PCTP	4.0	25	% w/v	PEG 1500
C2				0.1	M	PCTP	5.0	25	% w/v	PEG 1500
C3				0.1	M	PCTP	6.0	25	% w/v	PEG 1500
C4				0.1	M	PCTP	7.0	25	% w/v	PEG 1500
C5				0.1	M	PCTP	8.0	25	% w/v	PEG 1500
C6				0.1	M	PCTP	9.0	25	% w/v	PEG 1500
C7	0.2	M	Sodium chloride	0.1	M	HEPES	7.0	20	% w/v	PEG 6000
C8	0.2	M	Ammonium chloride	0.1	M	HEPES	7.0	20	% w/v	PEG 6000
C9	0.2	M	Lithium chloride	0.1	M	HEPES	7.0	20	% w/v	PEG 6000
C10	0.2	M	Magnesium chloride hexahydrate	0.1	M	HEPES	7.0	20	% w/v	PEG 6000
C11	0.2	M	Calcium chloride hexahydrate	0.1	M	HEPES	7.0	20	% w/v	PEG 6000
C12	0.01	M	Zinc chloride	0.1	M	HEPES	7.0	20	% w/v	PEG 6000
D1				0.1	M	MMT	4.0	25	% w/v	PEG 1500
D2				0.1	M	MMT	5.0	25	% w/v	PEG 1500
D3				0.1	M	MMT	6.0	25	% w/v	PEG 1500
D4				0.1	M	MMT	7.0	25	% w/v	PEG 1500
D5				0.1	M	MMT	8.0	25	% w/v	PEG 1500
D6				0.1	M	MMT	9.0	25	% w/v	PEG 1500
D7	0.2	M	Sodium chloride	0.1	M	Tris	8.0	20	% w/v	PEG 6000
D8	0.2	M	Ammonium chloride	0.1	M	Tris	8.0	20	% w/v	PEG 6000
D9	0.2	M	Lithium chloride	0.1	M	Tris	8.0	20	% w/v	PEG 6000
D10	0.2	M	Magnesium chloride hexahydrate	0.1	M	Tris	8.0	20	% w/v	PEG 6000

PACT continued:

D11	0.2	M	Calcium chloride dihydrate	0.1	M	Tris	8.0	20	% w/v	PEG 6000
D12	0.002	M	Zinc chloride	0.1	M	Tris	8.0	20	% w/v	PEG 6000
E1	0.2	M	Sodium fluoride					20	% w/v	PEG 3350
E2	0.2	M	Sodium bromide					20	% w/v	PEG 3350
E3	0.2	M	Sodium iodide					20	% w/v	PEG 3350
E4	0.2	M	Potassium thiocyanate					20	% w/v	PEG 3350
E5	0.2	M	Sodium nitrate					20	% w/v	PEG 3350
E6	0.2	M	Sodium formate					20	% w/v	PEG 3350
E7	0.2	M	Sodium acetate trihydrate					20	% w/v	PEG 3350
E8	0.2	M	Sodium sulfate					20	% w/v	PEG 3350
E9	0.2	M	Potassium sodium tartrate tetrahydrate					20	% w/v	PEG 3350
E10	0.02	M	Sodium/potassium phosphate					20	% w/v	PEG 3350
E11	0.2	M	Sodium citrate tribasic dihydrate					20	% w/v	PEG 3350
E12	0.2	M	Sodium malonate dibasic monohydrate					20	% w/v	PEG 3350
F1	0.2	M	Sodium fluoride	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F2	0.2	M	Sodium bromide	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F3	0.2	M	Sodium iodide	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F4	0.2	M	Potassium thiocyanate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F5	0.2	M	Sodium nitrate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F6	0.2	M	Sodium formate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F7	0.2	M	Sodium acetate trihydrate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F8	0.2	M	Sodium sulfate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F9	0.2	M	Potassium sodium tartrate tetrahydrate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F10	0.02	M	Sodium/potassium phosphate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F11	0.2	M	Sodium citrate tribasic dihydrate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
F12	0.2	M	Sodium malonate dibasic monohydrate	0.1	M	Bis-Tris propane	6.5	20	% w/v	PEG 3350
G1	0.2	M	Sodium fluoride	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G2	0.2	M	Sodium bromide	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G3	0.2	M	Sodium iodide	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G4	0.2	M	Potassium thiocyanate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G5	0.2	M	Sodium nitrate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G6	0.2	M	Sodium formate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G7	0.2	M	Sodium acetate trihydrate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G8	0.2	M	Sodium sulfate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G9	0.2	M	Potassium sodium tartrate tetrahydrate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G10	0.02	M	Sodium/potassium phosphate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G11	0.2	M	Sodium citrate tribasic dihydrate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
G12	0.2	M	Sodium malonate dibasic monohydrate	0.1	M	Bis-Tris propane	7.5	20	% w/v	PEG 3350
H1	0.2	M	Sodium fluoride	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H2	0.2	M	Sodium bromide	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H3	0.2	M	Sodium iodide	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H4	0.2	M	Potassium thiocyanate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H5	0.2	M	Sodium nitrate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H6	0.2	M	Sodium formate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H7	0.2	M	Sodium acetate trihydrate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H8	0.2	M	Sodium sulfate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H9	0.2	M	Potassium sodium tartrate tetrahydrate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H10	0.02	M	Sodium/potassium phosphate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H11	0.2	M	Sodium citrate tribasic dihydrate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350
H12	0.2	M	Sodium malonate dibasic monohydrate	0.1	M	Bis-Tris propane	8.5	20	% w/v	PEG 3350