



**Recombinant Expression and  
Characterization of Ligand Binding to  
Mutants Of 3-O-sulphotransferase-1**

**Niall Walsh B.Sc. M.Sc.**

Thesis presented for the Ph.D. Degree  
of the  
National University of Ireland, Galway

School of Chemistry,  
National University of Ireland, Galway

June 2014

Supervisor: Professor Robert Woods

# Table of Contents

Abstract.....	iii
List of tables.....	iv
List of figures.....	v-xi
List of abbreviations.....	xi-xii
Acknowledgments.....	xiv

## **Chapter 1: Introduction**

Introduction Part 1: Heparin and 3-O-sulphotransferases.....	2-21
Introduction Part 2: Application of Heparin in research.....	22-26
Introduction Part 3: Developing Heparin reagents.....	27-30

## **Chapter 2: Materials and Methods**

Materials and Methods.....	31-46
----------------------------	-------

## **Chapter 3: Results and Discussion**

Results and Discussion.....	48-108
-----------------------------	--------

## **Chapter 4: Conclusions and Future Perspectives**

Conclusions and Future Perspectives.....	109-115
--	---------

Appendix (A): Project Methods.....	116-142
------------------------------------	---------

Appendix (B): Experimental Conditions.....	143-147
--	---------

References.....	148-156
-----------------	---------

# Abstract

Heparan Sulphate (HS) is a highly charged glycosaminoglycan (GAG) that is closely related to Heparin, a related GAG with a higher degree of sulphation. Heparin is used in the development of anticoagulant drugs. Currently, there is a lack of useful reagents for studying and distinguishing between complex GAGs like heparin/heparan sulphate from other GAGs. As a result, GAGs are often contaminated with other GAGs like chondroitin sulphate upon isolation. To improve GAG characterization, we aim to engineer a heparin reagent through the conversion of the HS modifying enzyme 3-O-sulphotransferase isoform 1 (3-OST-1). 3-OST-1 transfers a sulphate group to the 3-OH position of an *N*-sulpho-glucosamine linked to a glucuronic acid. This modification is essential for heparin used in the development of anticoagulant drugs like Arixtra. Four 3-OST-1 mutants, E90Q, E90K, E90H and E90R were generated using site-directed mutagenesis. The binding affinity for 3-OST-1 modified heparin was measured by surface plasmon resonance using the anticoagulant drug Arixtra. The effect of side chain and charge at position 90 among the four mutants in binding to Arixtra was investigated. The mutant with the greatest binding affinity was determined. Here in is the report on the expression, purification and binding analysis of the mutants to date.

# List of tables

**Table 1:** Optimized thermocycling conditions and optimized site-directed mutagenesis reaction mixture for the generation of mutants

**Table 2:** SPR affinity and kinetics data of the interaction of the protein samples to Arixtra.

**Table 3:** Sonication parameters.

**Table 4:** Enzyme activity reaction mixture.

**Table 5:** Site-directed mutagenesis primers.

**Table 6:** Site-directed mutagenesis primers features.

**Table 7:** Site-directed mutagenesis reaction mixture and thermocycling conditions.

**Table 8:** Thermocycling conditions for the generation of 3-OST-1 mutants.

**Table 9:** Colony PCR conditions.

**Table 10:** SDS-PAGE gel composition.

# List of figures

**Figure 1:** Main building blocks of Heparan sulphate and Heparin, the positions of the sulphate groups highlighted in red. R=  $\text{SO}_3^-$  or an H atom, R'=  $\text{SO}_3^-$  or a  $\text{COOH}_3$ .

**Figure 2:** Illustration of the formation the Heparan sulphate chain: **(1)** HS synthesis begins with the addition of an *N*-acetyl-glucosamine residue to a core oligosaccharide consisting of Xyl-Gal-Gal-GlcA , catalysed by the enzyme  $\alpha$ -*N*-acetylglucosaminyltransferase 1. **(2)** The HS chain is extended through the sequential addition of *N*-acetyl-glucosamine residues followed by alternating glucuronic acid residues. This is achieved via complex of the enzymes EXT1 glycosyltransferase 1 and EXT2 glycosyltransferase 2. **(3)** After the synthesis of the HS polymer a series of modifications take place in the Golgi apparatus, via a series of sulphotransferases.

**Figure 3:** In vivo synthesis of heparan sulphate begins with the deacetylation and *N*-sulphation of glucosamine residues by *N*-deactylase/*N*-sulphotransferase. Next the glucuronic acid residues are C5-epimerized by C5-epimerase to form iduronic acid. The iduronic acid residues are then sulphated by 2-*O*-sulphotransferase. The glucosamine residues are sulphated at the 6- position by 6-*O*-sulphotransferase and finally at the 3- position by 3-*O*-sulphotransferase.

**Figure 4:** Common conformations of the main building blocks of Heparan sulphate and Heparin.

**Figure 5:** Panel A the active site of 3-OST-1 indicated by the arrow. Panel B the Phosphate-binding loop in relation to PAPS and the substrate is indicated. The 3-OH position of the acceptor sugar is circled.

**Figure 6:** Gate structure that has been noted for its importance in substrate specificity. Residues that form the gate structure in 3-OST-1 in yellow, residues that form the gate structure in 3-OST-3 in blue.

**Figure 7:** Heparin sulphate Antithrombin binding sequence represented here by the structure for Arixtra, with 3-*O*-sulphation indicated in red.

**Figure 8:** Illustration of residues known to be critical to enzymatic activity in 3-OST-1 (Edavettal, Lee et al. 2004).

**Figure 9:** The catalytic mechanism of 3-OST-1 begins with Glu-90 deprotonating the 3-OH position of a glucosamine residue. This deprotonated position attacks the sulphur

atom of the co-factor PAPS. The result is the formation of a transition state. This state is stabilized through Lys-68 of the phosphate binding loop. After a build-up of negative charge a sulphate is transferred to the 3 position on the glucosamine residue.

**Figure 10:** Count per minute average (CPMA) results of the liquid scintillation assay:

The CPMA results here are proportional to the amount of [<sup>35</sup>S] O<sub>3</sub> transferred to heparan sulphate by a sulphotransferases, in this case the wild type 3-OST-1. The test sample consisted of a newly expressed wild type sample, the positive control consisted of a sample (wild type) of verified activity and the negative control consisted of an inactive wild type expression. The CPMA results indicate an active protein expression with the comparison of the test sample to the controls as the [<sup>35</sup>S] O<sub>3</sub> transferred to substrate is detected in both the test and positive control samples based on the CPMA results.

**Figure 11:** Position of Glu-90 (catalytic nucleophile) (in 3-OST-1) in relation to the N-sulphated and 3-OH position of a glucosamine and the carboxyl group of the glucuronic acid residue of the substrate (heparin). Highlighted (red arrows) are the negatively charged groups of the glucuronic acid and the glucosamine residue that may interact with the catalytic nucleophile. Circled in red is the position that is sulphated by 3-OST-1 and the dashed lines indicate the distance between Glu-90 and the highlighted positions in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).

**Figure 12:** Pro-125 (in 3-OST-1) relative to a hydroxyl group and the N-sulphated position of a glucosamine residue in the substrate (heparin). Highlighted (red arrows) is a negatively charged group in the glucosamine residue that may interact with the position-125. The dashed lines indicate the distance between Pro-125 and the highlighted positions in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).

**Figure 13:** Asn-167 (in 3-OST-1) relative to an N-sulphated position of the glucosamine and the carboxyl group of the adjacent iduronic acid residue in the substrate (heparin). Highlighted (red arrows) are the negatively charged groups of the iduronic acid and the glucosamine residue that may interact with the position-167. The dashed lines indicate the distance between Asn-167 and the highlighted positions in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).

**Figure 14:** Lys-123 (in 3-OST-1) relative to the carboxyl group of glucuronic acid residue in the substrate (heparin). Highlighted (red arrows) is a negatively charged group of the glucuronic acid that may interact with the position-123. The dashed lines indicate

the distance between Lys-123 and the highlighted position in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).

**Figure 15:** Colony PCR samples (C) indicated from row 1-9, (1) M1 = Molecular weight ladder (DNA ladder exACTGene), row (2) – (9) = C1-8, row (10) M2= Low molecular weight ladder (Low Scale Ladder 100 bp) and row (11) = C9. Approximate size of the PCR amplified product is 1102 base pairs (bp).

**Figure 16:** Truncated nucleotide and amino acid sequence of the wild type 3-OST-1 used in the alignment of the sequencing result. The mutation site underlined in part A (nucleotide sequence) and in part B (amino acid sequence). The mutagenesis primer sites are both indicated in blue in part A and B.

**Figure 17:** Alignment of the gene sequencing obtained for the mutant E90Q with the template DNA in part A. In part B is an example of a portion of the nucleotide sequence of the mutants generated with the mutation highlighted in blue.

**Figure 18:** Pie chart of amino acid coding sequence of 3-OST-1 illustrating amino acids with hydrophobic side chains, positively charged side chains, negatively charged side chains and the glutamine content.

**Figure 19:** IMAC purification (3-OST-1) with 100% elution were the protein was expressed in different cell lines. (A) Expressed in *E. coli* (DE3) plySs (4.3 mg/ml in 200  $\mu$ l), (B) Expressed in *E. coli* (DE3) RIPL (5.1 mg/ml in 250  $\mu$ l) and (C) Expressed in *E. coli* (DE3) C41 (0.96 mg/ml in 500  $\mu$ l). (1) M= Molecular weight ladder, (2)= 100% elution. Estimated molecular weight of the protein of interest= 33 kDa.

**Figure 20:** Nucleotide sequence of the wild type 3-OST-1 used in this study showing rare codons linked to poor expression in *E. coli* (underlined in purple) (used Rare Codon Calculator to analyse sequence). The rare codons are as follows:

Arginine (AGG, AGA, CGA) leucine (CTA), isoleucine (ATA), and proline (CCC).

**Figure 21:** IMAC purification (3-OST-1) with 100% elution showing examples of the use of chemical chaperones, molecular chaperones and cell lysis buffer additives. (A) Protein expression that was lysed in the presence of l-arginine (chemical chaperone) then IMAC purified (5.6 mg/ml in 500  $\mu$ l), (B) Protein expression in the presence of Takara plasmid 1 (molecular chaperone) then IMAC purified (1.1 mg/ml in 250  $\mu$ l) and (C) Protein expression that was lysed in the presence of Triton X-100 (lysis buffer additive) then IMAC purified (1.3 mg/ml in 3 ml). (1) M= Molecular weight ladder, (2)= 100% elution. Estimated molecular weight of the protein of interest= 33 kDa.

**Figure 22:** Optimization of the expression protocol (from: (Edavettal, Lee et al. 2004)) for the wildtype and mutant 3-OST-1. (A) Analysis of the original expression protocol showing (i) SDS-PAGE analysis and (ii) western blot of a soluble protein expression (3-OST-1). The analysis shows abundant insoluble expression (highlighted by the red arrows) using the original protocol. (B) Analysis of the optimized expression protocol showing (i) SDS-PAGE analysis and (ii) western blot of a soluble protein expression (3-OST-1) shows improved soluble expression. Lanes: (1) M= Molecular weight ladder, (2) S= Soluble fraction, (3) IS= Insoluble fraction. 3-OST-1 is shown with the black arrow.

**Figure 23:** First round IMAC purification of wild type 3-OST-1 soluble fraction with a single 100% elution. (A) Chromatograph of IMAC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2) Soluble fraction applied to column (S1= soluble fraction 1) and (3)= First round 100% elution (E1= Elution 1). Estimated molecular weight of the protein of interest= 33 kDa.

**Figure 24:** Second round IMAC purification (3-OST-1) of the flow through (see figure 20) from the first round IMAC purification with a single 100% elution. (A) Chromatograph of IMAC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2)= IMAC flow through (F1= flow through fraction 1) and (3)= Second round 100% elution (E2= Elution 2). Estimated molecular weight of the protein of interest= 33 kDa.

**Figure 25:** Third round IMAC (3-OST-1) of the combined elutions from the first and second round IMAC's with a single 100% elution. (A) Chromatograph of IMAC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2)= combined elutions applied to column (S2= soluble fraction 2) and (3)= Third round 100% elution (E3= Elution 3). Estimated molecular weight of the protein of interest= 33 kDa.

**Figure 26:** HIC purification (3-OST-1) of the 100% elutions from the third round IMAC with a single 100% elution. (A) Chromatograph of HIC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2)= 100% elution 1 mg/ml in 5 ml (E4= Elution 4). Estimated molecular weight of the protein of interest= 33 kDa.

**Figure 27:** Flow chart showing the optimized protein purification protocol. After harvesting the soluble fraction from the protein expression in *E. coli*, three rounds of IMAC were performed. The first round involved purifying the soluble fraction followed by a single elution step. The flow through collected from the first round purification step was purified again in the second round purification, followed by a single elution step. The elutions from the first and second round purifications were combined and purified in the third round purification, followed by a single elution step. The elution from the third round purification was then purified by HIC with a single elution step. The samples loaded onto the columns are indicated in red.

**Figure 28:** Free biotin test indicating the presence of biotin in both the flow through fractions and the collected fraction. The flow through consists of biotin that passed through the spin column membrane. The flow through was collected after the first and second centrifuge steps. The collection consists of the concentrated heparin-biotin conjugate.

**Figure 29:** SDS-PAGE showing HIC purified protein used in the ELISA experiment, (1) M= Molecular weight ladder, (2) Wild type, (3) E90R, (4) E90K, (5) E90Q and (6) E90H. Molecular weight of the protein of interest= 33 kDa. The protein concentrations were at 1 mg/ml for each sample and the total yield as follows: Wild type 6 mg in 6 ml, E90R 2 mg in 2 ml, E90K 3 mg in 3 ml, E90Q 4 mg in 4 ml and E90H 8 mg in 8 ml. The protein yield was obtained in the same manner as described in Figures 20-24, with the final 100% elution from the HIC purification making up the protein yield detailed above.

**Figure 30:** ELISA data of the wild type and mutant binding to heparin. The data indicated was blanked subtracted. The blank control wells consisted of wells coated in poly-L-lysine (absent heparin). The absorbance was measured at a wavelength of 450 nm and was plotted against the protein concentration in the corresponding well.

**Figure 31:** SPR binding curve fit to a 1:1 interaction model with immobilized protein as the ligand and Arixtra as the analyte. **(A)** Wild type versus Arixtra (Res SD : 4.219), **(B)** E90R versus Arixtra (Res SD: 2.425), **(C)** E90K versus Arixtra (Res SD: 3.413), **(D)** E90Q versus Arixtra (Res SD: 3.625) and **(E)** E90H versus Arixtra (Res SD: 9.876). The residual standard deviation (Res SD) describes the statistical estimate of the fit of the experimental data to the model, a value below 10 is considered a good fit.

**Figure 32:** Glu-90 was mutated to a lysine, arginine, histidine and a glutamine. Highlighted above are the structures of the wild type and mutant residues with the

length of the amino acid side chains indicated and the charges of the side chains highlighted by the red plus symbol and the blue negative symbol.

**Figure 33:** Sensograms of the interaction of wild type 3-OST-1 with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.

**Figure 34:** Sensograms of the interaction of E90R with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.

**Figure 35:** Sensograms of the interaction of E90H with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.

**Figure 36:** Sensograms of the interaction of E90Q with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.

**Figure 37:** Sensograms of the interaction of E90K with Arixtra. (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.

**Figure 38:** Position-90 substituted with a glutamine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotomer that the glutamine residues conforms to, a clash may occur. The positions that a glutamine substituted at position-90 may dash with are indicated by the blue arrows.

**Figure 39:** Position-90 substituted with an arginine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotomer that the arginine residues conforms to, a clash may occur. The positions that an arginine substituted at position-90 may dash with are indicated by the blue arrows.

**Figure 40:** Position-90 substituted with a histidine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotomer that the histidine residues conforms to, a clash may occur. The positions that a histidine substituted at position-90 may dash with are indicated by the blue arrows.

**Figure 41:** Position-90 substituted with a lysine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotomer that the lysine residues conforms to, a clash may occur. The positions that a lysine substituted at position-90 may dash with are indicated by the blue arrows.

**Figure 42:** Illustration of key features in Surface Plasmon Resonance.

# List of abbreviations

**AT**→ Antithrombin.

**BCIP/NBT**→ 5-bromo-4-chloro-3'-indolyphosphate p-toluidine/nitro blue tetrazolium.

**BSA**→ Bovine serum albumin.

**CAS**→Computational Alanine Scanning.

**CPMA**→ Count per minute average.

**DEAE**→ Diethylaminoethyl.

**EDC**→1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride.

**ELISA**→ Enzyme linked immunosorbant assay.

**EST**→ Estrogen sulphotransferase.

**EXT-1**→ Exostonsin glycosyltransferase 1.

**EXT-2**→ Exostonsin glycosyltransferase 2.

**G**→ Glycoprotein.

**GAG**→ Glycosaminoglycans.

**GlcN** → Glucosamine.

**GlcNH<sub>2</sub>**→ N-unsubstituted glucosamine.

**GlcNAc** → N-acetylglucosamine.

**GlcNH<sub>2</sub>** → N-unsubstituted Glucosamine.

**GlcNS** → N-sulphoglucosamine.

**GlcUA** → Glucuronic acid.

**HIC**→ Hydrophobic interaction chromatography.

**HSV-1** → Herpes Simplex Virus Type 1.

**HVEM** → Herpes Virus Entry receptor.

**HS** → Heparan Sulphate.

**HPLC**→ High performance liquid chromatography

**IdoUA** → Iduronic acid.

**IMAC**→ Immobilized metal ion affinity chromatography

**LMWH**→ Low molecular weight heparin.

**MS**→ Mass Spectrometry.

**NMR**→Nuclear Magnetic Resonance.

**NDST** → N-deacetylase/ N-sulphotransferase.

**NHS**→ N-hydroxysuccinimide.  
**3-OST** → 3-O-sulphotransferase.  
**2-OST**→ 2-O-sulphotransferase.  
**6-OST**→ 6-O-sulphotransferase.  
**PAP**→ 3' phosphoadenosine-5' phosphate.  
**PCR**→ Polymerase chain reaction.  
**SDM**→ Site directed mutagenesis.  
**SHIPREC**→ Sequence homology independent protein recombination.  
**SM**→ Saturation mutagenesis.  
**SPR** → Surface Plasmon Resonance.  
**sTEP**→ Staggered extension protocol.  
**TIR**→ Total internal reflection.  
**TMB**→ 3,3',5,5'-tetramethyl benzidine.  
**UA**→ Uronic acid.  
**ULMWH**→ Ultra low molecular weight heparin.

# Acknowledgments

I would like to thank my supervisor Prof. Robert Woods for taking me on as a PhD student in his lab. I would also like to thank the School of Chemistry and Science Foundation Ireland for the financial support and providing the means to undertake my PhD studies.

I would like to thank Elisa Fadda and Nina Weisser for their help and for the great example they set over the years. To Jenifer Hendel, Oliver Grant, Joanne Martin, Valerie Murphy and Hannah Smith thanks for all your help. Throughout my time in the lab it has been great to get to know you all and the support you have given is really appreciated.

I would like to especially thank my parents Paula and Gerry without your support this would not have been possible. I would like to thank my brothers Ronan and Alan, my sisters Lisa and Ciara, your support has been great throughout the years.

# Chapter 1: Introduction

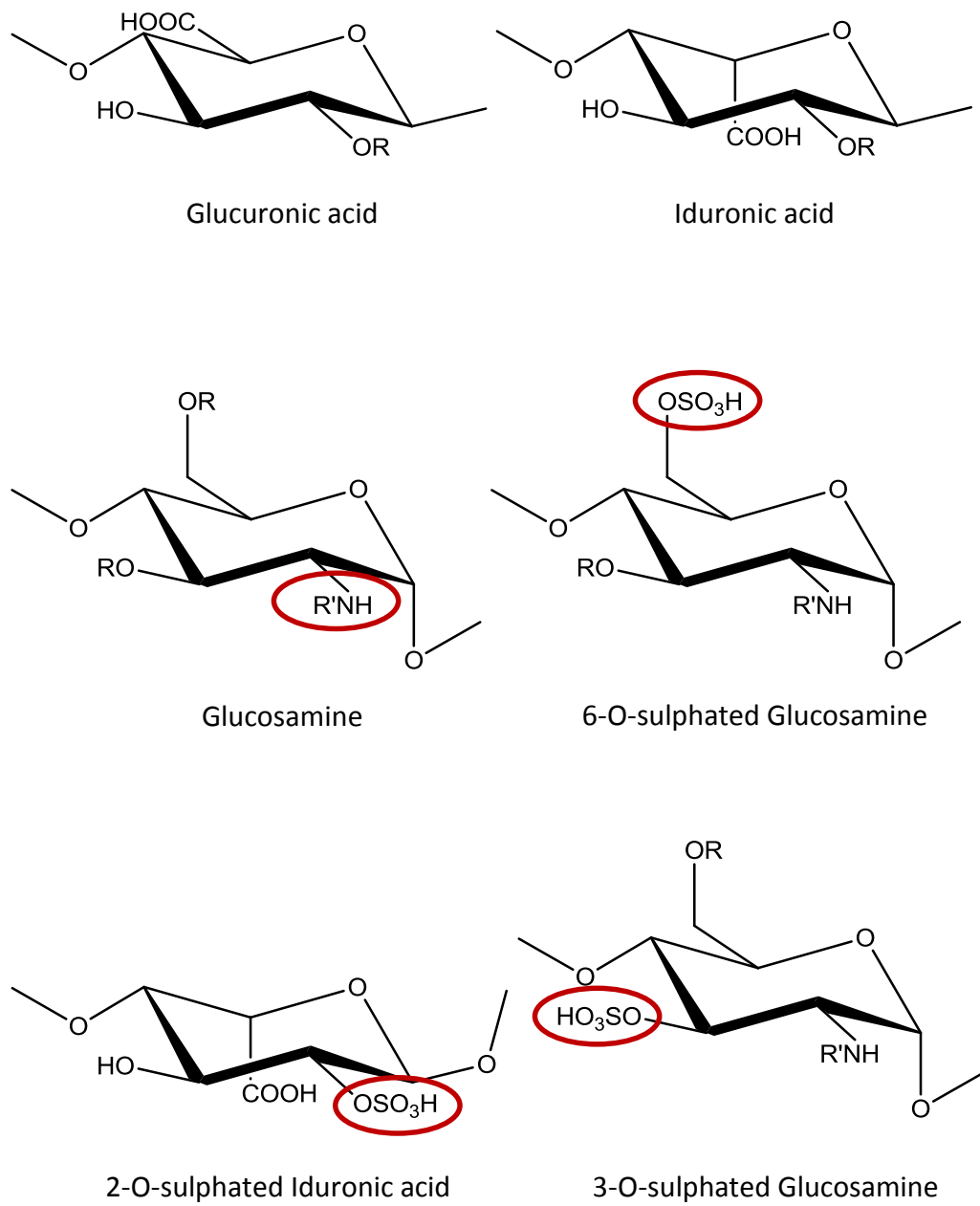
## **Introduction Part 1: Heparin and 3-O-sulphotransferases**

### **Differences between Heparin and Heparan sulphate**

Heparan Sulphate (HS) has been found to play a role in regulating a variety of physiological and pathological functions, from embryonic development and blood coagulation to facilitating viral infection (Esko and Selleck 2002) (Liu and Thorp 2002). Heparan Sulphate (HS) is closely related to Heparin (H), a linear polysaccharide made up of the same building blocks with however a higher degree of sulphation compared to HS (Yang, Solakyildirim et al. 2011) (Tkachenko, Rhodes et al. 2005). HS is a highly charged polysaccharide that is made up of regions of low sulphation with alternating regions of high sulphation with a molecular weight ranging between 5 and 70 kDa (Yang, Solakyildirim et al. 2011). HS is composed of repeating glucuronic and/or iduronic acid units that are 1→4 linked to glucosamine, with 32 different disaccharide building blocks and  $>10^{14}$  possible sequence variations (Yang, Solakyildirim et al. 2011) (Tkachenko, Rhodes et al. 2005).

Glucuronic acid is a carboxylic acid derivative of glucose formed through C-6 oxidation and iduronic acid is a C-5 epimer of glucuronic acid. Glucosamine is a monosaccharide amino sugar that includes an amine ( $\text{NH}_2$ ) at C-2 which can be acetylated. Iduronic acid residues can be sulphated at the 2-O position. Glucosamine residues can be sulphated at the N-, 6-O and/or the 3-O positions (See Figure 1).

In heparin the majority of glucosamine residues are N-sulphated and approximately 5% are N-acetylated. Heparin oligosaccharides are covalently bound to mast cells via serlycin core proteins, whereas HS is attached to the cell surface of most cells via mainly syndecans or glypicans (Tkachenko, Rhodes et al. 2005).



**Figure 1:** Main building blocks of Heparan sulphate and Heparin, the positions of the sulphate groups highlighted in red.  $R = \text{SO}_3^-$  or an H atom,  $R' = \text{SO}_3^-$  or a  $\text{COOH}_3$ .

**In vivo synthesis of heparan sulphate:**

HS synthesis begins with the addition of an *N*-acetyl-glucosamine residue to a core oligosaccharide consisting of Xyl-Gal-Gal-GlcA that is attached to a structural protein via a serine residue in the Golgi apparatus. The addition of this residue is catalysed by the enzyme  $\alpha$ -*N*-acetylglucosaminyltransferase 1 (Sugahara and Kitagawa 2002). The Heparin chain is extended through the sequential addition of *N*-acetyl-glucosamine residues followed by alternating glucuronic acid residues. This is achieved via complex of the enzymes EXT1 glycosyltransferase 1 and EXT2 glycosyltransferase 2 (Turnbull 2001) (Kreuger and Kjellén 2012) (See Figure 2).

After the synthesis of the HS polymer a series of modifications take place in the Golgi apparatus. First *N*-acetyl-glucosamine is deacetylated and then *N*-sulphated by *N*-deacetylase/*N*-sulphotransferase (NDST). NDST is a type 2 transmembrane protein that exists in four different isoforms which share 65-70% sequence identity (Aikawa and Esko 1999). Site-directed mutagenesis (SDM) has revealed that Lys-614, Phe-640, Glu-641, Glu-642, Asn-647 and Ile-643 are important for NDST activity (Sueyoshi *et al*, 1998)(Kakuta, Li *et al*. 2003).

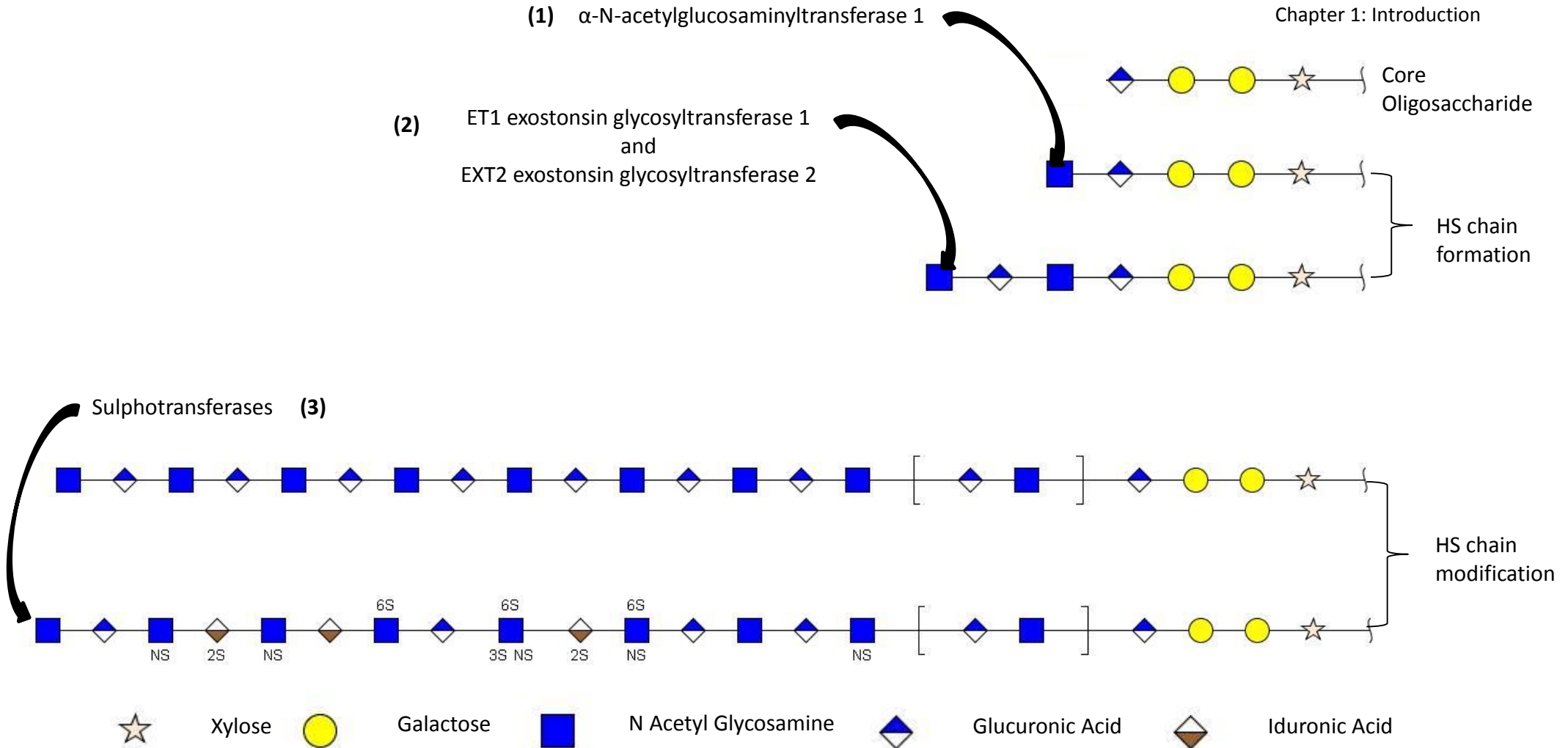
The glucuronic acid units are C-5 epimerized to form iduronic acid which can be 2-O-sulphated by 2-O-sulphotransferase (2-OST). In a study performed by Xu *et al*, alanine scanning mutagenesis was used to identify critical residues for the catalytic activity of HS 2-OST. The residues His<sup>140</sup> and His<sup>142</sup> were identified as being critical for enzymatic activity (Xu, Moon *et al*. 2008). C-5-epimerase is a type 2 transmembrane protein and 2-OST is a type 2 transmembrane protein that is thought to act as a trimer *in vivo*. (Pinhal, Smith *et al*. 2001) (Bethea, Xu *et al*. 2008) (Crawford, Olson *et al*. 2001).

The glucosamine residues can be 6-O-sulphated by different isoforms of 6-O-sulphotransferase (6-OST). 6-OST is a type 2 transmembrane protein that exists in three different isoforms (Habuchi, Tanaka *et al*. 2000). SDM studies performed on 6-OST-1 have revealed that a number of residues are critical for enzymatic activity. Mutations Leu-477, Leu-478 and Leu-483 lead to a marked decrease in enzymatic

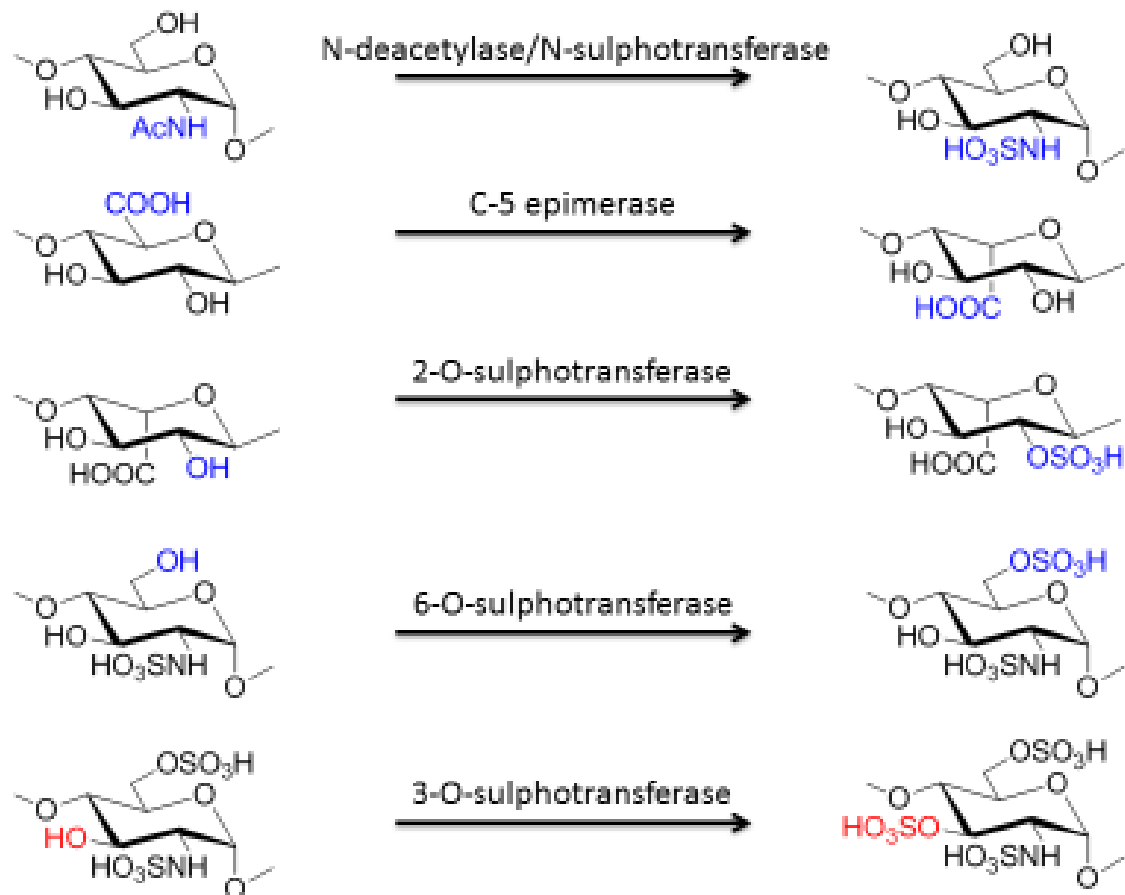
activity (Chen, Ichihara-Tanaka et al. 2004). The final modification involves the glucosamine residues, which become 3-O-sulphated by different isoforms of 3-O-sulphotransferase (3-OST). All of these modifications can occur at different amounts which result in the varying composition of HS.

Heparin is synthesized in a similar fashion in mast cells where it is attached to the core structural protein serglycin (Duelli, Rönnberg et al. 2009). Upon completion of heparin synthesis in mast cells the newly synthesised chains are cleaved by endo- $\beta$ -D-glucuronidase generating shorter chains that are stored in the cytoplasmic granules of the mast cells (Carlsson, Presto et al. 2008). HS is localised on the cell surface and the extra-cellular matrix of most cells presented as a glycosaminoglycan (GAG).

Modifications in the structure of heparin have been related to certain diseases and cancers. Studies have shown that the expression of 2-OST is upregulated in an LNCaP-c4-2B model of prostate cancer and is thought to aid its proliferation. The upregulation of 2-OST in the cancer model was suggested to be hypoxia induced (Ferguson and Datta 2011).

(1)  $\alpha$ -N-acetylglucosaminyltransferase 1(2) ET1 exostonsin glycosyltransferase 1  
and  
EXT2 exostonsin glycosyltransferase 2

**Figure 2:** Illustration of the formation the Heparan sulphate chain: **(1)** HS synthesis begins with the addition of an *N*-acetyl-glucosamine residue to a core oligosaccharide consisting of Xyl-Gal-Gal-GlcA , catalysed by the enzyme  $\alpha$ -N-acetylglucosaminyltransferase 1. **(2)** The HS chain is extended through the sequential addition of *N*-acetyl-glucosamine residues followed by alternating glucuronic acid residues. This is achieved via complex of the enzymes EXT1 glycosyltransferase 1 and EXT2 glycosyltransferase 2. **(3)** After the synthesis of the HS polymer a series of modifications take place in the Golgi apparatus, via a series of sulphotransferases.



**Figure 3:** In vivo synthesis of heparan sulphate begins with the deacetylation and N-sulphation of glucosamine residues by N-deacetylase/N-sulphotransferase. Next the glucuronic acid residues are C5-epimerized by C5-epimerase to form iduronic acid. The iduronic acid residues are then sulphated by 2-O-sulphotransferase. The glucosamine residues are sulphated at the 6- position by 6-O-sulphotransferase and finally at the 3- position by 3-O-sulphotransferase.

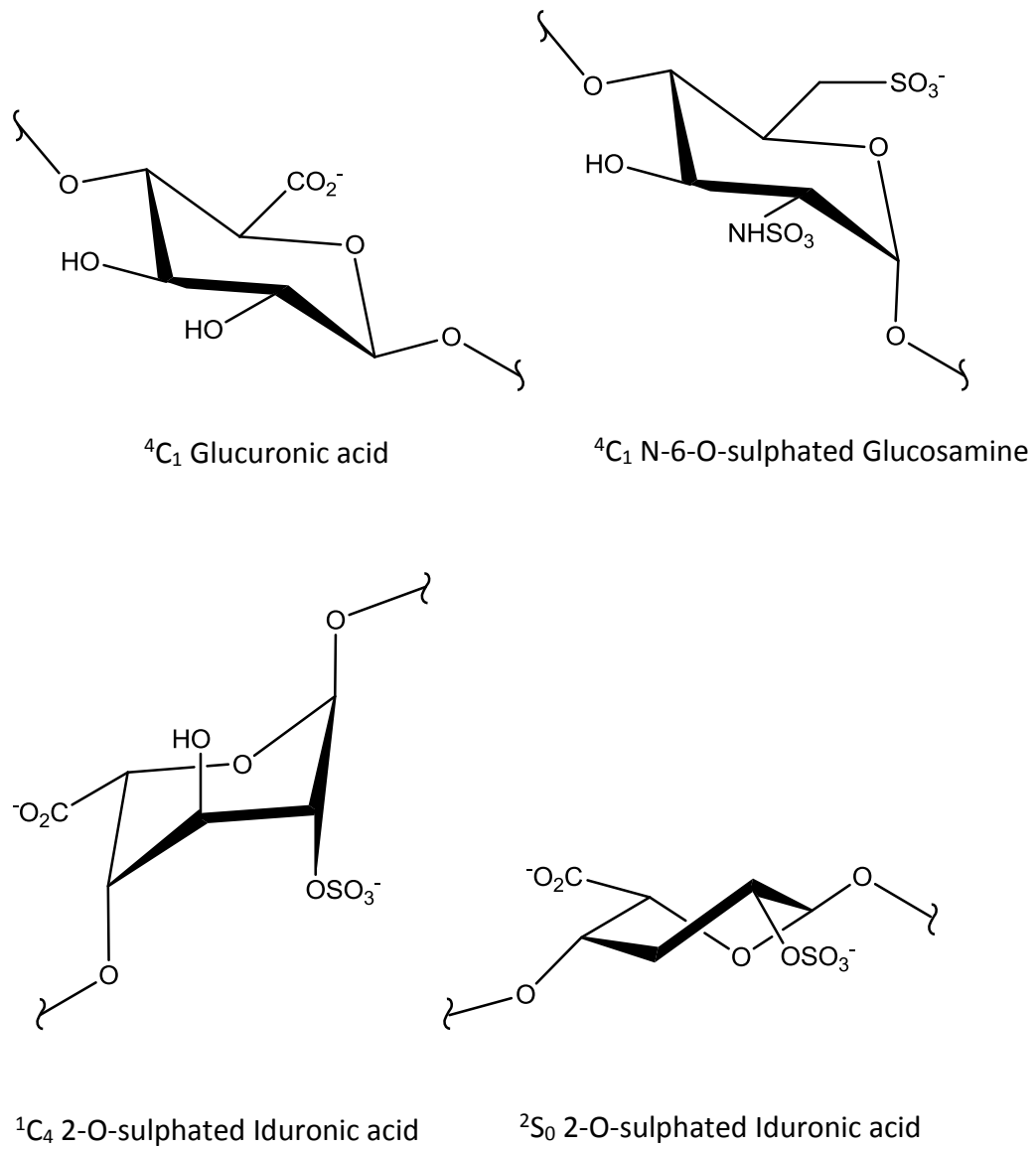
**Variability in sequence, domain and conformation:**

Although HS is structurally related to heparin, heparin is only found in mast cells. HS has a more varied structure and is less sulphated. Also, the minor sequence variants in HS are greater than that found in Heparin, thus allowing for a more diverse and complex structure.

HS and Heparin are charged polysaccharides with highly sulphated domains and alternating unmodified domains (Turnbull 2001). The modifications to heparin result in various domain structures forming. Certain regions can form that are unmodified and are termed N-acetylated (NA) domains due to the presence of glucuronic acid and N-acetyl-glucosamine repeats (Turnbull and Gallagher 1991). These spacer regions are contrasted by highly sulphated regions termed S-domains. The S-domains and the NA-domains can also be flanked by regions with glucosamine residues that are either N-acetylated or N-sulphated which are termed NA/NS domains (Lindahl, Kusche-Gullberg et al. 1998).

Nuclear Magnetic Resonance (NMR) studies have shown that the degree of sulphation of heparin preparations varies between animal species. For example porcine heparin is composed of a trisulphated disaccharide, at around 60-70% (2-O-sulphated iduronic acid linked to a 6-O and 3-O-sulphated glucosamine). However, in bovine heparin, at around 60-70%, this disaccharide has varying degrees of sulphation, 50% are N- and 6-O sulphated while 36% are not sulphated at the 6-O position (Aquino, Pereira et al. 2010).

NMR has shown that glucosamine and glucuronic acid residues are found in stable  ${}^4C_1$  conformations, where iduronic acids adopt  ${}^2S_0$  or  ${}^1C_4$  conformations (Ayotte and Perlin 1986) (See Figure 4). Heparin undergoes a greater amount of sulphation relative to HS. Over 90% of the uronic acid residues are epimerized to iduronic acid (Seo, Andaya et al. 2012).

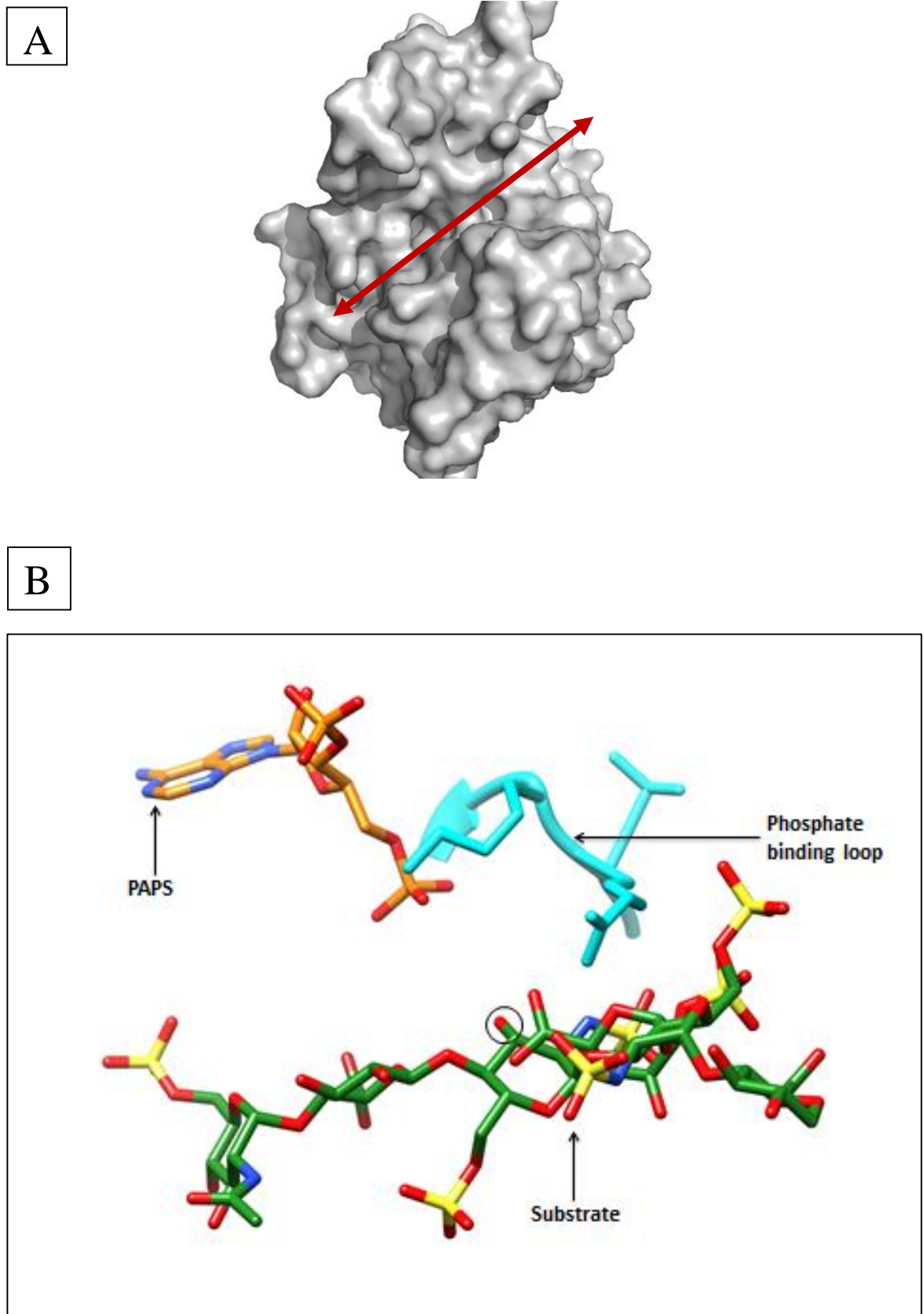


**Figure 4:** Common conformations of the main building blocks of Heparan sulphate and Heparin.

## **What are 3-O-Sulphotransferases?**

Sulphotransferases are a family of enzymes that can be organised into two categories based on their sub-cellular locations, namely Cytosolic or Golgi sulphotranferases, they can be further divided based on the position to which sulpho group is added (Kakuta et al, 1999). As mentioned in the previous section HS is synthesised in a number of enzymatic steps, key to which are the sulphotransferases. The core of the enzyme is composed of an  $\alpha/\beta$ -motif with a P-loop like structure called the phosphosulphate binding loop which plays a role in the binding of the sulpho donor from the 3'-phosphoadenosine 5'-phosphosulphate (PAPS) to a particular position on HS depending on the sulphotransferase involved (See Figure 4). The acceptor substrate binding site is located in a large cleft running across the active site (Moon, Edavettal et al. 2004). (See Figure 5)

3-OST a Golgi sulphotranferases exist in at least seven different isoforms (HajMohammadi, Enyoji et al. 2003). In terms of sequence homology, the amino acid sequences of the different isoforms have 50%-80% homology in the sulphotransferase domains (Liu and Rosenberg 2002). The different 3-O-sulphotransferase isoforms transfer a sulpho group from PAPS to the 3-OH position of the glucosamine units depending on the different saccharide sequences, thus resulting in modified HS with different biological activities. The 3-OST-1 isoform transfers a sulphate group to the 3-OH position of a N-sulphoglucosamine residue linked to a glucuronic acid, which results in antithrombin binding activity (Shworak, Liu et al. 1997). 3-OST-2, which is found in the brain, modifies HS in the brain tissue resulting in the formation of HS which assists in Herpes Simplex virus entry (O'Donnell, Tiwari et al. 2006).



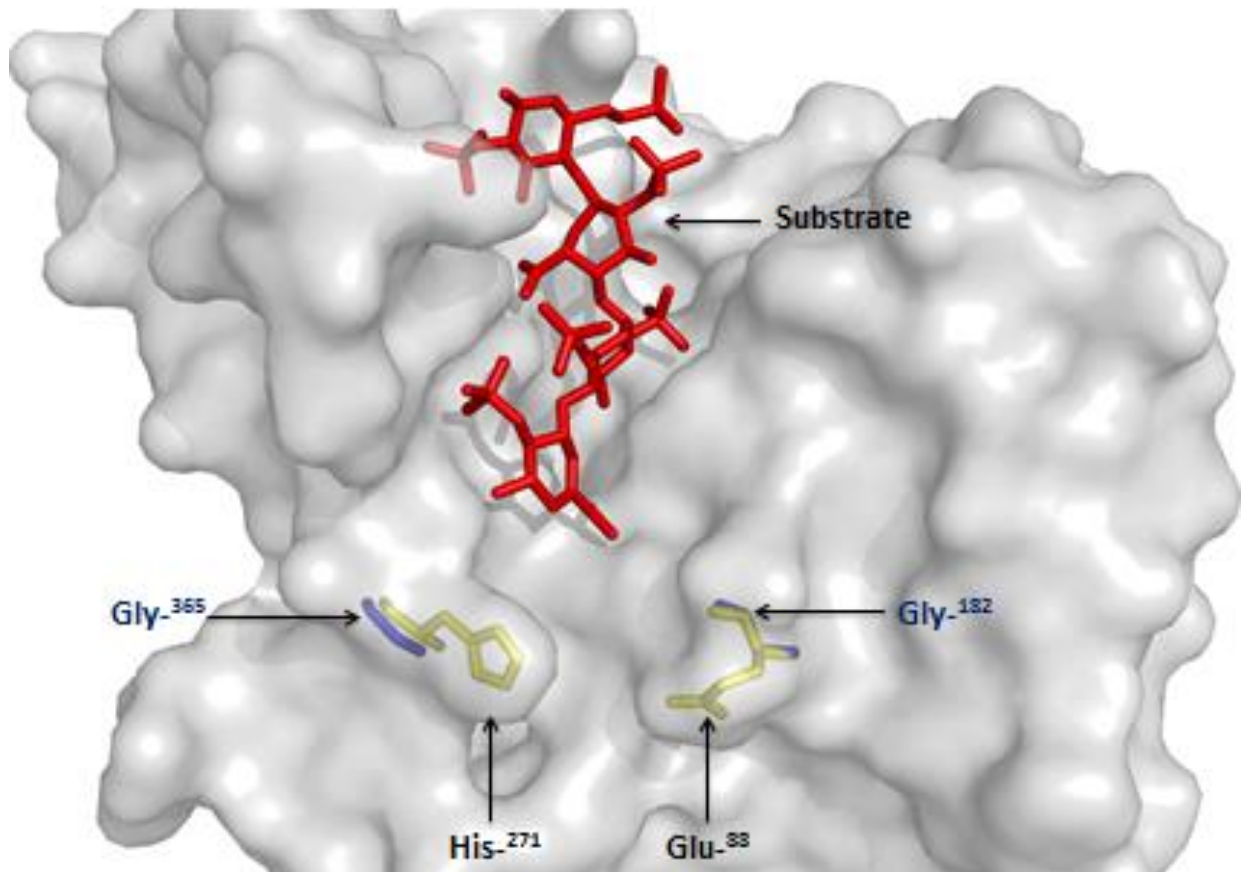
**Figure 5:** Panel A the active site of 3-OST-1 indicated by the arrow. Panel B the Phosphate-binding loop in relation to PAPS and the substrate is indicated. The 3-OH position of the acceptor sugar is circled.

3-OST-3 transfers a sulphate group to the 3-OH position of N-acetylglucosamine 1-4 linked to an iduronic acid. The 3-OST-3 isoform can be subdivided into 3-OST-3 A and 3-OST-3 B, with 3-OST-3 B sharing 92.2% homology with that of 3-OST-3 A and it produces identical HS products, namely assisting in Herpes Simplex virus entry (Shukla, Liu et al. 1999) (Liu, Shriver et al. 1999). The 3-OST-4 isoform is expressed predominantly in neurons and may play a role in assisting herpes simplex viral entry (Lawrence, Yabe et al. 2007). HS modified by 3-OST-5 has been shown to bind to Antithrombin and glycoprotein D (Xia, Chen et al. 2002). HS modified by 3-OST-6 has been found to assist in viral entry (O'Donnell, Kovacs et al. 2010).

### **Substrate recognition of 3-OST-1 and 3-OST-3:**

In 3-OST-1 modified HS, the oligosaccharide is sulphated at the 3-OH position of *N*-sulpho-glucosamine linked to a glucuronic acid. However, 3-OST-3 transfers the sulpho to the 3-OH position of the *N*-un-substituted glucosamine linked to a 2-O sulphated iduronic acid. In a study performed by Xu *et al*, an analysis of the crystal structure showed that in 3-OST-1 and 3OST-3 there are residues that form a “gate” across the large cleft structure as seen in Figure 5. Mutational analysis indicates these residues are important for substrate specificity (Xu, Moon et al. 2008). In 3-OST-1 residues Glu-88 and His-271 form a “gate” across the large cleft. This study also indicates that the width of the gate is important for substrate specificity. In 3-OST-1 the gate is 6.7Å in width and in 3-OST-3 the gate is 14.2Å in width (Xu, Moon et al. 2008) (See Figure 6).

The overall conformation and sulphation pattern of heparin/HS has also been suggested to play a role in substrate recognition between 3-OST-1 and 3-OST-3 (Moon, Xu et al. 2012). Lys-259 is thought to play a role in the substrate interaction of 3-OST-3. The in 3-OST-1 this residue is an asparagine, Asn-167 (Moon, Edavettal et al. 2004). 3-OST-1 and 3OST-3 are found to be expressed in a variety of different tissues and they share similar tissue distributions (Mochizuki, Yoshida et al. 2008).

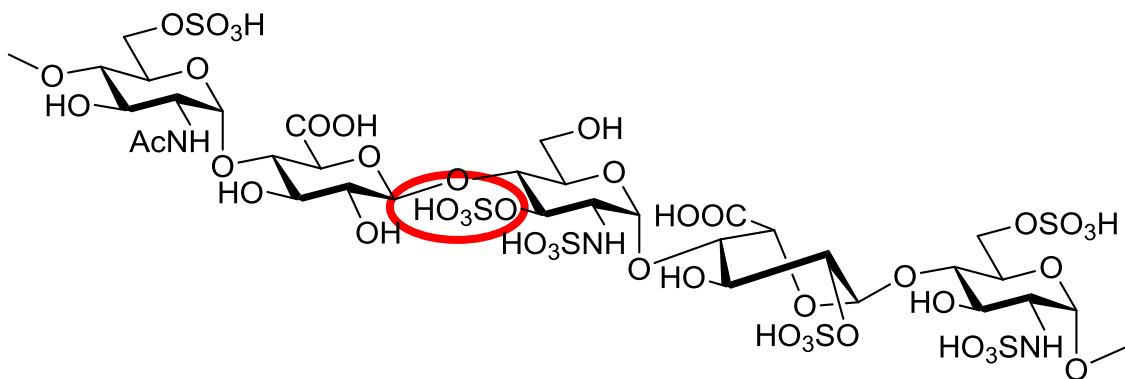


**Figure 6:** Gate structure that has been noted for its importance in substrate specificity. Residues that form the gate structure in 3-OST-1 in yellow, residues that form the gate structure in 3-OST-3 in blue.

**3-OST-1 modified Heparan Sulphate and the Blood Coagulation Pathway:**

3-OST-1 and 3-OST-5 modified HS binds to Antithrombin (AT) leading to a conformational change in AT to its active form. This active form binds to the blood coagulation factors Xa and thrombin leading to their inactivation (Li, Johnson et al. 2004). Important to the ability of HS to activate AT is the sulpho group located at the 3-OH position (See Figure 7). Atha *et al*, has shown that HS which lacks this essential sulpho group have been found to have 20,000-fold less AT binding activity (Atha, Lormeau et al. 1985). Research has shown that only a small fraction of HS isolated from animal tissue will bind to AT (Li, Johnson et al. 2004).

Studies on the AT binding sequence of HS has led to the development of a number of important anti-coagulant drugs, one of which is Arixtra. This synthetic pentasaccharide contains the binding partner for AT. Arixtra acts as an indirect inhibitor of the blood coagulation factor Xa through the binding of Arixtra to AT in the same manner as that of HS where it is used in the treatment of venous thromboembolisms.



**Figure 7:** Heparin sulphate Antithrombin binding sequence represented here by the structure for Arixtra, with 3-O-sulphation indicated in red.

### **3-OST-1 modified Heparin in the pharmaceutical industry:**

In principal, the large scale preparation of HS involves a number of steps beginning with the preparation of the tissues from which HS is to be isolated, typically porcine intestine. The crude HS is isolated and purified by cation chromatography to remove any unwanted material (Linhardt and Gunay 1999). However in 2008, contaminated heparin led to a number of reports of diarrhoea, dizziness and headaches and in some patients pharyngeal edema and tachycardia. A total of 256 deaths were associated with heparin use from batches supplied from China. The main cause of the contamination was attributed to over sulphated chondroitin sulphate (Blossom, Kallen et al. 2008).

Isolates are typically tested for bacterial endotoxins. A coagulation assay is typically performed, which involves the amount of HS expressed in 1 mg that will cause 1 ml of sheep plasma to half-clot after incubation at 20 °C. Ion exchange chromatography and pyrolysis can be used to determine the level of sulphation. Capillary electrophoresis can be used to give an indication of the level of AT binding sites in a particular batch (Desai, Wang et al. 1993). NMR can be used to determine the ratios of iduronic acid to glucuronic acid and the ratio of N-acetyl to N-sulpho-glucosamine (Toida, Yoshida et al. 1997).

The enriched heparin can be analysed by strong anion exchange high performance liquid chromatography (HPLC) for purity followed by determination of the molecular weight by size exclusion chromatography (Guo, Condra et al. 2003) (Seo, Andaya et al. 2012). Screening of heparin for contaminants as required by the FDA is carried out using Capillary electrophoresis and Nanoscale magnetic resonance (Desai, Wang et al. 1993) (Toida, Yoshida et al. 1997).

The above techniques provide a useful means of isolating and characterising complex carbohydrates, however there is an urgent need for reagents to analyse such biological samples, where purified GAGs would benefit from carbohydrate specific reagents. Currently there is a lack of useful reagents for studying and distinguishing between

complex GAGs like heparan sulphate and chondroitin sulphate. Converting heparan sulphate processing enzymes into high affinity carbohydrate binders would provide a useful means of overcoming this problem and would also aid in studying heparan sulphate. A carbohydrate specific biosensor would potentially provide a useful approach to this problem (Packer, von der Lieth et al. 2008).

Pharmaceutical grade heparin generates up to 3 billion USD worldwide. Antithrombin binding to heparan sulphate occurs in around 1-10% of heparan sulphate isolated from animal tissue and 30% of pharmaceutical grade heparin (Li, Johnson et al. 2004). Variations occur in the heparin isolated from different sources in relation to their coagulation, thrombosis and bleeding effects and there is a greater degree of 3-O-sulphation in porcine derived Heparin (Tovar et al, 2012).

Low molecular weight heparins (LMWH) are produced primarily through various methods of depolymerisation and fractionation of native heparin. These products have reduced side effects such as the excessive bleeding that can be seen with unfractionated heparin. There are additionally varying degrees of 3-O-sulphated glucosamine residues present in LMWH's like tinzaparin, enoxaparin and dalteparin.

Current methods of generating LMWH's can cleave the glycosidic bonds both outside and inside the AT binding sequence. In addition within LMWH's there are both high affinity and non-antithrombin binders reducing their efficacy. The ability to enrich heparin isolates for specific structural features like glucuronic acid  $\beta$  1-4 linked to a 3-O-sulphated glucosamine prior to the depolymerization of the native heparin has the potential to increase the anticoagulant activity of LMWH's.

**3-OST-3 modified Heparan sulphate and herpes simplex virus:**

Herpes Simplex virus type 1 (HSV-1) and HSV-2 are members of the Herpesviridae family. HSV infection is attributed to the cause of mucutaneous infections such as oral-facial infections, cutaneous infections, herpes genitalis, perineal herpes and neonatal or disseminated herpes (Kleymann 2003). HSV is made up of double-stranded linear DNA contained in a capsid surrounded by a tegument and a glycoprotein envelope. HSV infect cells that carry viral entry mediators and negatively charged structures such as HS. A key characteristic of HSV is the establishment of latent infection in the host autonomic nervous system (Kleymann 2003).

There are a number of viral glycoproteins that play a critical role in viral entry. Glycoproteins C and B have been shown to mediate the binding of the virus to HS. Mutagenesis studies have shown that glycoproteins (g) D, gB, gH and gL are required for viral entry when they work together (Spear, Eisenberg et al. 2000). However research has shown that the absence of gD, gB, gH and gL or any single glycoprotein of the above will allow viral host cell binding but will prevent viral entry (Garner 2003).

The mechanism by which HS assists herpes simplex type 1 (HSV-1) viral entry begins with the interaction of the viral envelope glycoprotein's C and B and HS on the host cell surface (Mettenleiter 1989) (Trybala, Bergström et al. 1994). In the context of glycoprotein C and HS binding, studies show that binding requires a minimum of 10-12 residues that contain IdoUA2S and GlcNS (or Ac) 6S for successful interaction (Feyzi, Trybala et al. 1997) (Herold, WuDunn et al. 1991).

Glycoprotein D reacts to a number of entry receptors such as nectin-1, Herpes Virus Entry Mediator (HVEM) and 3-O-sulphotransferase (3-OST)-modified HS (Geraghty, Krummenacher et al. 1998). The N-terminal region of gD contains the binding site which spans 1-260 amino acids (Whitbeck, Muggeridge et al. 1999). Research has indicated that the 3-OST-modified HS binding site on gD is near the HVEM binding site (Carfí, Willis et al. 2001).

Research performed by Liu *et al* has shown through sequence analysis that a purified gD-binding octasaccharide from a HS affinity library modified by 3-OST-3 binds to gD, however the exact contribution of each sulphate group to binding of gD is not known (Liu, Shriver et al. 2002). This does suggest the possibility of a structural similarity between the purified octasaccharide and the actual 3-OST-modified HS-gD binding site. Research has shown through X-ray crystallography that the interaction between gD and HVEM is centred around a protruding tyrosine of HVEM that fits into a pocket on gD, researchers indicated that the pocket may be the binding site for 3-OST-modified HS (Carfi, Willis et al. 2001).

### **3-OST-1 and important residues and structural features:**

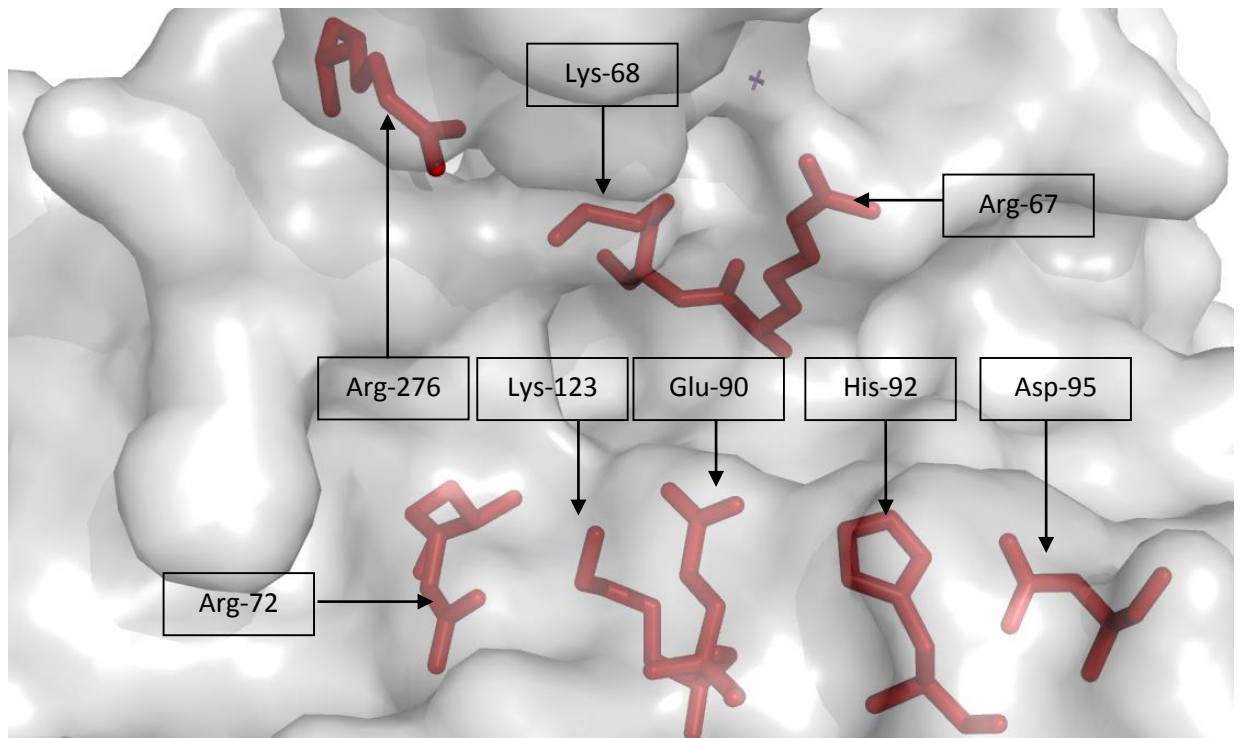
A study of the 3-OST-1 precursor protein cDNA shows that the enzyme undergoes the removal of an amino-terminal leader sequence leading to the formation of a Golgi intraluminal resident consisting of around 290 residues. The enzyme contains a conserved domain of around 260 residues that makes up the sulphotransferase domain (Shworak, Liu et al. 1997). 3-OST-1 is composed of an amino-terminal SPLAG domain fused to a carboxyl-terminal sulphotransferase domain. Unlike previously isolated Golgi enzymes 3-OST-1 does not contain a cytoplasmic domain, the enzyme lacks sufficient length to span membrane as an alpha-helix (Shworak, Liu et al. 1999).

A number of amino acids have been determined experimentally to be important for binding and enzymatic activity. SDM is a laboratory technique involving site specific mutations often employed to study protein structure function relationships. Alanine scanning mutagenesis is an extension of this principal. The process involves mutating specific amino acid residues to an alanine and using assays to determine the effect of the mutation on protein structure function relationships (Matthews 1996).

Moon *et al*, have shown that mutation of Arg-268 in 3-OST-1 results in a significant loss of enzymatic activity suggesting that the binding cleft extends beyond the gate structure generated by His-271 and Glu-88. Heparin polysaccharides that extend from the acceptor sugar at the non-reducing end come into contact with the Arg-268 residue. This suggests the importance of polysaccharide length on enzymatic activity

with respect to substrate interactions away from the catalytic site (Moon, Xu et al. 2012).

Mutagenesis studies can also be employed using mutations other than mutating to an alanine which is typically employed as alanine mutations are less likely to affect protein structure upon mutation (Lefèvre, Rémy et al. 1997). For example in a study performed by Xu *et al*, analysis of 3-OST-1 residues Glu-88 and His-271 found a “gate” structure in which mutagenesis studies reveal that they are essential for substrate specificity (Xu, Moon et al. 2008). Another example involved a mutagenesis study performed by Edavettal *et al* has shown that there are a number of residues that are essential to the enzymatic activity of 3-OST-1. In a study performed by Edavettal et al, mutagenesis studies showed that Arg-67, Lys-68, Arg-72, Glu-90, His-92, Asp-95, Lys-123 and Arg<sup>276</sup> are essential to enzymatic activity (Edavettal, Lee et al. 2004) (See Figure 8).

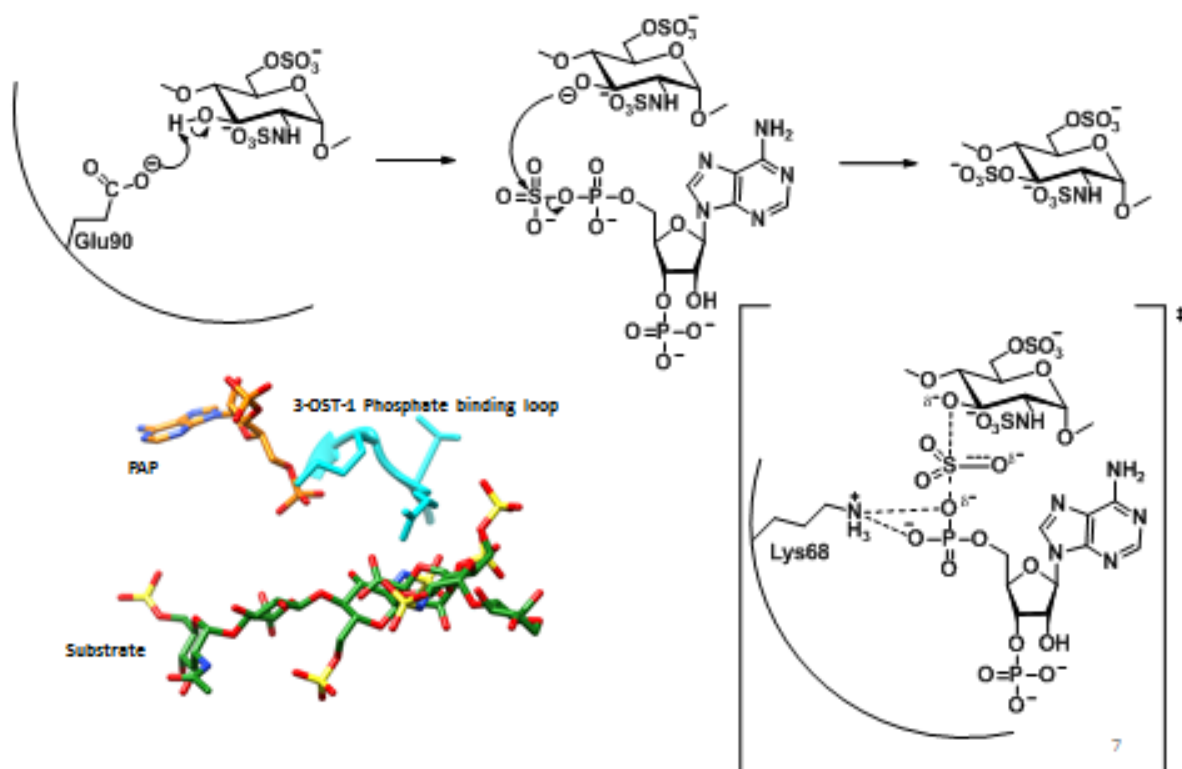


**Figure 8:** Illustration of residues known to be critical to enzymatic activity in 3-OST-1 (Edavettal, Lee et al. 2004).

### **Catalytic mechanism of 3-OST-1:**

The catalytic mechanism of 3-OST-1 is thought to occur through a similar mechanism as that of estrogen sulphotransferase (EST) and other cytosolic sulphotransferases (Gorokhov, Perera et al. 2000). In a study performed by Edavettal *et al*, the crystal structures of 3-OST-1 and PAPS were superimposed over that of EST in complex with PAPS and 17 $\beta$ -estradiol. Analysis of the crystal structures and mutational analysis revealed that 3-OST-1 and EST may share a similar catalytic mechanism (Edavettal, Lee et al. 2004).

The catalytic mechanism of EST is thought to occur through a S<sub>N</sub>-2 like displacement mechanism where the acceptor hydroxyl group acts as a nucleophile. This hydroxyl group becomes deprotonated by a conserved histidine. After which the acceptor group attacks the sulphur atom of the PAPS donor. This results in the formation of a transition state between PAPS and the acceptor group. A negative charge builds up on the leaving PAP group that is stabilised by a conserved lysine in the phosphate-binding loop. This forms a hydrogen bond with the bridging oxygen of the acceptor group and the sulphur atom of the sulphate group (SO<sub>3</sub>) being donated. This lysine is conserved in both EST and 3-OST-1. However 3-OST-1 does not contain a structurally equivalent histidine residue, a glutamic acid (Glu-90) is thought to act as the catalytic base and has been revealed to be essential for sulphotransferase activity (Edavettal *et al*, 2004). (See Figure 9)



**Figure 9:** The catalytic mechanism of 3-OST-1 begins with Glu-90 deprotonating the 3-OH position of a glucosamine residue. This deprotonated position attacks the sulphur atom of the co-factor PAPS. The result is the formation of a transition state. This state is stabilized through Lys-68 of the phosphate binding loop. After a build-up of negative charge a sulphate is transferred to the 3 position on the glucosamine residue.

## **Introduction Part 2: Application of Heparin in research**

### **What is a biosensor?**

Biosensors are defined as an integrated biological receptor-transducer device which is capable of providing selective quantitative or semi-quantitative analytical information. The biosensor recognition element is designed to transfer information about a particular analyte of interest to a transducer (Thévenot, Toth et al. 2001).

In developing protein based biosensors a number of factors are considered including: choosing the biosensor receptor element, defining the immobilization of the receptor or analyte and how the biosensor could potentially be applied. A key factor in biosensor receptor immobilisation is choosing the most efficient means of immobilization to the sensor surface. For example, protein based biosensors are commonly immobilized through covalent attachment of the protein via amine coupling to the sensor surface (O'Shannessy, Brigham-Burke et al. 1992) (Johnsson, Löfås et al. 1995).

Another important aspect of immobilisation in protein-based biosensors is defining the orientation of the biological recognition element. Such considerations are essential for the optimisation of the binding properties. For example, proteins that contain a free sulphate group have been shown to be immobilised with the correct orientation by coupling of thiol via cysteine onto a maleimide-derived carboxymethyl surface (Khilko, Corr et al. 1993). An alternative to biosensor receptor immobilization is the immobilization of the analyte. This approach mimics how analytes are presented *in vivo* and the development of such biosensors provide a means to study receptor analyte binding interactions (Gilligan, Schuck et al. 2002).

SPR based biosensors provide a powerful tool in studying carbohydrate-protein interactions and as a tool SPR based biosensors have a variety of applications (Guo 2012). A key aspect in this approach is choosing the biosensor receptor element with carbohydrate binding proteins providing a means of studying such interactions.

### **Carbohydrate binders and use as biosensors:**

Carbohydrate binding interactions would benefit greatly from the development of biosensors based on carbohydrate processing enzymes which have clearly defined site-specific specificity. Such biosensors, for example, have the potential application as reagents for the analysis of the binding interactions of GAGs and other complex carbohydrates. Lectin based biosensors provide an important means of studying lectin-polysaccharide binding interactions. In a study performed by Ballerstadt and Schultz, an assay for measuring glucose levels was developed using a lectin based biosensor. In this study the lectin based biosensor was developed from concanavalin-A which was subsequently immobilized on Sephadex beads. This lectin based biosensor assay took advantage of the high specificity and affinity of concanavalin-A for glucose (Ballerstadt and Schultz 2000).

### **Benefits of biosensors to heparin research:**

Biosensors based on carbohydrate processing enzymes like heparan sulphotransferases have the potential to help study HS from different sources and different species through taking advantage of the site specific nature of this class of enzymes. HS and other GAGs are typically isolated and purified by size exclusion chromatography, strong anion exchange chromatography, ion pairing reverse phase high pressure liquid chromatography or micro preparative polyacrylamide gel electrophoresis (Ziegler and Zaia 2006) (Volpi and Maccari 2006). Biosensors would provide a means to study such isolates where a reagent based on carbohydrate processing enzyme would be greatly needed (Packer, von der Lieth et al. 2008). Essential to the development of this biosensor technology is defining the interaction between the biosensor receptor element and the analyte and whether the biosensor receptor element is in solution or immobilized on a sensor surface

### **Limitations in the methods for heparin research:**

In applying protein based biosensors to study heparin beyond the choice of the receptor element, there are other important considerations. Without proper considerations such factors may limit the application of protein based biosensors to heparin research. There are inherent difficulties with protein immobilization such as the loss of activity upon immobilization, non-specific protein absorption, improper orientation and conformation (Kröger, Katerkamp et al. 1998). The random orientation of a protein upon immobilization can occur through immobilization through non-specific functional groups. However, through site-directed mutagenesis, mutations can be introduced that allow for the immobilization of a protein with a defined orientation.

There are ways to overcome the inherent difficulties in protein immobilization associated with non-specific protein absorption and incorrect orientation of immobilized proteins. In a study by Liu *et al*, involving a synaptosomal-associated protein containing a His-tag, the protein was immobilized on silicon nanowire through the covalent interaction between the His-tag and nickel ions that were chelated to nitrilotriacetic acid coating on the immobilizing surface (Yi-Chi, Nathalie et al. 2010).

In the development of biosensors based on carbohydrate processing enzymes there are advantages to immobilizing the analyte as opposed to the biosensor receptor element. Important to protein based biosensor development and the study of HS and sulphotransferase interactions, is mimicking the cellular interactions between HS and sulphotransferase. This can be achieved using Surface Plasmon Resonance (SPR) and the immobilization of the substrate to the sensor surface.

### **Heparan Sulphate and Surface Plasmon Resonance:**

SPR is a useful tool for the study between HS and HS binding partners. In a study performed by Ricard-Blum *et al*, the interaction between endostatin and HS was investigated thorough the use of SPR. In this study, HS was immobilized to the SPR sensor surface through the interaction between biotinylated HS and streptavidin which

was covalently immobilized to the sensor surface (Negishi, Pedersen et al. 2001). Alternatively, an emerging method of HS immobilization is the immobilization of HS or heparin through the covalent immobilization of albumin-heparin conjugate through amine coupling (Yang, Solakyildirim et al. 2011). In a study performed by Munoz et al, an affinity, kinetic and structural study of the interaction between 3-OST-1 and HS was performed using SPR (Muñoz, Xu et al. 2006). A benefit of using SPR to characterise macromolecular interactions is there are potential further downstream applications of the analytes.

### **Downstream applications to heparin research:**

Biosensors can be used to affinity purify ligands in addition to analysing biomolecular interactions and binding parameters. For example, research performed by Mattei et al showed the potential of applying SPR as a means of studying endopolygalacturonase glycosylation and as a means of affinity capturing of the analyte for downstream investigation by mass spectroscopy (Mattei, Cervone et al. 2001) (Nedelkov and Nelson 2003).

The use of biosensor surfaces for the purification and identification of ligands has been shown to be more efficient through the elution and recovery of the surface bound ligand, prior to further downstream analysis (Desai, Wang et al. 1993). Research has shown that there is also the potential for the use of biosensors in combination with micro-affinity purification for the elution and recovery of high concentration of ligands from sensor surface (Petitou, Herault et al. 1999).

As with protein based biosensors in general and with affinity purification applications, the development of a specific protein receptor is essential. In applying protein based biosensors to heparin research, of critical importance is the development of specific reagents due to the heterogeneous nature of heparin.

**Need for specific Heparin reagents:**

Fundamental to utilising HS and SPR for the study of the interactions between HS and HS binding proteins, is the development of specific binding reagents. Through protein engineering heparan sulphotransferases can potentially be developed as specific heparin reagents. Techniques such as Surface Plasmon resonance can be used to measure binding affinity of a mutated protein as a means of testing the effect of a site-directed mutation and validating the amino acid residues deemed capable of increasing binding free energy (Turnbull, Powell et al. 2001) (Ayotte and Perlin 1986) (Seo, Andaya et al. 2012).

There are multiple potential uses for these reagents in heparin research from increasing our understanding for the catalytic mechanism of heparin to the use of reagents in affinity purification systems. In exploring such potential applications there are a number of different techniques available to develop specific reagents.

## **Introduction Part 3: Developing Heparin reagents**

### **Analysing Protein-Ligand Interaction using a computational approach:**

A critical aspect of protein engineering is the identification of amino acids which are deemed essential to protein-ligand interaction and binding. These residues are considered essential typically when their mutation would lead to a loss in binding free energy or activity. Experimentally, alanine scanning mutagenesis is a means of introducing alanine mutations to probe for amino acid residues important for protein activity. This process can also be undertaken through a computational approach.

Computational Alanine Scanning (CAS) is a useful tool to identify target residues which have a role in protein structure and function. However target residues identified through the use of CAS need to be verified experimentally. In a study performed by Tuncbag *et al.* the hot spots (critical residues) identified computationally were 70% accurate in relation to the experimentally identified hot spots (Tuncbag, GURSOY *et al.* 2009).

### **Use of mutagenesis to enhance affinity:**

Enhancing affinity can be achieved using mutagenesis through the development of a mutant library. When constructing a mutant library there are a number of approaches available.

### **Error-prone PCR:**

The use of error-prone PCR provides a means of generating a mutant library. The process of error-prone PCR involves increasing the frequency of the incorporation of mutations in a PCR cycle through altering the components of a polymerase chain reaction (PCR) reaction and adjusting the thermocycling conditions (Eckert and Kunkel 1991). Mutations can be introduced through the introduction of unbalanced nucleotide concentrations (Shafikhani, Siegel *et al.* 1997). The introduction of high

concentrations of manganese can lead to mutations by decreasing the specificity of DNA polymerase in conjunction with unbalanced nucleotide (Shafikhani, Siegel et al. 1997). In error-prone PCR, Taq DNA polymerases typically are used due to the lack 3'→5' exonuclease activity of this polymerase. Increasing the concentration of this enzyme, increases the frequency of incorporation of mutations (Landgraf and Wolfes 1993). In generating a library of mutant genes, the incorporation of mutations brought on by the use of error-prone PCR can lead to the generation of mutants which are subsequently transformed into competent cells. The transformants produced can be screened for beneficial properties such as increased ligand binding (Cadwell and Joyce 1992). In a study by Leemhuis *et al*, error-prone PCR was used to generate a mutant library that yielded mutants of cyclodextrin glycosyltransferase with 90-fold increase in hydrolytic activity (Leemhuis, Rozeboom et al. 2003).

### **DNA shuffling:**

DNA shuffling is another method of generating a mutant library. The process involves digesting a target DNA sequence with DNase I to generate DNA fragments with random lengths which are subsequently reassembled using PCR. The resulting chimeric DNA sequences are subsequently expressed and the mutant proteins are screened for beneficial properties. Incremental truncation for the creation of hybrid enzymes (ITCHY) was developed from DNA shuffling. The process begins with the digestion of two parental genes with exonuclease III to yield truncated DNA. The 5' fragments of one gene are fused to the 3' fragment of another to generate a chimeric mutant protein library (Ostermeier and Benkovic 2001).

### **Sequence homology independent protein recombination:**

Sequence homology independent protein recombination (SHIPREC) is a method that involves fusion of two genes and the subsequent digestion by DNase I in the presence of manganese to yield fragments of DNA of random length. The fragments are then treated with  $S_1$  nuclease to produce blunt ended DNA which is fused together through circularization. In a study by Sieber et al, screening of a SHIPREC mutant library yielded

a soluble human cytochrome P450 enzyme with increased solubility in bacterial cytoplasm compared to the wild type enzyme even though the hybrid gene shared only 16% amino acid sequence identity (Sieber, Martinez et al. 2001).

### **Staggered extension protocol:**

Another method of generating a mutant library called the staggered extension protocol (stEP) involves using repeated cycles of denaturation and short annealing and extension stages of the template sequences. The process yields growing DNA fragments which anneal to different template sequences based on complementarity which in turn get extended further until a full length sequence is generated (Zhao, Giver et al. 1998). In a study performed by Zhao *et al*, screening a mutant library using this process generated a mutant subtilisin E with a half-life 50 times greater than the wild type enzyme (Zhao, Giver et al. 1998).

### **Saturation mutagenesis:**

Saturation mutagenesis (SM) is a process which involves mutating a critical amino acid residue to all possible residues and then screening the mutant library generated for improved protein function. In a study by Hogrefe *et al*, a kit supplied by Qiagen was used to perform site-directed mutagenesis on up to three sites simultaneously (Hogrefe, Cline et al. 2002). In a study performed by Wang *et al*, SM was used to generate a novel pyridine nucleotide-dependant dehydrogenase and the study showed that SM can be performed at multiple sites simultaneously (Wang, Zhang et al. 2007).

### **Site-directed mutagenesis:**

Site-directed mutagenesis (SDM) has the potential to provide answers to a specific question in a more direct and controlled way than that of techniques like DNA shuffling. In addition, SDM has the potential to answer specific questions that do not necessarily require undertaking the extensive approach used in saturation mutagenesis. In a study performed by Bethea et al. structurally guided mutagenesis

was used to study the catalytic mechanism of 2-OST, through the use of SDM insight was provided into residues important for substrate binding (Bethea, Xu et al. 2008).

In a study performed on a therapeutic antibody to methamphetamine, single point mutants were found to improve affinity. A mutant of scfv S93T showed a 3.1 fold improvement of affinity for methamphetamine and amphetamine. The I37M and mutant Y34M mutant showed a 94 and 8 fold affinity enhancement for amphetamine respectively using SDM (Thakkar, Nanaware-Kharade et al. 2014).

Single point mutations have the potential to enhance affinity and I aim to use SDM to develop a heparin specific reagent. Preliminary computational work shows that in the case of 3-OST-1 complexed with heparin that the catalytic base Glu-90 contributes negatively to the binding of the sulphated product. We aim to investigate the effect of amino acid substitutions at position 90, specifically the effect of the addition of certain amino acid side chains and the introduction of a positive charge at position 90. We aim to show that the change in side chain and the introduction of a positive charge will enhance the binding to 3-O-sulfated heparin with respect to the wild type. This effect will be investigated through Surface plasmon resonance (SPR) and Enzyme linked immunosorbant assay (ELISA).

The development of such a reagent has great potential in heparin research as a heparin specific reagent. Also with possible downstream applications in affinity purification to target 3-OST-1 modified heparin which would have a potential benefit to the pharmaceutical industry.

# Chapter 2: Materials and Methods

## **Materials**

### **Acknowledgment to Prof. Jian Liu:**

I hereby acknowledge the contribution of Prof. Jian Liu of Chapel Hill North Carolina in providing the protein expression vector (pET 28a) used in the generation of the 3-OST-1 mutants, the [<sup>35</sup>S] PAPS and the heparin used in the enzymatic activity assay. I would also like to acknowledge the opportunity to undertake the initial training in North Carolina.

### **Kits used:**

- Stratagene QuickChange II Site-directed mutagenesis kit<sup>®</sup> (Agilent) (Product code: 200521).
- QIAprep Spin miniprep kit<sup>®</sup> (Qiagen) (Product code: 27106).
- Invitrogen Platinum *Pfx* DNA polymerase<sup>®</sup> (Invitrogen) (Product code: 11708-013).
- Phusion high fidelity DNA polymerase<sup>®</sup> (New England Biolabs) (Product code: M0530S)

### **DNA and protein ladders:**

- Pre-stained, Ez-Run Rec ladder (Fisher Scientific) (Product code: 10785674).
- DNA ladder exACTGene 13 bands general purpose ladder digested DNA 1kb 300-1000kb (Fisher Bioreagents) (Product code: 10489883).
- Low Scale Ladder 100 bp (Thermoscientific) (Product code: BP25881-200)

### **Antibodies:**

- Penta-His antibody, BSA free (Qiagen) (Product code: 34660).
- ImmunoPure Goat-anti-mouse IgG, Alkaline Phosphatase conjugate (Life technologies) (Product code: 6-21060).
- Monoclonal anti-poly His HRP (Alpha Diagnostics) (Product code: HISP12-HRP).

### **Restriction Enzymes:**

- *Dpn I* (New England Biolabs) (Product code: R01765).

### **Culture media:**

#### **Protein expression media:**

3.2% Tryptone, 2% yeast extract and 0.5% NaCl

### **Materials:**

- Amicon ultra-15 centrifugal filter units® (Millipore) (Product code: UFC901008).
- Cellulose nitrate membrane filters 0.45 nm filter (Sartorius) (Product code: 13006-50-N).
- Ministart syringe filter 0.45 nm filter (Sartorius) (Product code: 17574-Q).
- Western blotting paper (BioRad) (Product code: 1703966).
- Immunobilon-P PVDF transfer membrane (Thermoscientific) (Product code: IPVH00010).
- Streptavidin Bio Bind Assembly 96 well plates (Thermoscientific) (Product code: 95029293)
- Maxisorb 96 well plates (eBioscience) (Product code: 11190932).
- Microtest 96 well plates (Sarstedt) (Product code: 82-1581).
- Minimal Binding 96 well plate (Greiner) (Product code: 650901).
- Heparin (Alpha Aesar) (Product code: 9041-08-1).
- DEAE sepharose (GE healthcare) (Product code: 71-0709-01).
- Ni sepharose (Qiagen) (Product code: 30230).

### **Columns:**

- 5 ml His trap FF crude column® (GE healthcare) (Product code: 11-0004-58).
- 1 ml Phenyl HP HIC column® (GE healthcare) (Product code: 17-1351-01).

## **Buffers**

### **SDS-PAGE running buffer:**

0.025M Tris, 0.192M Glycine, 0.1% SDS.

### **3×SDS-PAGE Sample buffer:**

187 mM Tris HCL, 6% SDS, 30% Glycerol, 0.03% Bromophenol blue, pH6.8.

### **0.25% Coomassie Blue:**

In 95% ethanol

### **10% Ammonium persulphate solution**

### **10% Sodium dodecyl sulphate solution**

### **IMAC A:**

25mM Tris, 500mM NaCl, 10mM Imidazole, pH 11.5.

### **IMAC B:**

25mM Tris, 500mM NaCl, 500mM Imidazole, pH 11.5.

### **HIC starting buffer:**

25mM Tris, 1M NaCl, pH 11

### **HIC elution buffer:**

25mM Tris, pH11.

### **Cell lysis:**

20mM MOPS, 500mM NaCl, 100mM Na<sub>2</sub> CO<sub>3</sub>, 10mM Imidazole, 0.5% Triton X-100, 0.5% Tween-20, pH 11.5.

### **Sonication buffer:**

25 mM Tris HCL, 10 mM imidazole, 500 mM NaCl, pH 7.5

**Storage buffer:**

25mM Tris, 500mM NaCl, 2% glycerol, pH 10.5.

**Phosphate buffered saline:**

As per manufacturer (Fisher Scientific)

**Phosphate buffered saline, 0.1% Tween-20**

**Acetate buffers pH 4-5.5.**

**Sorrensens buffers pH 5.6**

5 ml of 133 mM  $\text{Na}_2\text{HPO}_4$  solution in 95 ml of 133 mM  $\text{KH}_2\text{PO}_4$

**Staining solution 1:**

50% ethanol, 10% acetic acid.

**Staining solution 2:**

5% ethanol, 7.5% acetic acid, 0.25% coomassie blue in 95% ethanol.

**Western transfer buffer:**

48mM Tris, 39mM Glycine, 20% methanol, 1.3mM SDS, pH8.2.

**UPAS buffer:**

50 mM NaAcO, 150 mM NaCl, 6 M urea, 1 mM EDTA, 0.01% triton X-100, pH 5.5.

**250mM NaCl solution**

**1M NaCl solution**

### **Software**

- Unicorn 5.11
- GraphPad Prism 5
- Nanodrop 2000
- Scrubber 2
- ExPASy Prot Param tool
- Rare Codon Calculator (RaCC)
- DNA/RNA sequence translator ([www.fr33.com](http://www.fr33.com))
- Clustal W2
- SPR autolink 1-1-9
- SKANIT software 3.0

### **Equipment**

- ÄKTA purifier 10
- Nanodrop spectrophotometer 2000c
- Thermo Scientific Multiscan FC Microplate Photometer
- Reichert SPR SR7000DC
- Veriti 96 well thermocycler
- BioRad transblot SD semidry transfer cell

## Methods

### *Small-scale protein expression:*

Using a glycerol stock containing the 3-OST-1 expression plasmid provided by Prof. Jian Liu, a scraping was used to inoculate 5 ml of LB media with 50 µg/ml of kanamycin. The inoculated culture was then grown at 37°C overnight, shaking at 250 rpm. The 5ml culture was used to inoculate 100 ml of LB media supplemented with 50 µg/ml of kanamycin using a 1:1000 inoculation ratio. The culture was grown at 37°C 250 rpm until the optimal absorbance was reached. Upon reaching OD<sub>600</sub> of 0.6 the culture was incubated at 22°C for 30 minutes then expression was induced with the addition of 0.2 mM IPTG. The culture was then incubated for a further 18 hours.

The 100 ml culture was centrifuged at 10000 rpm for 20 min and the supernatant was discarded. The culture pellet was re-suspended in 12.5 ml of sonication buffer and the suspension was sonicated four times (See Appendix B for sonication conditions Table 2). The suspension was centrifuged at 10000 rpm for 1 hour. The supernatant was then sterile filtered using a 0.45 µm filter.

### *IMAC gravity flow purification:*

The supernatant was then purified using gravity flow immobilized metal ion affinity chromatography (IMAC). Using a 10 ml column the membrane was hydrated using 1 ml of water after which 800 µl of Ni-sepharose was added and allowed to compact. The resin was washed with 10 ml of sonication buffer. The supernatant was passed through the column 1 ml at a time and the column was washed with a further 10 ml of sonication buffer. The bound protein was then eluted in 3 ml of sonication buffer with 500 mM imidazole.

### Enzyme activity assay:

The catalytic activity of the purified enzyme was then determined. A reaction mixture was made for the test sample, the positive control (sample of known activity) and the negative control (denatured protein sample) (See Appendix B for the enzyme activity reaction mixture table 3). The reaction mixtures were incubated at 37°C for 1 hour in a water bath followed by boiling the sample for 1 min. The reaction mixture was then stored on ice for 2 min. The reaction was centrifuged at 14000 rpm for 1 min, the supernatant was discarded and the pellet was resuspended in 100 µl of UPAS buffer.

Diethylaminoethyl (DEAE) chromatography was used to purify the heparin prior to analysis by liquid scintillation counting. A DEAE column was prepared by washing the column three times with water followed by the addition of 200 µl of DEAE resin. The column once the resin was compacted was washed with 1 ml of a 250 mM NaCl solution a total of three times. The washes were repeated with a 1 M NaCl solution and UPAS buffer. The resuspended reaction mixture was then applied to the column which was subsequently washed three times with UPAS buffer and 250 mM NaCl. The heparin was eluted from the column using 1 ml of a 1 M NaCl solution.

Using the DEAE chromatography purified heparin the activity of the enzyme was assessed using liquid scintillation counting. To 2 ml of liquid scintillation fluid 100 µl of DEAE chromatography purified heparin was added and was mixed thoroughly. The sample was then analysed by a liquid scintillation counter.

### Generation of mutants:

Using 5 ml culture generated as previously described, plasmid DNA was isolated using a Qiagen miniprep kit using a modified protocol. The 5 ml culture was centrifuged at 4200 rpm for 10 min and the supernatant was discarded. The culture pellet was resuspended in 500 µl of P1 buffer and the suspension was incubated on ice for 10 min. To this 500 µl of P2 buffer was added and the suspension was continuously mixed for 3 min. The solution was returned to ice and 750 µl of N3 buffer was added, the

suspension was then incubated on ice for 10 min. After centrifuging the sample the procedure was carried out as per kit instructions.

Using oligonucleotide primers, mutants of 3-OST-1 were generated through site-directed mutagenesis (SDM). Oligonucleotide primers were designed to be complementary to the 3-OST-1 nucleotide sequence. This nucleotide sequence consisted of the catalytic domain with an N-terminal his-tag as in the PDB 1VKJ (Edavettal *et al*, 2004). Using optimized conditions (See Appendix B for SDM conditions and primer sequences Tables 4-7) the resulting SDM product was then treated with 1.5  $\mu$ l of *Dpn* I, the reaction was then incubated at 37°C for 4 hour.

The *E. coli* cell lines HB2151, BL21 Gold, RIPL and TG1 were prepared using a similar protocol as described by Chung *et al* (Chung and Miller 1993). Overnight 5 ml cultures were grown for each cell line in LB media at 37°C with shaking at 250 rpm. The following morning the cells were diluted in fresh media to give a 1:100 dilution in 50 ml. The cultures were grown at 37°C with shaking at 250 rpm until an OD<sub>600</sub> of 0.2-0.5 was reached. The cultures were split into two 50 ml tubes and were incubated on ice for 10 min. The culture was centrifuged at 3000 rpm at 4°C for 10 min, the supernatant was then discarded. The culture pellet was resuspended in TSS buffer at 10% the volume spun down from which aliquots were made and were stored at -80°C.

Previously prepared competent cells were aliquoted into 50  $\mu$ l and stored on ice for 30 min. To the appropriate aliquot 5  $\mu$ l of *Dpn*1 digested site-directed mutagenesis product was added, a positive control consisted of 1  $\mu$ l of template DNA in 50  $\mu$ l of the competent cell and the negative control consisted of a 50  $\mu$ l aliquot of the competent cells. The aliquots were incubated for a further 30 min before undergoing heat shock at 42°C for 90 sec. The aliquots were incubated on ice for two min before being transferred to 500  $\mu$ l of TB media and incubated at 37°C for 1 hour at 100 rpm. The aliquots were brought up to 1 ml with fresh media containing 10  $\mu$ g/ml of kanamycin and were incubated for a further hour. The concentration of kanamycin was adjusted to 50  $\mu$ g/ml and the aliquots were incubated for another hour before being centrifuged at 13000 rpm for 1 min. After discarding the supernatant the pellet was

resuspended in 200 µl of TB media and the suspension was plated on LB agar plates supplemented with 50 µg/ml of kanamycin and 1% glucose. The plates were incubated at 37°C overnight. The site-directed mutagenesis products were successfully transformed into the following *E. coli* competent cells and were stored at 4°C:

- E90Q into *E. coli* HB2151.
- E90H into *E. coli* HB2151.
- E90K into *E. coli* (DE3) BL21 Gold.
- E90R into *E. coli* TG1.

### Assessing the SDM transformed product:

Colony polymerase chain reaction (PCR) was used to assess the colonies generated from the *Dpn I* digested site-directed mutagenesis product transformations prior to sending the samples for sequencing. Colonies were selected and a scrapping was added to the colony PCR reaction mixture. Forward and reverse primers were used to amplify the appropriate sequence including the sequence encoding for the protein of interest (See Appendix B for the colony PCR reaction conditions Table 8).

Agarose gel electrophoresis was used to analyse the colony PCR product. A 1% agarose gel was made by adding 0.45 g of agarose to 45 ml of 1× TAE buffer. The solution was microwaved for 2 min. Before pouring the solution into the gel apparatus 2 µl of SYBR safe was added and the apparatus was set up then left to set for 30 min. To the appropriate well 10 µl of a 1kb DNA ladder and the low molecular weight ladder was added followed by the appropriate samples consisting 5 µl of the colony PCR product and 2 µl of type 2 gel loading buffer. The gel was run at 80 volts for 45 min. The resulting gel was then visualized under UV light.

Using the appropriate samples from the transformation reaction (chosen using the colony PCR results) 5ml overnight cultures were generated as described previously from which glycerol stocks were generated. Glycerol stocks were prepared by taking 750 µl of culture and adding it to 250 µl of glycerol. The stock were then stored at -80°C. As previously described plasmid DNA was isolated from the cultures generated and they were subsequently sent for gene sequencing.

*Large-scale protein expression of wild type and mutants:*

5 ml cultures were generated from glycerol stocks of the mutants and the wild type. Plasmid DNA was isolated as before which was then transformed into *E. coli* (DE3) RIPL cells as described previously. Using glycerol stocks of the newly transformed samples, 5ml cultures were inoculated using protein expression media supplemented with 1% glucose, 35 µg/ml streptomycin and 50 µg/ml kanamycin. The cultures were incubated at 37°C 300 rpm overnight.

Using 10 ml cultures generated, 900 ml of protein expression media (See materials section) was inoculated at 1:100 ratio. The media was supplemented with 4 ml of glycerol, 100 ml of phosphate buffer, 35 µg/ml streptomycin and 50 µg/ml kanamycin. The culture was grown at 37°C 200 rpm, upon reaching an OD<sub>600</sub> of 0.6 the culture temperature was reduced to 18°C and was incubated for a further hour and a half. Protein expression was induced with 0.2 mM IPTG and then was incubated for a further 18 hours.

The 1L culture was centrifuged at 10,000 rpm for 20 min after which the supernatant was discarded. The culture pellet was resuspended in 200 ml of cell lysis buffer and was sonicated in 50 ml aliquots as described previously. The pH was monitored and adjusted to 11.5 when necessary prior to incubating the cell lysate on ice for 1 hour. The cell lysate was sonicated as before a second time and was centrifuged at 10,000 rpm for 1 hour. The suspension was then filtered through a 0.45 µm filter.

*IMAC purification:*

The soluble fraction generated was applied to a 5 ml His-trap Fast flow column at 2 ml/min after equilibrating the column with IMAC A. The column was then washed with 14 column volumes of IMAC A and the bound protein was then eluted with 100% IMAC B. The flow through from the first purification was applied to the column again as before and the bound protein was eluted as before. The 100% elutions were combined and were concentrated using an aminicon filter. The concentrated protein was diluted

in 100 ml of IMAC A then was purified as described above. The sample was then concentrated and buffer exchanged into 30 ml of HIC starting buffer.

*HIC purification:*

The IMAC purified sample was applied to a 1 ml Phenyl high performance hydrophobic interaction chromatography (HIC) column that was previously equilibrated with HIC starting buffer. The protein sample was loaded onto the column at 1ml/min. The column was then washed with at least 10 column volumes of HIC starting buffer and the bound protein was eluted in 100% HIC elution buffer. The eluted fraction was then buffer exchanged into storage buffer and was stored at 1 mg/ml. The sample was stored at 4°C.

*SDS-PAGE of protein purification:*

The protein purifications were analysed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) to assess the quality of the purifications. The gel was made by first pouring the resolving layer and allowing it to polymerize while covered by a layer of 70% ethanol. Once polymerized the stacking layer was poured and the well comb was placed into the gel apparatus and the layer was allowed to polymerize (For gel composition see Appendix B Table 9).

The samples were prepared by adding 10 µl of protein sample along with 5 µl of sample buffer. The gel apparatus was set up and the gel running buffer was added. To the appropriate well 6.5 µl of protein ladder was added followed by adding the samples appropriate wells. The gel was run at 100 volts for 15 min followed by 180 volts for 45 mins. The gel was carefully removed upon completion of the run and was then placed in staining solution 1. The gel was microwaved for 1 min and was placed on a rocker for 30 min. The staining solution 1 was replaced with staining solution 2 and the gel was microwaved for 1 min. The gel was placed on a rocker overnight to allow the gel to develop fully. The gel was then placed into Millipore water to remove

excess stain with the process being repeated until the excess stain was completely removed.

### Western blotting of protein purification:

The protein purifications were analysed by Western blotting to verify the presence of the target protein. A SDS-PAGE gel was run, as previously described, the gel was then placed into Western transfer buffer. The blotting paper was also soaked in Western transfer buffer prior to the activation of the transfer membrane. The transfer membrane, Immobilon P PVDF, was activated in methanol which was washed with water prior to being transferred to Western transfer buffer. Using a BioRad Trans-Blot SD semi-dry electrophoretic transfer cell the western blot was assembled. The transfer was run at 40V for 40 min. The transfer membrane was transferred to 2% Skimmed milk in PBS and was incubated at room temperature gently rocking for 1 hour. The membrane was washed three times at 10 min intervals using PBST. The primary antibody consisted of a murine anti-penta-His monoclonal Ab, the antibody was added to the membrane at a dilution of 1:2500 in 10 ml of blocking buffer. The antibody was incubated at room temperature for 1 hour with gentle rocking. The membrane was washed as before and the secondary antibody was added. The secondary antibody consisted of a goat-anti-mouse alkaline phosphatase conjugate. The secondary antibody was added to the membrane at a dilution of 1:2500 in 10 ml of blocking buffer. The antibody was incubated at room temperature for 1 hour with gentle rocking. The membrane was washed as before and the membrane was developed using 5-bromo-4-chloro-3'-indolyphosphate p-toluidine/nitro blue tetrazolium (BCIP/NBT) solution which was incubated at room temperature in the dark for 15 min then the membrane was washed in water.

### Indirect ELISA of wild type and mutant binding to heparin:

Using a 96 well polystyrene plate (Sarstedt) each well was coated with poly-L-lysine at concentration of 3 µg/ml and a final volume of 100 µl per well. The plate was incubated at room temperature for 1 hour to allow passive adsorption of the poly-L-

lysine to the plate. The plate was washed with PBST three times and was patted dry. To all wells except the controls (PBS only), heparin was added at a concentration of 50 µg/ml and a final volume of 100 µl per well. The plate was incubated overnight at room temperature. The plate was washed as before and each well was coated with 100 µl of 2% BSA in PBS as a blocking buffer. The plate was incubated at 37°C for 1 hour. The plate was washed as before and the protein of interest was then added to the plate using a 1:2 dilution across the plate beginning with a concentration of 20 µg/ml and a final volume of 100 µl per well in blocking buffer. The plate was incubated at 37°C for 1 hour. The plate was washed as before and an anti-penta-His horseradish peroxidase (HRP) conjugated antibody was added to each well at a 1:10000 dilution in blocking buffer at a final volume of 100 µl per well. The plate was incubated at 37°C for 1 hour. The plate was washed again as before and the ELISA 3,3',5,5'-tetramethyl benzidine (TMB) substrate was added at a final volume of 100 µl per well. The plate was then incubated at room temperature until colour developed in the wells. The reaction was stopped with the addition 0.2 M H<sub>2</sub>SO<sub>4</sub> at a final volume of 100 µl per well. Using a Thermo Scientific Multiscan FC Microplate Photometer the absorbance of the wells was measured at 450 nm. Graphpad Prism 5 was used to plot the data.

*Indirect ELISA of heparin-biotin conjugated sample:*

Using 500 µl of a 10 mg/ml stock of heparin-biotin conjugate (Sigma) the sample was applied to a spin column with a 10 K MWCO. The heparin-biotin conjugate was centrifuged at 4000 g for two 10 min intervals. The first and second flow through along with the concentrated fraction was collected separately. The samples consisting of the first and second flow through from above and the concentrated conjugate sample were then analysed. Analysis was carried out by indirect ELISA to ascertain the presence of free biotin in the heparin-biotin conjugated sample. Using a pre-coated streptavidin and pre-blocked 96 well plate a concentration of 50 µg/ml of each test sample was added to the designated wells at a final volume of 100 µl per well, with a blank well acting as a control. The plate was then incubated at room temperature for 1 hour. The plate was then washed three times with PBST and patted dry. A streptavidin-HRP conjugated antibody was added at a 1:10000 dilution in PBS at a final volume of 100 µl per well. The plate was then incubated at room temperature for 1 hour. The

plate was washed again as before and the ELISA TMB substrate was added at a final volume of 100  $\mu$ l per well. The plate was then incubated room temperature until colour developed in the wells. The reaction was stopped with the addition 0.2 M  $\text{H}_2\text{SO}_4$  at a final volume of 100  $\mu$ l per well. Using a Thermo Scientific Multiscan FC Microplate Photometer the absorbance of the wells was measured at 450 nm. Graphpad Prism 5 was used to plot the data.

### *SPR analysis of wild type and mutant binding to Arixtra:*

Kinetic analysis of the wild type and mutant interaction with Arixtra was carried out using a Reichert SR7500DC SPR instrument. The protein of interest was immobilized via amine coupling to a carboxymethyl-dextran coated chip. Activation of the chip was carried out using 40 mg of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC) and 10 mg of N-hydroxysuccinimide (NHS). Immobilization of the protein of interest was done at 20  $\mu$ g/ml of the protein of interest in 10 mM sodium acetate buffer pH 5. To block any remaining active sites 1M ethanolamine pH 8.5 stock was used. The running buffer in this experiment consisted of PBST. Using each preparation 400  $\mu$ l was injected over the chip as follows. A flow rate of 20  $\mu$ l/min was set. Then the EDC/NHS mixture was injected over the chip for 7 min to activate the chip. A 10 sec dissociation was set followed by four 500  $\mu$ l washes. The protein of interest was injected over the chip (left side only) for 10 min in the acetate buffer to allow pre-concentration of the protein of interest to the chip surface. A 1 min dissociation was set followed by a 5 min wait. Then the flow rate was then returned to both flow channels and a 500  $\mu$ l wash was performed. To block any remaining active sites on the chip, ethanolamine was injected over the chip for 10 min. This was followed by 10 sec dissociation and two 500  $\mu$ l wash steps. The system was then allowed to equilibrate.

The kinetics experiment was carried out by first preparing dilutions of Arixtra in running buffer with the running buffer also acting as the buffer blank. The dilutions used were 62.5  $\mu$ M, 93.5  $\mu$ M, 125  $\mu$ M, 188  $\mu$ M, 250  $\mu$ M, 375  $\mu$ M, 500  $\mu$ M and 750  $\mu$ M of Arixtra. The dilutions were injected over the chip preceding and followed by buffer blanks. The flow rate was set at 5  $\mu$ l followed by a 30 sec wait. The blanks and the

samples were injected over the chip for 12 min followed by 10 min dissociation in order from the lowest concentration to the highest sequentially as described above. Before each subsequent injection three 500  $\mu$ l washes were carried out. The data generated was then analysed using the analysis software scrubber 2 and BIAevaluation software 3.2.

# Chapter 3: Results and Discussion

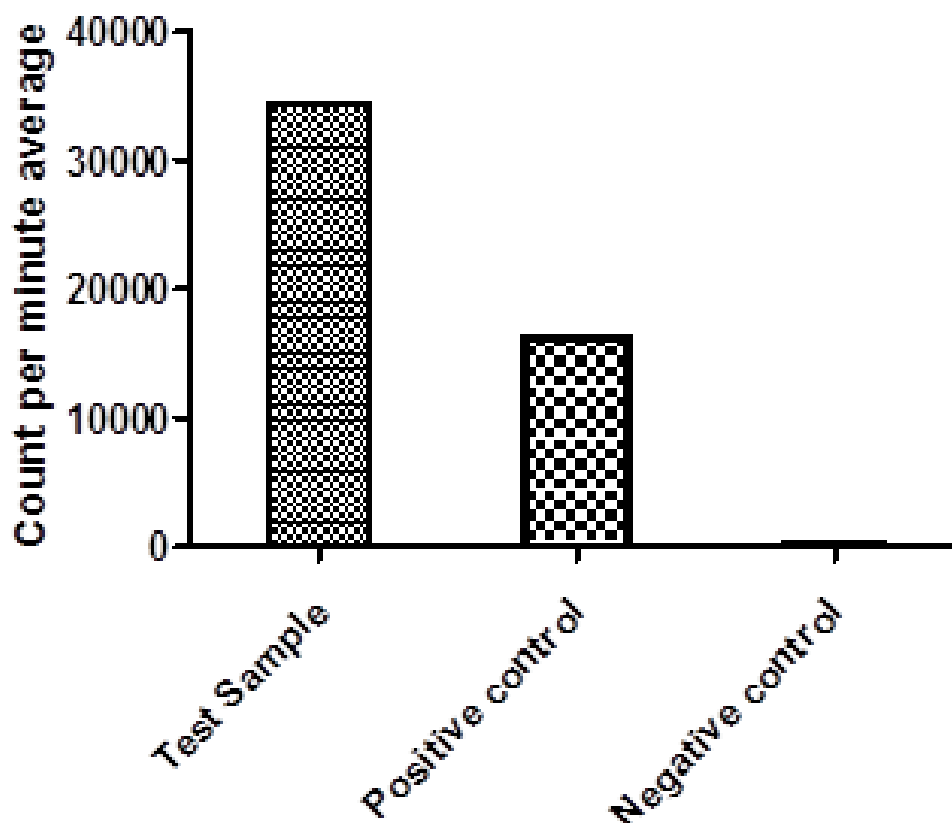
## **Results and Discussion**

The aim of this project was to characterize the binding of 3-OST-1 mutants to heparin and a heparin analogue (Fondaparinux, Arixtra). In characterizing the effect on binding of mutations to the catalytic base, the importance of charge-charge interactions was to be ascertained. Specifically, the effect of the introduction of a positive charge, and variations in side chain length at position 90, was studied. The purpose of these mutations was to investigate the possibility of enhancing affinity for the enzymatic product relative to the wild type enzyme through the interaction of the 3-O-sulphate with a positively charged residue at position 90. Work began with characterizing the enzymatic activity of the 3-OST-1 construct used in this project.

### **Analysing enzymatic activity:**

Expression of the 3-OST-1 wild type enzyme was carried out using the protocol developed by Edavettal et al. After the cells were pelleted, the soluble protein was isolated via sonication and purified by gravity flow IMAC (see appropriate section in Materials and Methods) (Edavettal, Lee et al. 2004).

Using the purified protein, an assay was carried out to determine the relative activity of the protein preparation. An enzymatic reaction was carried out by the incubation of the protein with heparin and  $1 \times 10^7$  cpm [ $^{35}\text{S}$ ] 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to catalyse the 3-O-sulphation of heparin. After the incubation period the heparin was then purified by Diethylaminoethyl (DEAE) chromatography. The relative degree of 3-O-sulphation was measured by liquid scintillation counting allowing the activity of the protein preparation to be inferred. The results were measured as a count per minute average (CPMA) which is proportional to the amount of [ $^{35}\text{S}$ ] O<sub>3</sub> transferred to heparan sulphate by the enzyme. A protein of known activity acted as the positive control and is compared to a test sample. In the case of this experiment a boiled protein acted as the negative control. The CMPA results obtained were compared to assess the activity of the protein sample (See Figure 10).



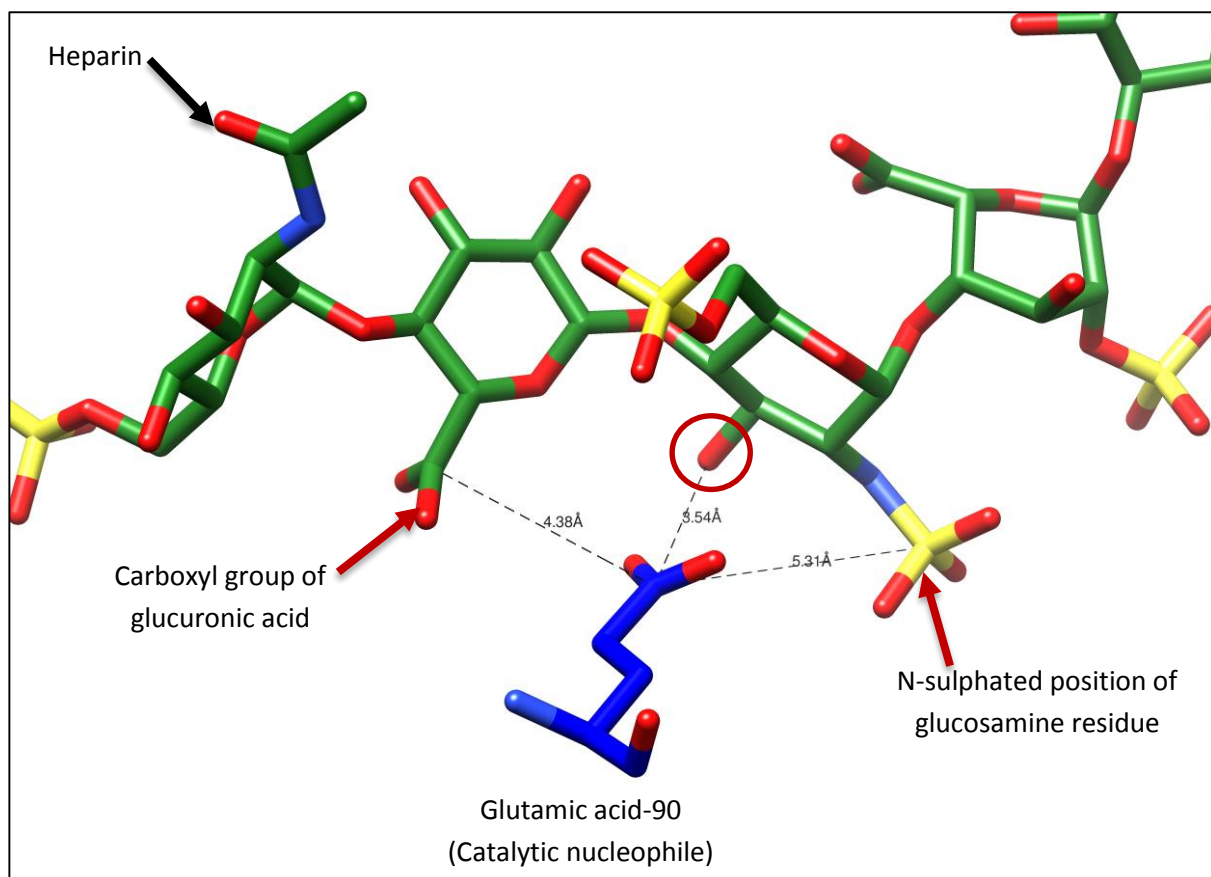
**Figure 10:** Count per minute average (CPMA) results of the liquid scintillation assay: The CPMA results here are proportional to the amount of [ $^{35}\text{S}$ ]  $\text{O}_3$  transferred to heparan sulphate by a sulphotransferases, in this case the wild type 3-OST-1. The test sample consisted of a newly expressed wild type sample, the positive control consisted of a sample (wild type) of verified activity and the negative control consisted of an inactive wild type expression. The CPMA results indicate an active protein expression with the comparison of the test sample to the controls as the [ $^{35}\text{S}$ ]  $\text{O}_3$  transferred to substrate is detected in both the test and positive control samples based on the CPMA results.

The results of the enzyme activity assay show the activity of the test expression relative to the positive and negative controls and by comparison the results indicate the test expression is active. Upon confirming the activity of the protein expression, the corresponding nucleotide sequence was used in the generation of single point mutants of 3-OST-1.

### **Generation of mutants:**

The selection of mutants was based on a number of structural features of 3-OST-1. Prior, research has shown glutamic acid-90 acts as the catalytic nucleophile in 3-OST-1. The catalytic nucleophile is located along a large cleft that acts as the substrate recognition site (Edavettal, Lee et al. 2004).

The substrate recognition site contains a large positively charged region. The catalytic nucleophile is located in negatively charged patch that flanks this region. Additionally, the PAPS co-factor binding site is located in a channel that is embedded within the protein and upon binding it presents the sulphate group to of Glu-90. Therefore the focus in designing the 3-OST-1 mutants was centred around this region and in residues that may interact with the acceptor sugar (See Figure 11).



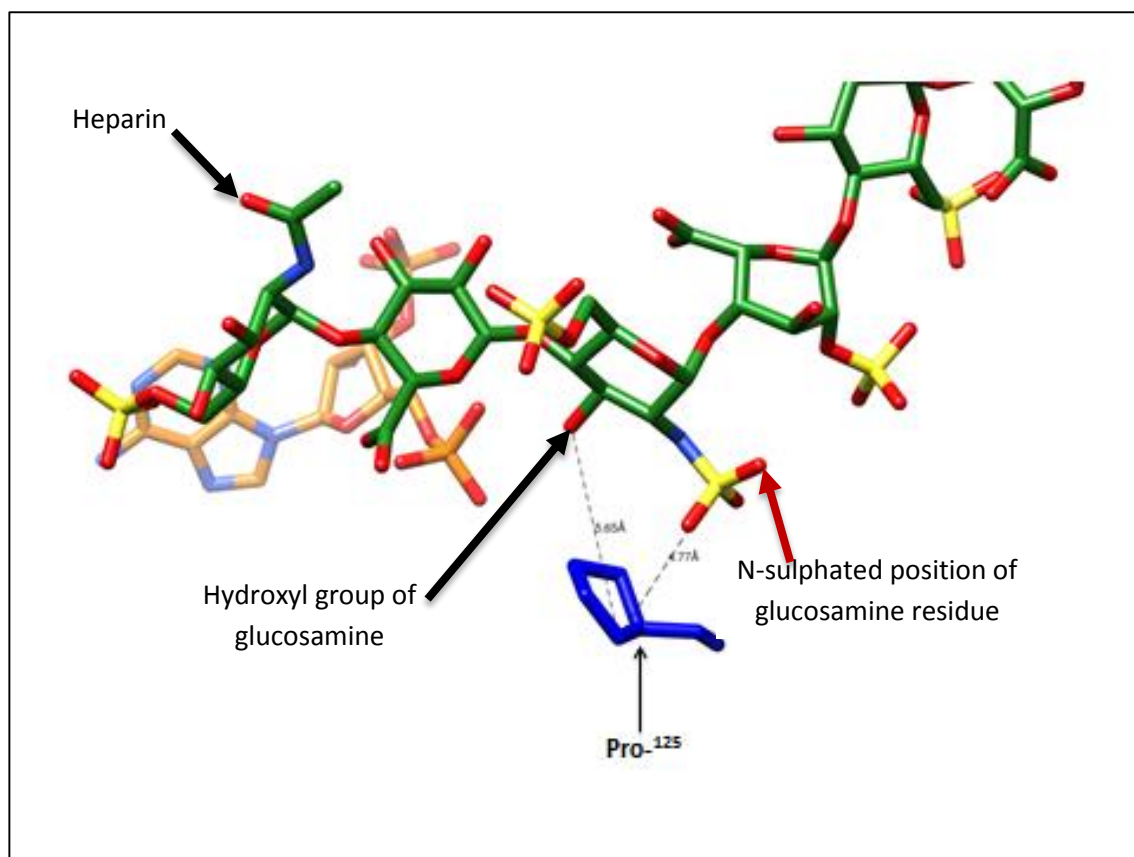
**Figure 11:** Position of Glu-90 (catalytic nucleophile) (in 3-OST-1) in relation to the N-sulphated and 3-OH position of a glucosamine and the carboxyl group of the glucuronic acid residue of the substrate (heparin). Highlighted (red arrows) are the negatively charged groups of the glucuronic acid and the glucosamine residue that may interact with the catalytic nucleophile. Circled in red is the position that is sulphated by 3-OST-1 and the dashed lines indicate the distance between Glu-90 and the highlighted positions in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).

When selecting residues for mutation a number of positions were considered when studying the crystal structure. The focus was on the residues that may interact with the glucosamine of the acceptor sugar that is essential for antithrombin. Crystal structure and mutational analysis of 3-OST-1 have shown that positively charged residues play important roles in heparin binding and interactions with sulpho groups (Edavettal, Lee et al. 2004; Moon, Xu et al. 2012).

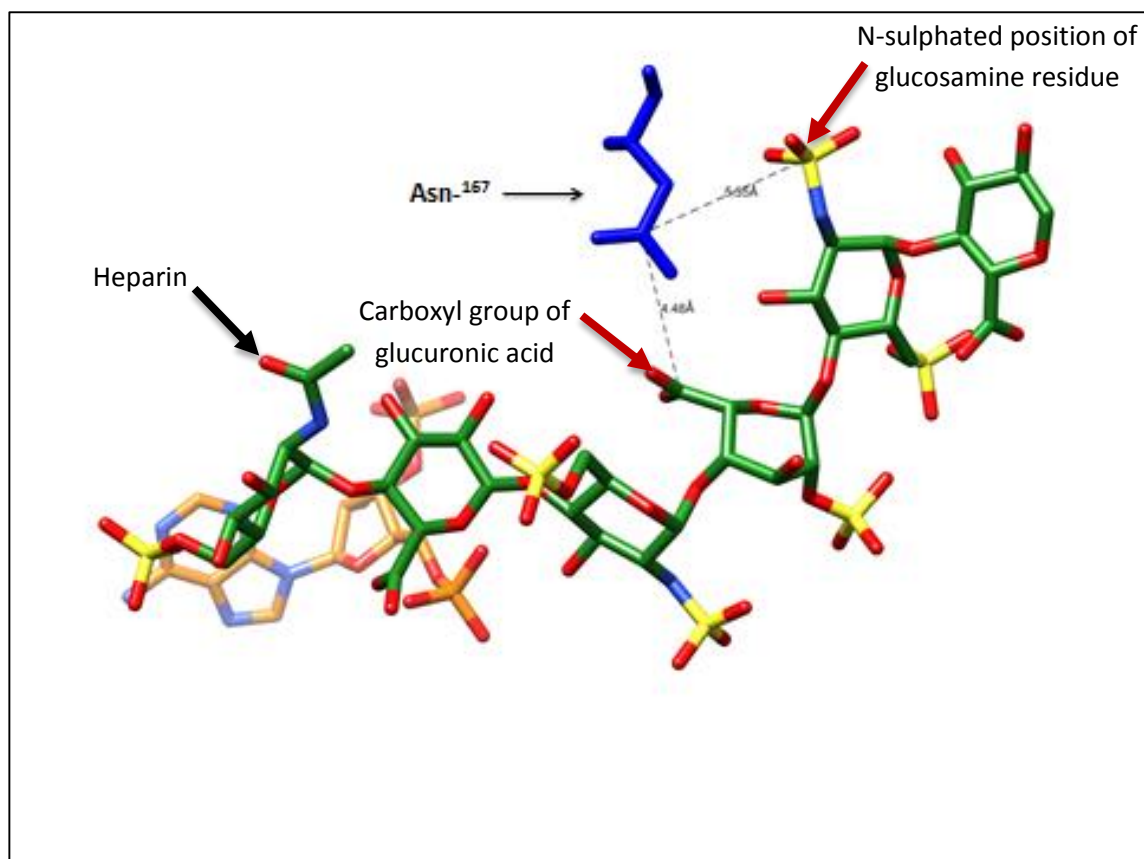
Pro-125 is located adjacent to position 90 and is close to the acceptor sugar in heparin, one might therefore consider this residue as a possible mutation site. However this position was not considered due to the structural role of proline residues (Edavettal, Lee et al. 2004) (See figure 12)

Flanking position 90 on the other side is Lys-123 which has been shown to be important for protein activity. Mutation of this residue to an alanine has been shown to result in a 51% drop in protein activity (Edavettal, Lee et al. 2004). This would suggest that removal of the positive charge at this position may be responsible for the drop in activity and therefore this was the reason that this position was not considered (See Figure 13).

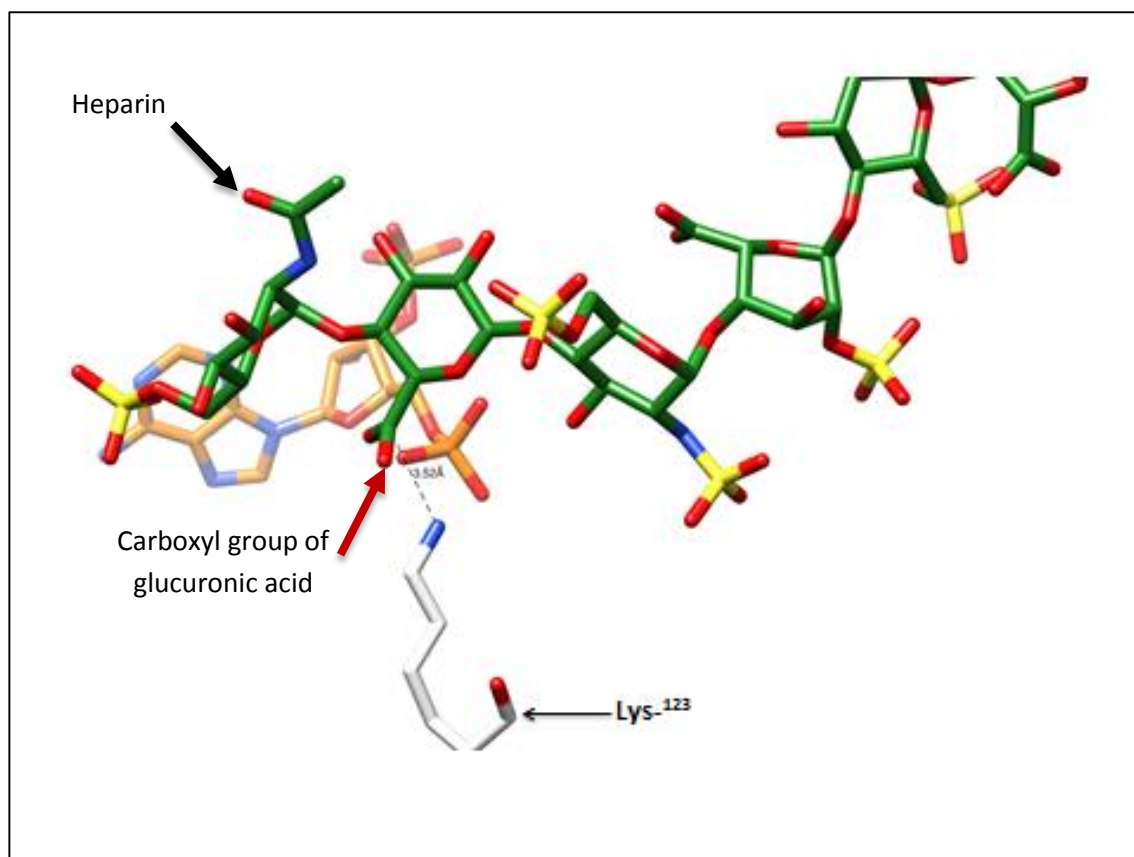
Asn-167 is located near the carboxyl position (4.45Å) of the iduronic acid adjacent the acceptor sugar and structurally one could suggest the possibility that it may play a role in substrate binding. This site however was not selected due to the distance from the 3-OH position of the acceptor sugar (See Figure 14) (Edavettal, Lee et al. 2004).



**Figure 12:** Pro-125 (in 3-OST-1) relative to a hydroxyl group and the N-sulphated position of a glucosamine residue in the substrate (heparin). Highlighted (red arrows) is a negatively charged group in the glucosamine residue that may interact with the position-125. The dashed lines indicate the distance between Pro-125 and the highlighted positions in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).



**Figure 13:** Asn-167 (in 3-OST-1) relative to an N-sulphated position of the glucosamine and the carboxyl group of the adjacent iduronic acid residue in the substrate (heparin). Highlighted (red arrows) are the negatively charged groups of the iduronic acid and the glucosamine residue that may interact with the position-167. The dashed lines indicate the distance between Asn-167 and the highlighted positions in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).



**Figure 14:** Lys-123 (in 3-OST-1) relative to the carboxyl group of glucuronic acid residue in the substrate (heparin). Highlighted (red arrows) is a negatively charged group of the glucuronic acid that may interact with the position-123. The dashed lines indicate the distance between Lys-123 and the highlighted position in the heparin molecule (PDB:3UAN) (Moon, Xu et al. 2012).

Research has shown that substituting the carboxyl group side chain of Glu-90 for that of a glutamine abolishes catalytic activity yet does not prevent substrate binding (Muñoz, Xu et al. 2006). Therefore this position became the primary target in the generation of mutants to bind to 3-OST-1 modified heparin

The positively charged amino acids in the substrate binding site are of particular importance to binding in relation to their interactions with the sulpho and carboxyl groups in heparin. This information narrowed the range of amino acid substitution considered (Moon, Xu et al. 2012).

Studies have indicated that arginine and lysine residues play important roles in heparin binding and of the three positively charged amino acids, arginine contains the longest side chain and is the strongest base (Caldwell, Nadkarni et al. 1996). The three asymmetrical nitrogen atoms in the guanidinium group of arginine allow interactions in three possible directions compared to one in lysine residues (Borders, Broadwater et al. 1994; Donald, Kulp et al. 2011).

Interestingly in cytosolic sulphotransferases the catalytic nucleophiles are thought to be histidine residues have been found to be predominantly conserved amongst other cytosolic sulphotransferases (Weinshilboum, Otterness et al. 1997). This is in contrast to glutamic acid of Golgi sulphotransferases like 3-OST-1. In 3-OST-1 His<sup>92</sup> plays an important role in catalytic activity and is thought to be important for substrate recognition (Edavettal, Lee et al. 2004). Taken together the three positively charged amino acids are known to play important roles in 3-OST-1 binding to heparin (Edavettal, Lee et al. 2004).

The E90Q mutant of 3-OST-1 has been shown to bind to heparan sulphate with similar affinity to wild type 3-OST-1 (Edavettal, Lee et al. 2004). This suggests that the presence of a glutamine at position 90 does not negatively affect the binding of the substrate. Therefore based on all the considerations the mutations chosen were E90R, E90K, E90Q and E90H. The mutants were generated using SDM and mutant

primers were designed to be complimentary to the parental DNA sequence and were used in the optimized SDM reaction.

A number of components of this reaction needed to be optimized in order to generate the desired mutants. In the reactions the primer concentrations were kept constant for the majority of the test experiments to aid primer annealing. Different buffer concentrations for both the *Pfx* and phusion kits were investigated in addition to varying the dNTP and plasmid DNA concentrations. As a whole these alterations were employed to find the optimal conditions to generate the desired mutants.

To assess the validity of these alterations to the protocol, with each variable changed in the protocol the corresponding SDM reaction was *Dpn I* digested and transformed into chemically competent cells. The success or failure of these alterations to the protocol was determined by the presence of transformation colonies and analysis of the subsequent DNA sequencing results. Typically the failure to generate transformation colonies indicated the success or failure of the alterations to the protocol.

In conjunction with varying the reaction mixture components, the thermocycling conditions were altered. Different denaturing, annealing and extension temperature and times were also investigated. Variations in the denaturation temperature beyond kit recommendations were examined in addition to the introductions of DMSO to the reaction mixture to aid in the denaturation of super-coiled DNA. Increasing the annealing temperature in conjunction with using the enhancer solution provided with the kit allowed an annealing temperature closer to the primer melting temperature to be used, while still maintaining the guidelines set out by the kit. Variations in the extension time and temperature were investigated to allow for sufficient time generate the desired mutant. The number of cycles used was also varied in addition to experimenting with using different DNA polymerases, namely *Pfx*, phusion and phi 29.

As with the variations to the SDM reaction mixture, typically the failure to generate transformation colonies indicated the success or failure of the alterations to the protocol. When transformations colonies were generated a continuing problem was the presence of parental DNA upon analysis of the sequencing results. With optimizations to the thermocycling and reaction mixture conditions the E90Q mutant was successfully generated (see Table 1). However generating the remaining mutants using the same conditions proved to be difficult. Due to the presence of parental DNA and the poor transformation efficiency the focus was turned to improving the *Dpn I* digestion and transformation protocols.

Therefore in addition to variations in the reaction mixture and thermocycling conditions the *Dpn I* digestion required optimization. Increasing the duration of digestion and the volume of the enzyme in some cases was required to increase the efficacy of the removal of the parental DNA. This led to a reduction in the occurrence of false positives in regards to the generation of transformation colonies.

Poor transformation efficiency even in the positive controls led the focus of the investigations to improving the transformations efficiency. Alterations to the transformation protocol were undertaken in conjunction with increased *Dpn I* digestion. The uses of different competent cells were also investigated. Increasing the transformation recovery time in the presence of the appropriate antibiotic led to an increase in the number of transformation colonies. The colonies generated were further examined prior to gene sequencing using colony PCR so as to select the appropriate colonies for DNA sequencing.

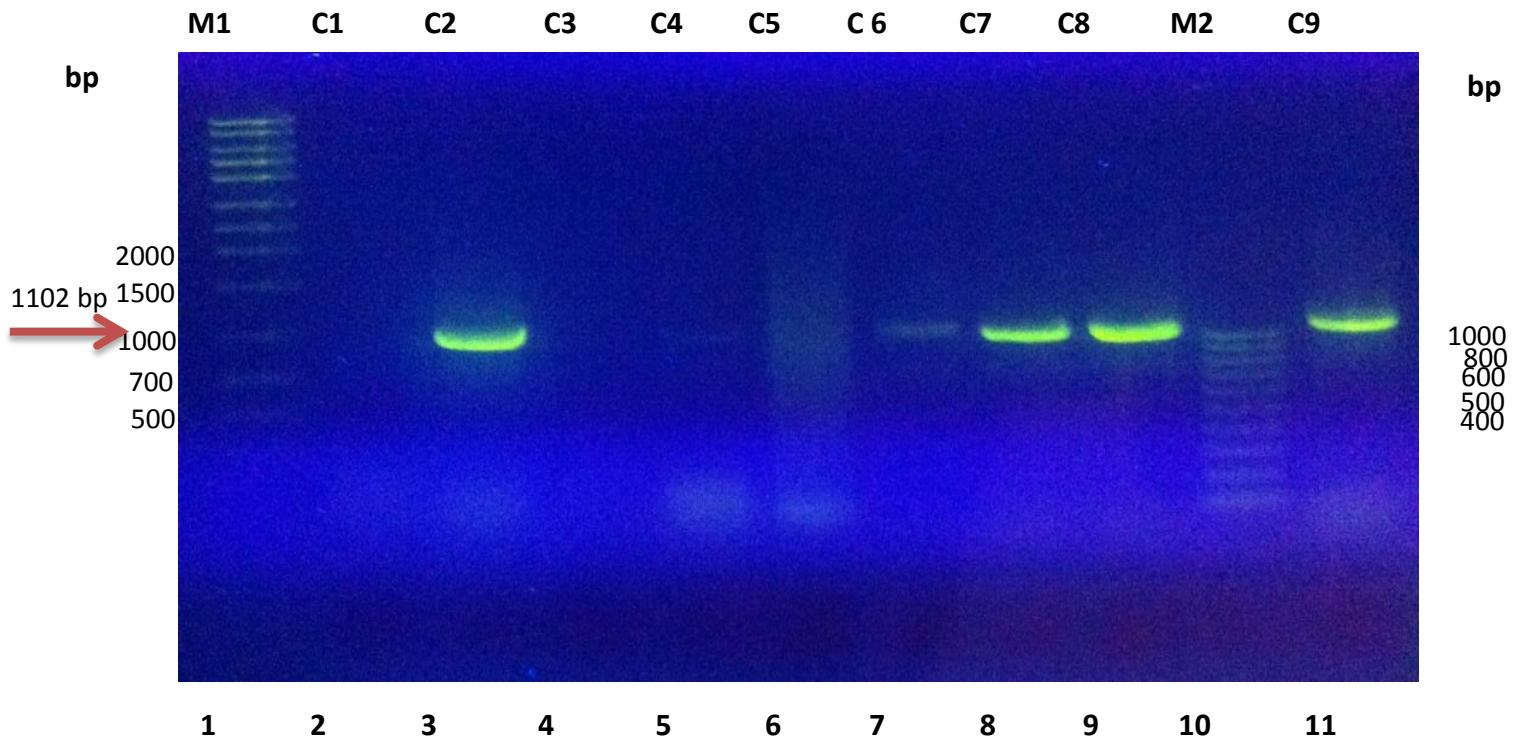
### **Agarose gel of Colony PCR amplification of site-directed mutagenesis transformants:**

Initially upon obtaining transformation colonies, the plasmid DNA was isolated from the samples being tested, which were then sent for sequencing. The sequencing

results obtained were often problematic in that the sequencing either failed, possibly due to poor DNA purity, or template DNA was sequenced in the samples sent. A combination of increasing transformation efficiency and improving the plasmid DNA purity was employed using modified transformation and plasmid DNA isolation protocols.

Colony PCR allowed for the testing for the presence of the target nucleotide sequence in the sample being sent for sequencing. In performing colony PCR the T 7 primer sets under the reaction conditions employed were designed to amplify the nucleotide sequence. The resulting amplification would be around 1102 base pairs in length.

The sample generated was then analysed by agarose gel electrophoresis to ascertain the size of the colony PCR amplified product. The colonies from which the reaction was performed were used to inoculate agar plats. The results of the colony PCR were therefore very useful choosing which samples were sent for gene sequencing (See Figure 15). However the presence of parental DNA remained a problem that was overcome in sending multiple samples for sequencing.

- Agarose gel of colony PCR samples:

**Figure 15:** Colony PCR samples (C) indicated from row 1-9, (1) M1 = Molecular weight ladder (DNA ladder exACTGene), row (2) – (9) = C1-8, row (10) M2= Low molecular weight ladder (Low Scale Ladder 100 bp) and row (11) = C9. Approximate size of the PCR amplified product is 1102 base pairs (bp).

**Sequencing results for site-directed mutagenesis:**

Samples were sent for sequencing using LGC genomics, the sequence data obtained was then analysed where the nucleotide sequence was converted to an amino acid sequence and the correct reading frame was identified. The sequence was then aligned with the parental DNA sequence to determine if the correct mutation was achieved.

Getting to the point of generating all four mutants required extensive optimizations to multiple protocols in consort to successfully produce the desired mutants. There were a number of obstacles that needed to be overcome to achieve this. The thermocycling conditions as well as the reaction mixture components of the SDM required optimization. Maintaining the primer concentration in excess of the template DNA concentration in addition to using the optimal DNA polymerase appears to be critical. Additionally increasing the annealing temperature in the presence of the enhancer buffer allowed an annealing temperature to be used closer to the calculated melting temperature of the primers, while still not exceeding the conditions suggested by the kit. Doing so appears to be of importance considering the conditions investigated.

Optimizing the plasmid DNA isolation protocol allowed for increased plasmid DNA purity. The poor sequencing results was often due to the presence of parental DNA or in other cases poor background signal in the sequencing results. This often resulted in failed sequencing. However upon optimization of the plasmid isolation procedure, the success of obtaining sequencing results greatly improved.

In the end, transformation efficiency was a persistent problem, optimization to the protocol in combination with the preceding experiments (see above) allowed for the successful generation of the 3-OST-1 mutants (See Figures 16, 17 and Table 1).

**Truncated nucleotide sequence**

```

1 ATGGGCACAG CATCCAATGG TTCCACACAG CAGCTGCCAC AGACCATCAT CATTGGGGTG
61 CGCAAGGGTG GTACCCGAGC CCTGCTAGAG ATGCTCAGCC TGCATCCTGA TGTTGCTGCA
121 GCTGAAAACG AGGTCCATTT CTTTGACTGG GAGGAGCATT ACAGCCAAGG CCTGGGCTGG
181 TACCTCACCC AGATGCCCTT CTCCTCCCCT CACCAGCTCA CCGTGGAGAA GACACCCGCC
241 TATTTCACTT CGCCCAAAGT GCCTGAGAGA ATCCACAGCA TGAACCCAC CATCCGCCTG
301 CTGCTTATCC TGAGGGACCC ATCAGAGCGC GTGCTGTCCG ACTACACCCA GGTGTTGTAC
361 AACCACCTTC AGAAGCACAA GCCCTATCCA CCCATTGAGG ACCTCCTAAT GCGGGACGGT
421 CGGCTGAACC TGGACTACAA GGCTCTCAAC CGCAGCCTGT ACCATGCACA CATGCTGAAC
481 TGGCTGCGTT TTTTCCCGTT GGGCCACATC CACATTGTGG ATGGCGACCG CCTCATCAGA
541 GACCCTTTCC CTGAGATCCA GAAGGTCGAA AGATTCTCTGA AGCTTTCTCC ACAGATCAAC
601 GCCTCGAACT TCTACTTTAA CAAAACCAAG GGCTTCTACT GCCTGCGGGA CAGTGGCAAG
661 GACCGCTGCT TACACGAGTC CAAAGGCCGG GCGCACCCCC AGGTGGATCC CAACTACTT
721 GATAAACTGC ACGAATACTT TCATGAGCCA AATAAGAAAT TTTTCAAGCT CGTGGGCAGA
781 ACATTCGACT GGCCTGGA

```

(A)

**Truncated amino acid sequence**

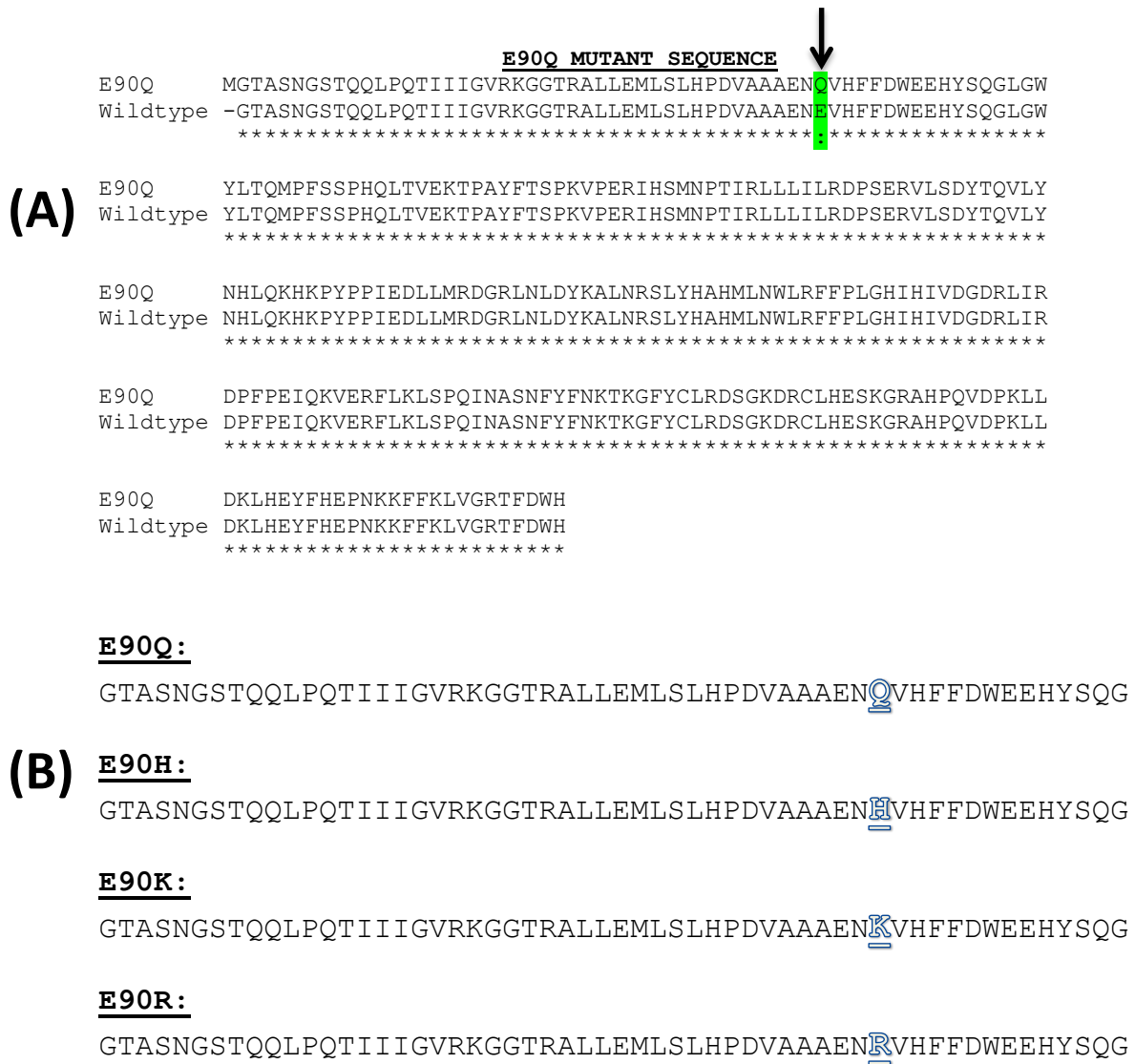
```

MGTASNGSTQQLPQTIIIGVRKGGTRALLEMLSLHPDVAAENEVHFFDWEEHYSQGLGWYLTQMPFSSPHQLTVEKTPAYFTSPKVPERIHSMNPTIRLLILRDPSE
RVLSDYTQVLYNHLQKHKPYPIEDLLMRDGRNLNDYKALNRSLYHAHMLNWLRFPLGHIHIVDGDRLIRDPFPEIQKVERFLKLSPOINASNIFYFNKTKGFYCLRDSG
KDRCLHESKGRAHPQVDPKLLDKLHEYFHEPNKKFFKLVGRTFDWHSTOP

```

(B)

**Figure 16:** Truncated nucleotide and amino acid sequence of the wild type 3-OST-1 used in the alignment of the sequencing result. The mutation site underlined in part A (nucleotide sequence) and in part B (amino acid sequence). The mutagenesis primer sites are both indicated in blue in part A and B.



**Figure 17:** Alignment of the gene sequencing obtained for the mutant E90Q with the template DNA in part A. In part B is an example of a portion of the nucleotide sequence of the mutants generated with the mutation highlighted in blue.

**Table 1:** Optimized thermocycling conditions and optimized site-directed mutagenesis reaction mixture for the generation of mutants:

Temperature		Time		Number of cycles		Stages	
E90K, E90Q, E90H	E90R	E90K, E90Q, E90H	E90R	E90K, E90Q, E90H	E90R	E90K, E90Q, E90H	E90R
94°C	95°C	2 min	5 min	1	1	1	1
94°C	95°C	30 sec	30 sec	20	30	2	2
55°C	55°C	30 sec	30 sec				
68°C	68°C	6 min	7 min				
68°C	68°C	10 min	10 min	1	1	3	3
4°C	4°C	∞	∞	∞	∞	4	4

Reagents *	Volume (μl)				Concentration
	E90R	E90K	E90H	E90Q	
Millicule water	26.0	26.3	32.1	29.3	
Stock dNTP	1	1	1	0	
2 mM dNTP	0	0	0	7.5	
1X Amplification buffer	0	5	5	5	
1X Enhancer buffer	0	5	5	0	
1X Phusion buffer	5	0	0	0	
5% DMSO	5	0	0	0	
Template DNA	∞	∞	∞	∞	
50 mM MgSO <sub>4</sub>	1	1	1	1	
Forward primer	1	1	1	1	
Reverse primer	1	1	1	1	
Phusion high fidelity DNA polymerase	0.4	0	0	0	
Platinum pfx DNA polymerase	0	0.4	0.4	0.4	

**Table 1:** Thermocycling conditions and optimized SDM reaction mixture for the generation of mutants, ∞ based on plasmid DNA concentrations

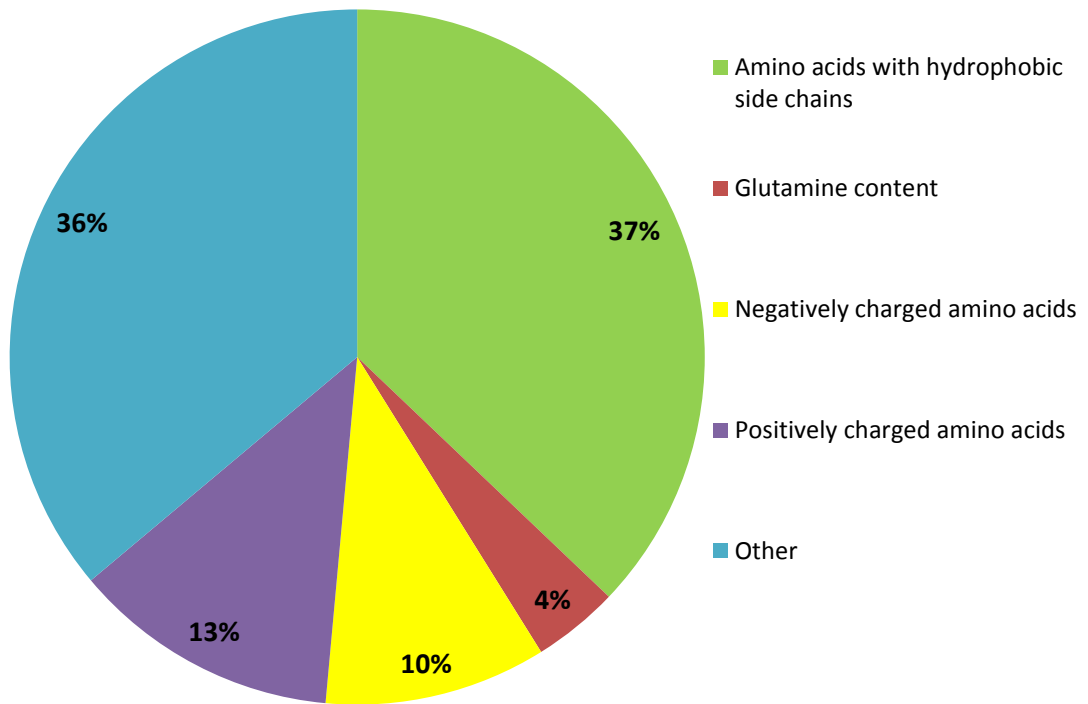
## **Purification of wild type and mutants:**

Protein expression and purification of the mutant and wild type following the protocol used by Edavettal et al. proved to be very problematic (Edavettal, Lee et al. 2004). A new protein expression, cell lysis and protein purification strategy needed to be developed in order to obtain sufficiently pure protein for further analysis. The strategy employed had its beginnings in investigating why 3-OST-1 would express insolubly. There are a number of potential problems that could cause a protein to express insolubly. Amino acid sequence analysis can show trends in amino acid composition that may contribute to insolubility. Christendat et al. have shown through sequence analysis of 424 non-membrane proteins from *Methanobacterium thermoautotrophicum* that there is a trend in the amino acid composition of insoluble proteins. Insoluble proteins tended to have lower glutamine contents (<4%), less negatively charged residues (<17%), a higher amount of hydrophobic residues with aromatic rings (>7.5%) and that insoluble proteins tend to have more hydrophobic stretches (Christendat, Yee et al. 2000). A sequence analysis of the 3-OST-1 construct used in this project revealed a similar trend with additional information indicated below:

- An aliphatic index of 79.79 indicates that the protein is thermostable (IKAI 1980).
- An instability index of 44.57 indicates that the protein is unstable (Guruprasad, Reddy et al. 1990).
- Grand average of hydropathicity (GRAVY) of -0.600, GRAVY indicates the solubility of proteins with negative values being hydrophilic (Kyte and Doolittle 1982).

Although the trend was similar one must acknowledge that protein expression was carried out in *E. coli* and the trend may not necessarily account for insolubility. Interestingly the relative abundance of hydrophobic residues would suggest however a possible explanation for the insolubility of the protein. Hydrophobic interactions are have been reported to be important in inclusion body formation and the relative abundance of these residues may play a role in the insolubility of the protein of interest (Upadhyay, Murmu et al. 2012) (See Figure 18).

### Amino acid composition

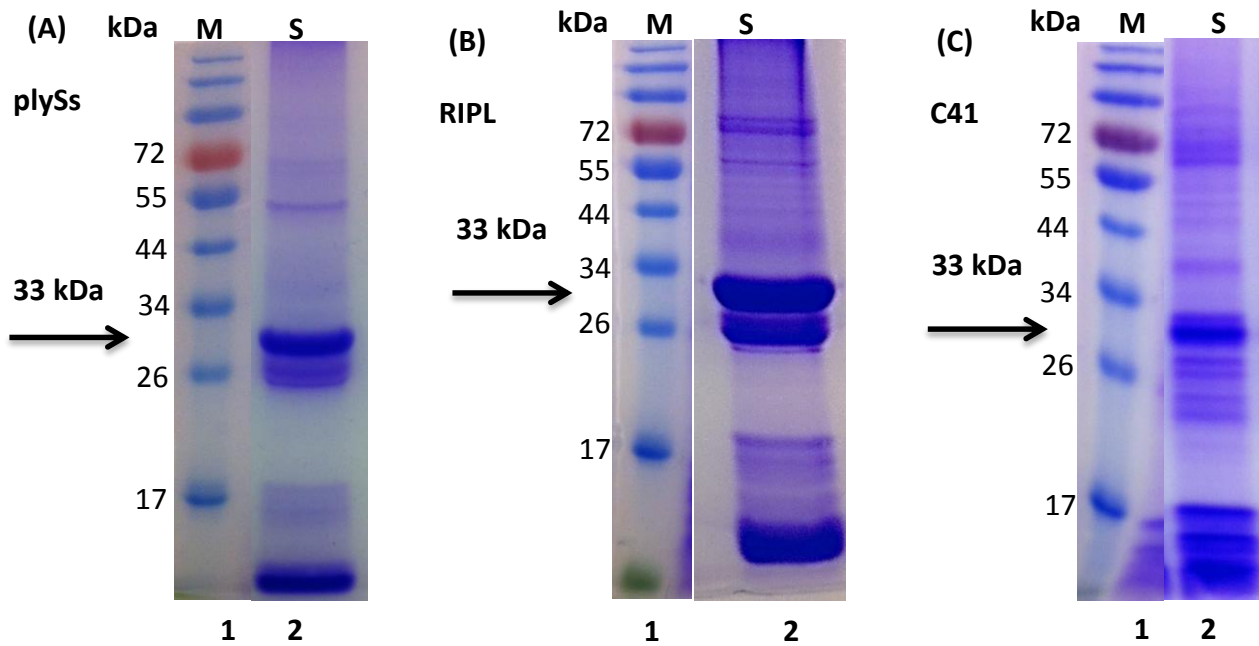


**Figure 18:** Pie chart of amino acid coding sequence of 3-OST-1 illustrating amino acids with hydrophobic side chains, positively charged side chains, negatively charged side chains and the glutamine content.

**Protein expression:**

In order to obtain sufficient soluble expression a number of approaches were undertaken. The inducer concentration along with the induction time, temperature and duration were varied in different cell lines. Different expression strains were investigated for their ability to improve soluble expression. *E. coli* (DE3) pLysS was tested as the strain has been shown to reduce basal expression. This strain produces small amounts of the T7 lysosyme which reduces the amount of basal expression that can be induced due to the presence of toxic gene constructs (Studier 1991). *E. coli* (DE3) C41 was investigated as a viable expression host as the strain has been used to successfully express toxic proteins, membrane proteins and or globular proteins (Miroux & Walker). The *E. coli* strain BL21 (DE3) RIPL was studied as the strain is used for the expression of proteins with rare codons (See Figure 19). The strain contains extra copies of tRNA's that aid in the expression of proteins containing rare codons, sequence analysis showed there are rare codons present in the protein used for this project.

The RIPL strain proved to be the most successful in increasing soluble expression, however it should be noted that it was only a contributing factor. Based on the assessment of the SDS-PAGE gels and the presence of rare codons in the nucleotide sequence of the protein of interest, *E. coli* (DE3) RIPL was chosen as the strain used in this study (see Figure 19 and 20). However it should be noted that low protein yield and purity were a significant issue. Therefore efforts were made to enhance soluble protein yields.



**Figure 19:** IMAC purification (3-OST-1) with 100% elution were the protein was expressed in different cell lines. (A) Expressed in *E. coli* (DE3) plySs (4.3 mg/ml in 200  $\mu$ l), (B) Expressed in *E. coli* (DE3) RIPL (5.1 mg/ml in 250  $\mu$ l) and (C) Expressed in *E. coli* (DE3) C41 (0.96 mg/ml in 500  $\mu$ l). (1) M= Molecular weight ladder, (2)= 100% elution. Estimated molecular weight of the protein of interest= 33 kDa.

ATG GGC ACA GCA TCC AAT GGT TCC ACA CAG CAG CTG CCA CAG ACC ATC ATC ATT GGG GTG  
 CGC AAG GGT GGT ACC CGA GCC CTG CTA GAG ATG CTC AGC CTG CAT CCT GAT GTT GCT GCA  
 GCT GAA AAC GAG GTC CAT TTC TTT GAC TGG GAG GAG CAT TAC AGC CAA GGC CTG GGC TGG  
 TAC CTC ACC CAG ATG CCC TTC TCC TCC CCT CAC CAG CTC ACC GTG GAG AAG ACA CCC GCC TAT  
 TTC ACT TCG CCC AAA GTG CCT GAG AGA ATC CAC AGC ATG AAC CCC ACC ATC CGC CTG CTG  
 CTT ATC CTG AGG GAC CCA TCA GAG CGC GTG CTG TCC GAC TAC ACC CAG GTG TTG TAC AAC  
 CAC CTT CAG AAG CAC AAG CCC TAT CCA CCC ATT GAG GAC CTC CTA ATG CGG GAC GGT CGG  
 CTG AAC CTG GAC TAC AAG GCT CTC AAC CGC AGC CTG TAC CAT GCA CAC ATG CTG AAC TGG  
 CTG CGT TTT TTC CCG TTG GGC CAC ATC CAC ATT GTG GAT GGC GAC CGC CTC ATC AGA GAC  
 CCT TTC CCT GAG ATC CAG AAG GTC GAA AGA TTC CTG AAG CTT TCT CCA CAG ATC AAC GCC  
 TCG AAC TTC TAC TTT AAC AAA ACC AAG GGC TTC TAC TGC CTG CGG GAC AGT GGC AAG GAC  
 CGC TGC TTA CAC GAG TCC AAA GGC CGG GCG CAC CCC CAG GTG GAT CCC AAA CTA CTT GAT  
 AAA CTG CAC GAA TAC TTT CAT GAG CCA AAT AAG AAA TTT TTC AAG CTC GTG GGC AGA ACA  
 TTC GAC TGG CAC TGA

**Figure 20:** Nucleotide sequence of the wild type 3-OST-1 used in this study showing rare codons linked to poor expression in *E. coli* (underlined in purple) (used Rare Codon Calculator to analyse sequence). The rare codons are as follows: Arginine (AGG, AGA, CGA) leucine (CTA), isoleucine (ATA), and proline (CCC).

To aid protein expression changes in the culture media were investigated to identify the optimal media for soluble protein expression. Different inoculation ratios were also investigated from 1:50, 1:100 to 1:1000. The antibiotic concentrations were varied in order to investigate the optimal concentration to maintain selective pressure on the expression vector (Smith and Bidochka 1998). In addition the culture aeration was varied in an effort to improve soluble expression. The shaking rates were varied from 150, 200, 250 and 300 rpm and the use of baffled culture flasks were tested. Oxidative stress has been shown to hinder culture growth which in turn will effect protein expression (Conter, Gangneux et al. 2001).

The alterations to the expression protocol were assessed through analysis of IMAC purified samples using SDS-PAGE. However a continuing problem was low protein yield and insufficient purity. To overcome this, methods to enhance soluble protein expression were investigated from colony selection, molecular chaperones to the use of chemical chaperones.

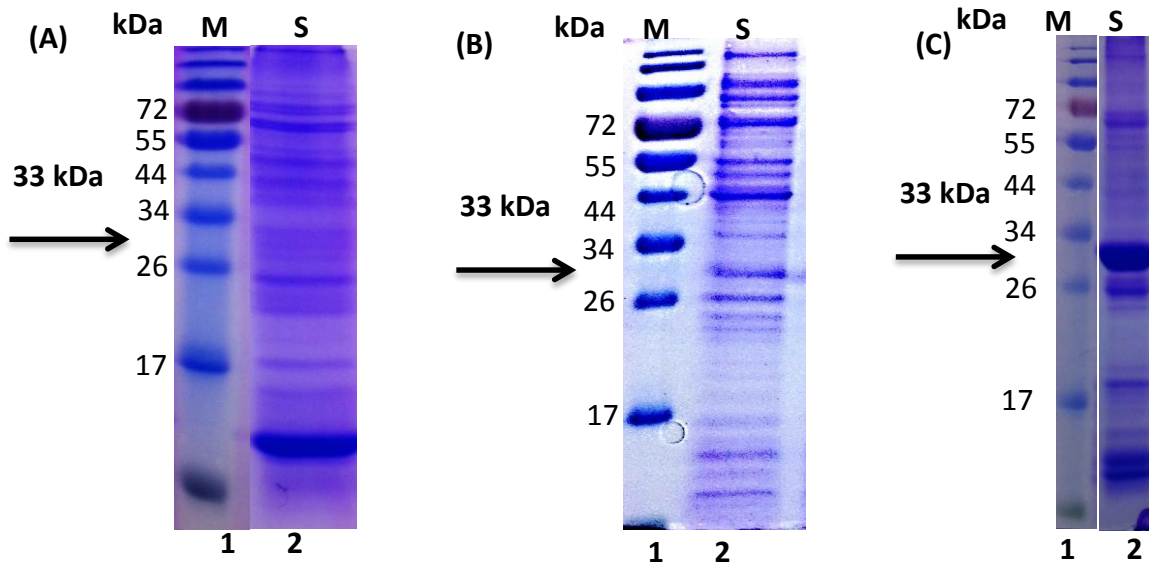
The use of a colony selection protocol was investigated in an attempt to enhance soluble protein expression. The method involves the screening of newly transformed colonies for improved soluble expression using a double colony selection method (Murray, Chen et al. 2010). This method was employed as a means of selecting high protein expression colonies. The analysis of SDS-PAGE gels for these experiments showed no significant increase in soluble expression.

To test the effect on soluble expression the use of molecular chaperones were explored through the use of the takara plasmid kit. Chaperone systems can be employed to improve soluble expression by preventing inclusion body formation. The Hsp 70 chaperone family reduces protein aggregation and aids in the proteolysis of mis-folded proteins (Mogk, Mayer et al. 2002). The Hsp 100 chaperone family has been shown to improve solubilization and disaggregation (Schlieker, Bukau et al. 2002). The Hsp 60 chaperone family additionally has been shown to reduce aggregation and inclusion body formation (Kuczynska-Wisnik, Kędzierska et al. 2002). The effect of the use of the molecular chaperones on enhancing soluble expression was not significant upon analysis of SDS-PAGE gels.

The use of chemical chaperones to improve soluble expression was also explored. The introduction of chemical chaperones like L-arginine, glycylglycine,  $\beta$ -cyclodextrin, sorbitol and PEG was explored. Studies have shown that the addition of chemical chaperones can aid in the solubility of proteins through assisting in correct protein folding and the suppression of aggregation (Prasad, Khadatre et al. 2011) (Ghosh, Rasheedi et al. 2004) (Nath and Rao 2001) (Yazdanparast and Khodarahmi 2005) (Majumder, Basak et al. 2001). As with the use of the colony selection protocol and the use of molecular chaperones, the effect of the use of the chemical chaperones on enhancing soluble expression was not significant upon analysis of SDS-PAGE gels.

In order to increase the soluble expression yields, variations in the cell lysis protocol were explored. Different methods of cell lysis were tested from freeze-thaw, sonication, and enzymatic cell lysis to combinations of all three. The duration, pH and temperature of the cell lysis were also investigated. The components of the lysis buffer were also varied with the addition of different additives to improve cell lysis and soluble protein yield. Chaotropic agents can be used to prevent protein aggregation through prevention of intermolecular interactions, the use of mild chaotropes like  $MgCl_2$  and  $CaCl_2$  were investigated (Xie and Timasheff 1997). The use of kosmotropic agents that stabilize intramolecular protein interactions was also investigated as a means of stabilizing the native structures and preventing aggregation. The use of kosmotropes like  $MgSO_4$  and  $(NH_4)_2SO_4$  were investigated (Neagu, Neagu et al. 2001). The inclusion of L-arginine to the lysis buffer was tested as a means of improving proteins solubility through prevention of protein aggregation (Golovanov, Hautbergue et al. 2004). To aid cell lysis and to improve soluble protein yield the addition of detergents were tested. The effect of non-ionic, zwitterionic and ionic detergents were investigated. The use of detergents like sarkosyl have been reported as a means to purify natively folded proteins from inclusion bodies in the presence of non-ionic and zwitterionic detergents. The use of sarkosyl has been shown to be problematic when purifying proteins using affinity chromatography (Tao, Liu et al. 2010).

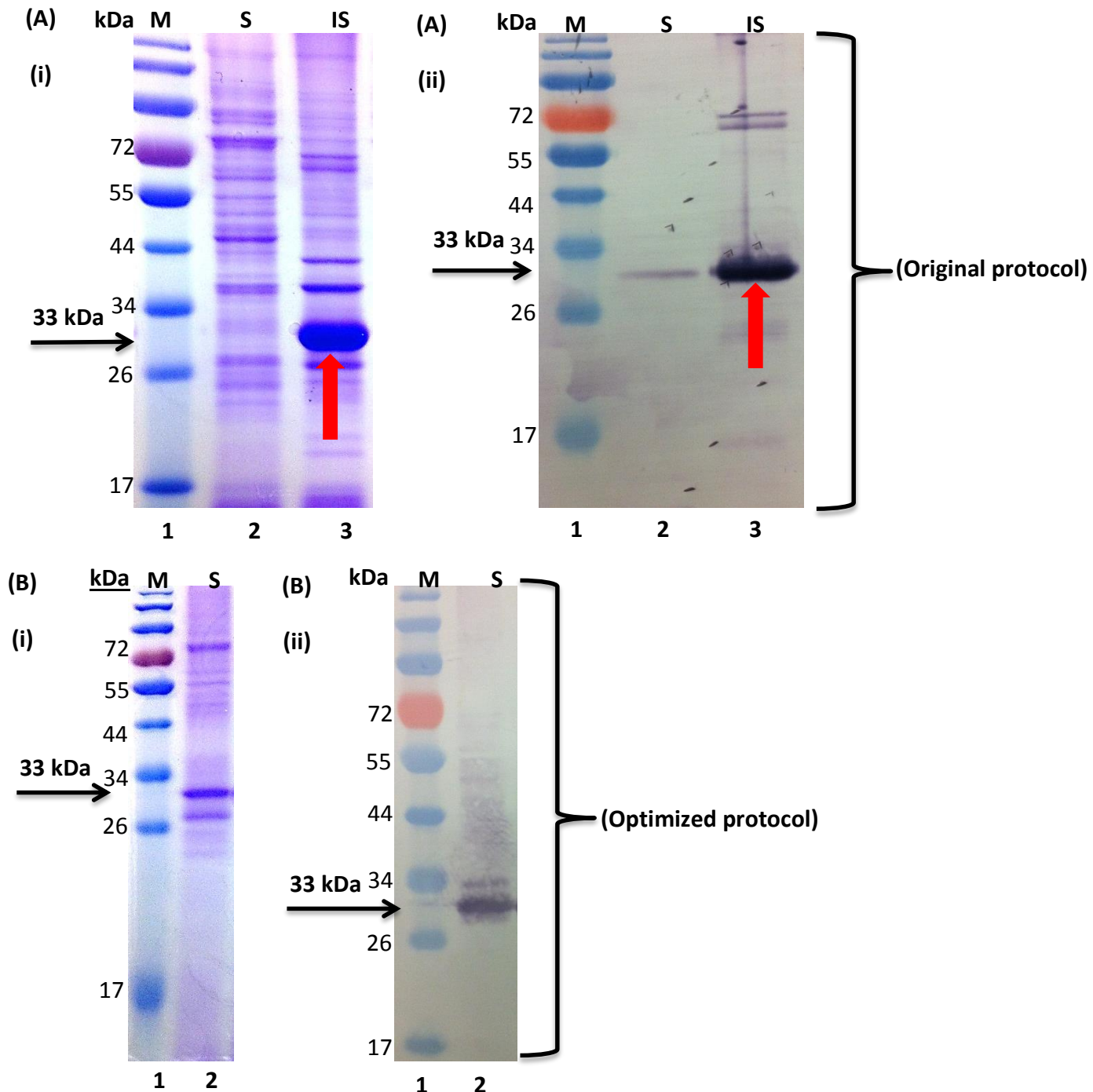
The use of chemical chaperones, molecular chaperones and cell lysis buffer additives was met with varying degrees of success upon analysis of SDS-PAGE gels (see Figure 21), however the issue of protein contamination remained.



**Figure 21:** IMAC purification (3-OST-1) with 100% elution showing examples of the use of chemical chaperones, molecular chaperones and cell lysis buffer additives. (A) Protein expression that was lysed in the presence of L-arginine (chemical chaperone) then IMAC purified (5.6 mg/ml in 500  $\mu$ l), (B) Protein expression in the presence of Takara plasmid 1 (molecular chaperone) then IMAC purified (1.1 mg/ml in 250  $\mu$ l) and (C) Protein expression that was lysed in the presence of Triton X-100 (lysis buffer additive) then IMAC purified (1.3 mg/ml in 3 ml). (1) M= Molecular weight ladder, (2)= 100% elution. Estimated molecular weight of the protein of interest= 33 kDa.

To obtain sufficiently pure protein for downstream experiments a number of variations to the purification strategy were examined. During IMAC purification the concentrations of the imidazole was varied to reduce co-purification of contaminants. The addition of different wash steps were explored from tween-20, triton X-100 to glycerol. The introduction of reducing agents was investigated to reduce co-purification of contaminants in addition to varying the ionic strength of the IMAC buffers. Different purification methods were investigated from ion exchange chromatography, size exclusion chromatography to hydrophobic interaction chromatography. The purification of protein from inclusion bodies was explored. However protein was being lost in the wash steps during the inclusion body preparation. This led to the focussing on the development of an appropriate lysis buffer to obtain increased soluble expression yield.

In an attempt to obtain sufficiently pure protein for downstream experiments a number of aspects of the protein expression and purification protocol was looked at. There were issues in choosing an appropriate protein expression cell line, in addition to poor protein yield and purity. Optimizations to the protein expression and purification protocol to overcome the issues with poor protein yield and purity were investigated. The main focus was on optimizing the protein expression protocol and the methods of harvesting soluble expression. This lead to the development of an optimized protocol (see Figure 22) with further optimization needed in the protein purification protocol (see Figure 23) detailed below.



**Figure 22:** Optimization of the expression protocol (from: (Edavettal, Lee et al. 2004)) for the wildtype and mutant 3-Ost-1. (A) Analysis of the original expression protocol showing (i) SDS-PAGE analysis and (ii) western blot of a soluble protein expression (3-Ost-1). The analysis shows abundant insoluble expression (highlighted by the red arrows) using the original protocol. (B) Analysis of the optimized expression protocol showing (i) SDS-PAGE analysis and (ii) western blot of a soluble protein expression (3-Ost-1) shows improved soluble expression. Lanes: (1) M= Molecular weight ladder, (2) S= Soluble fraction, (3) IS= Insoluble fraction. 3-Ost-1 is shown with the black arrow.

Optimizing the soluble protein expression protocol was achieved through a combination of optimizations to the expression, cell lysis and purification protocols (See Figure 19). Protein expression using *E. coli* (DE3) RIPL proved to be beneficial to the expression of soluble protein and overcoming the presence of rare codons. In addition expression was carried out in the presence of kanamycin and streptomycin. The latter although not noted to be required to maintain the pACYC based plasmid, including streptomycin was beneficial to the expression protocol. The inoculation ratio of 1:100 proved to be beneficial to protein expression. Generating litre cultures using a 4 litre flask with a shaking rate of 200 rpm proved to be of particular importance in regards to protein yield. This was likely due to sufficient aeration which is a known factor in achieving soluble expression. Modifying the culture media recipe was also of significant importance to protein expression. Starting with the recipe for Super broth, tryptone was used in the place of peptone, the addition of glycerol provided an alternative carbon source to the culture and the addition of phosphate buffer allowed the culture media pH to be maintained. Finally in regards to protein expression, reducing culture temperature and increasing the expression time in the presence of a lower concentration of inducer proved to be critical to soluble protein expression.

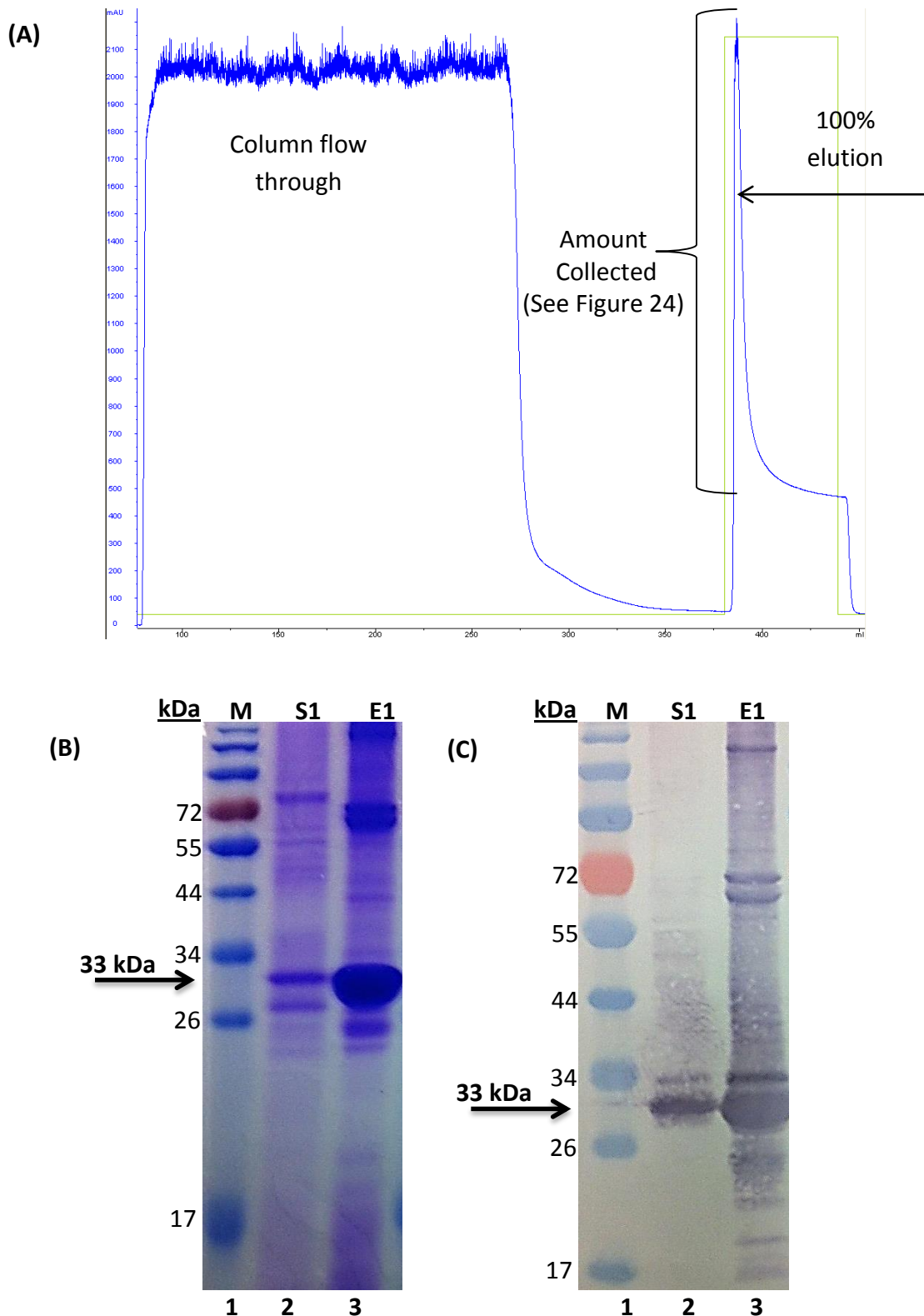
The means of cell lysis was found to be of paramount importance to soluble protein yield, in particular the choice of lysis buffer was of great significance. The strategy that led to obtaining soluble protein involved a combination of membrane solubilisation to disrupt hydrophobic and electrostatic interactions between the protein of interest and the cell membrane and secondly the strategy involved efficient sonication (Everberg, Leiding et al. 2006). The method began with re-suspending the culture pellet in a lysis buffer with an optimized recipe. Starting with membrane solubilisation buffer additions to the recipe was added to improve protein solubility.

To aid protein solubility NaCl was added to increase the ionic strength of the buffer and imidazole was added to aid in the IMAC purification. The ionic strength of the lysis buffer was important. Although increasing the ionic strength increases hydrophobic interactions, a balance between the ionic strength and the hydrophobicity was found. Triton X-100 and tween 20 was added to aid in the solubilisation of the cell membrane and to reduce electrostatic interactions in combination with increasing the ionic

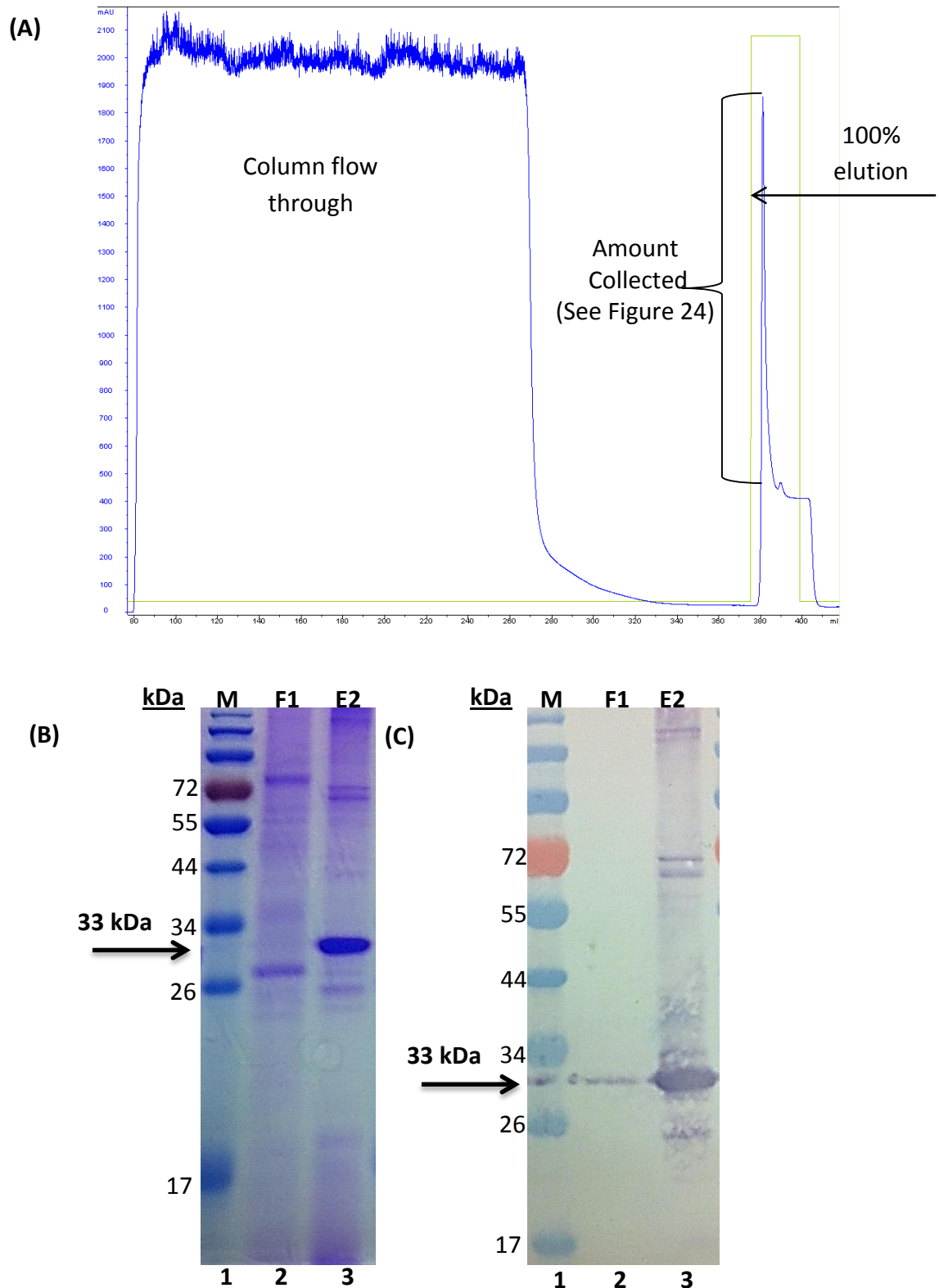
strength of the buffer. MOPS was added not as a buffering agent but to provide increased negative charge to the buffer which was hypothesized to compete with the binding of DNA to the protein of interests positively charged regions through the interactions of the sulphonic acid group in MOPS. In addition research has shown that the zwitteronic buffer MOPS interacts with lipid bilayers and is suggested to lead to a weakening of the van der Waal interactions and repulsion of electrostatic interactions in membranes (Koerner, Palacio et al. 2011). The pH of the buffer was set at 11.5 to aid in protein solubility and to reduce hydrophobic interactions and the cell suspension was kept on ice on ice at all times. After sonication the lysate was prepared for purification.

The successful purification strategy involved multiple rounds of IMAC purifications followed by purification by HIC. The strategy employed here did not use a gradient elution as commonly employed in the literature. Instead the protein was purified by extensive washing steps and purification using multiple rounds of IMAC. HIC was then used as a final polishing step.

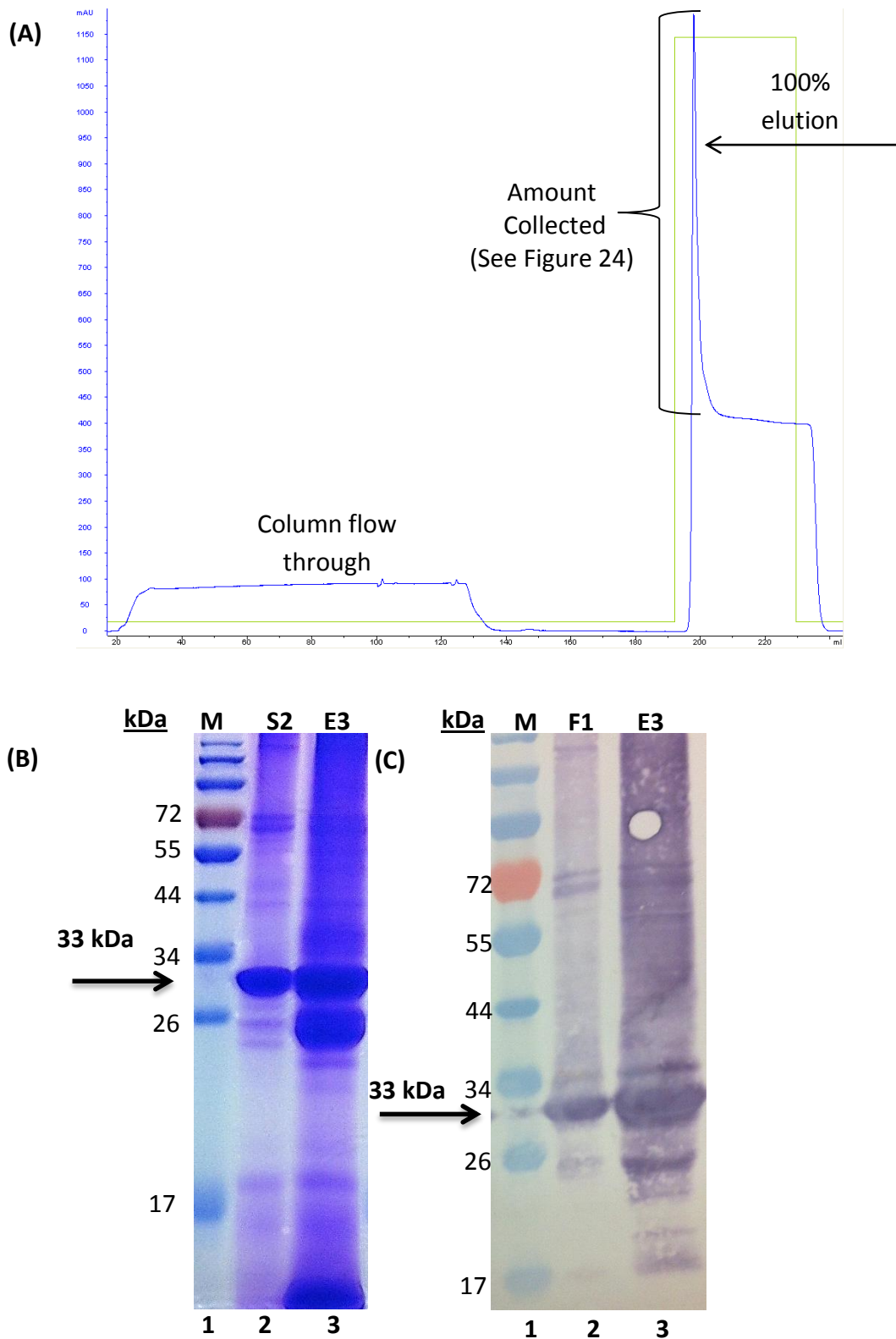
The protein purifications were analysed by SDS-PAGE and western blotting, an example can be seen in Figure 23-26.



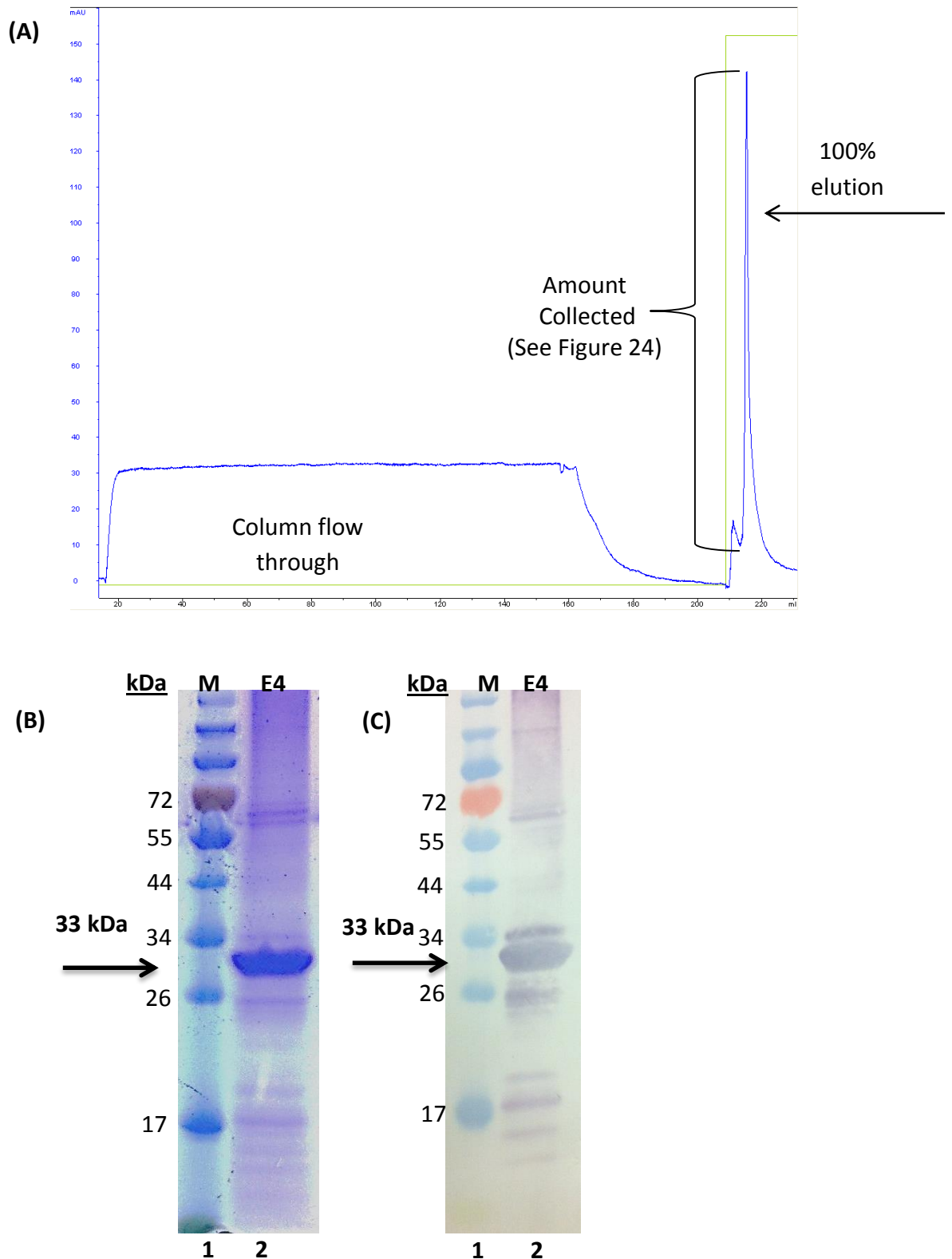
**Figure 23:** First round IMAC purification of wild type 3-OST-1 soluble fraction with a single 100% elution. (A) Chromatogram of IMAC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2) Soluble fraction applied to column (S1= soluble fraction 1) and (3)= First round 100% elution (E1= Elution 1). Estimated molecular weight of the protein of interest= 33 kDa.



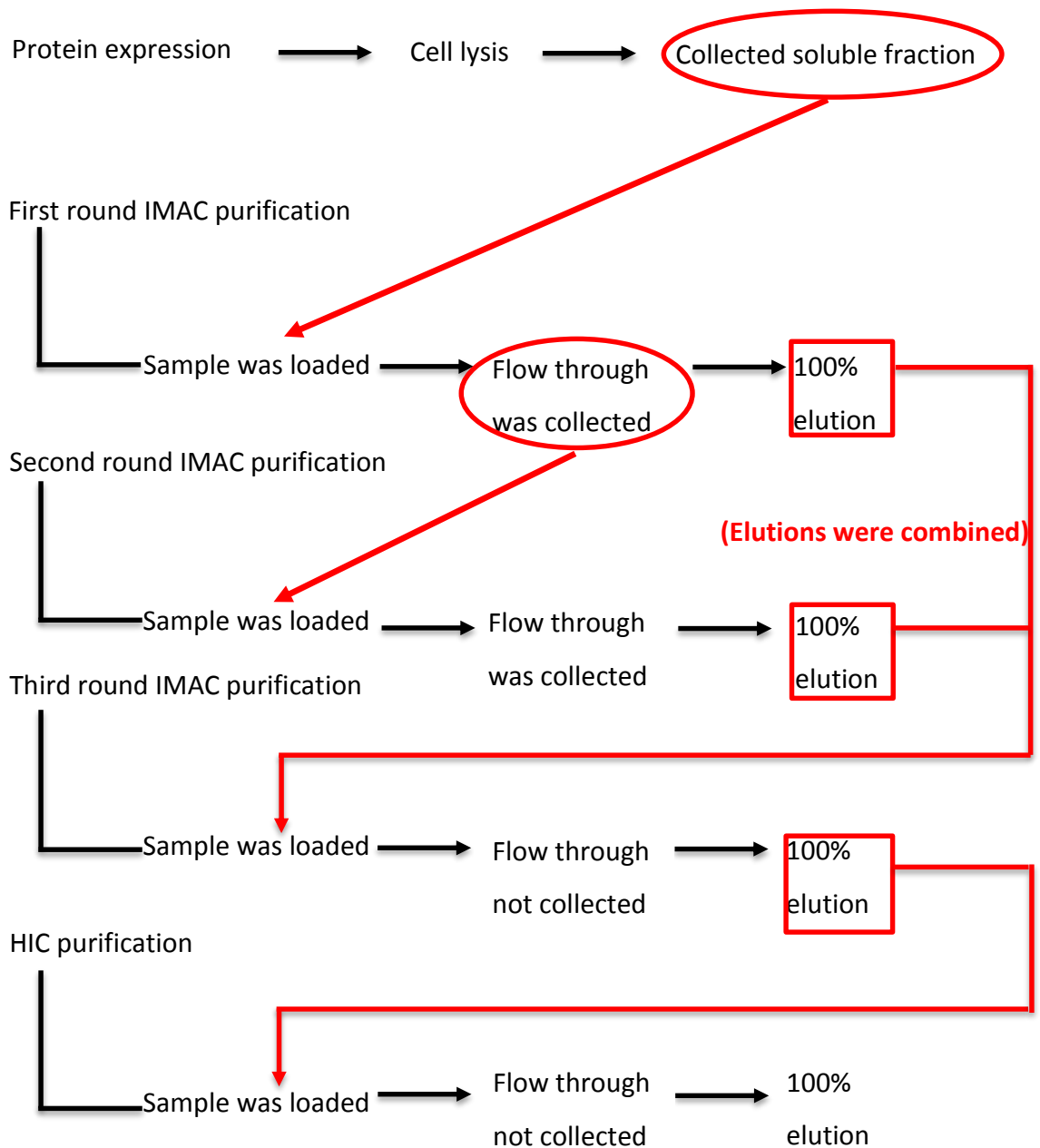
**Figure 24:** Second round IMAC purification (3-OST-1) of the flow through (see figure 20) from the first round IMAC purification with a single 100% elution. (A) Chromatogram of IMAC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2)= IMAC flow through (F1= flow through fraction 1) and (3)= Second round 100% elution (E2= Elution 2). Estimated molecular weight of the protein of interest= 33 kDa.



**Figure 25:** Third round IMAC (3-OST-1) of the combined elutions from the first and second round IMAC's with a single 100% elution. (A) Chromatograph of IMAC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2)= combined elutions applied to column (S2= soluble fraction 2) and (3)= Third round 100% elution (E3= Elution 3). Estimated molecular weight of the protein of interest= 33 kDa.



**Figure 26:** HIC purification (3-OST-1) of the 100% elutions from the third round IMAC with a single 100% elution. (A) Chromatogram of HIC purification, (B) SDS-PAGE analysis of purification and (C) western blot analysis of purification. (1) M= Molecular weight ladder, (2)= 100% elution 1 mg/ml in 5 ml (E4= Elution 4). Estimated molecular weight of the protein of interest= 33 kDa.



**Figure 27:** Flow chart showing the optimized protein purification protocol. After harvesting the soluble fraction from the protein expression in *E. coli*, three rounds of IMAC were performed. The first round involved purifying the soluble fraction followed by a single elution step. The flow through collected from the first round purification step was purified again in the second round purification, followed by a single elution step. The elutions from the first and second round purifications were combined and purified in the third round purification, followed by a single elution step. The elution from the third round purification was then purified by HIC with a single elution step. The samples loaded onto the columns are indicated in red.

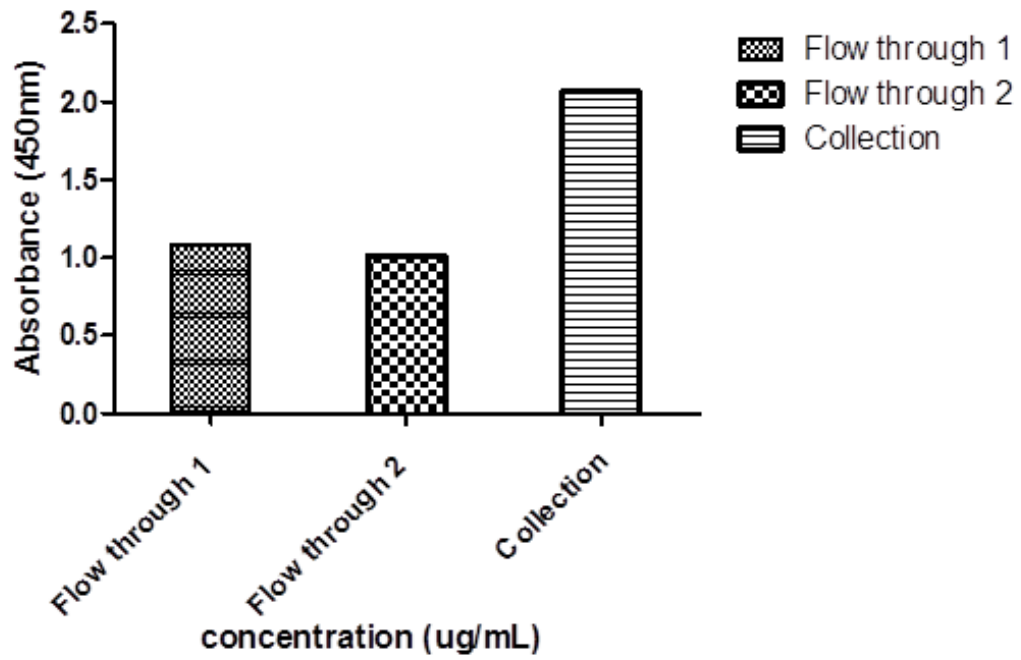
## **Characterization of wild type and mutants:**

To test that the wild type and the mutants bind to heparin prior to performing SPR the use of Enzyme Linked Immunosorbant Assay (ELISA) to confirm binding was employed. The development of an ELISA protocol required optimization to effectively detect protein binding to heparin.

The two predominant methods used for an ELISA are the use of conjugates like heparin-biotin and heparin-bovine serum albumin (BSA) (Foxall, Holme et al. 1995) (Barth, Schäfer et al. 2003). With the limited commercial availability of the heparin-BSA conjugate the use of the heparin-biotin conjugate was investigated.

The method involves the immobilization of the heparin-biotin to a streptavidin coated plate. Initially the immobilization of the heparin-biotin to a streptavidin was tested using SPR. Streptavidin was immobilized to a carboxymethyl chip via amine coupling. The heparin-biotin conjugate was then injected onto the chip and the response was monitored. The level of immobilization was very low giving a response far below that expected of a heparin-biotin conjugate. Upon repeating the immobilization procedure a protein-biotin conjugate was injected over the streptavidin coated chip and as before there was no significant response.

The potential of free biotin in the sample was investigated as free biotin would out compete the binding of the heparin-biotin conjugate. To investigate this possibility the free biotin was removed through filtration. The heparin-biotin conjugate was applied to a spin column with a 10,000 molecular weight cut off and all the fractions were collected. The fractions were then analysed by ELISA to verify the presence of the free biotin and the heparin-biotin conjugate using a streptavidin-HRP conjugate (See Figure 28). It should be noted that further concentration of the sample may have removed more free-biotin from the sample.



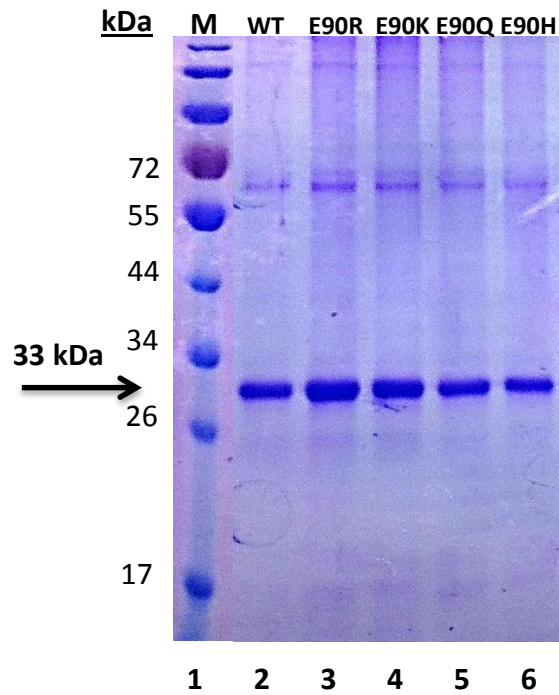
**Figure 28:** Free biotin test indicating the presence of biotin in both the flow through fractions and the collected fraction. The flow through consists of biotin that passed through the spin column membrane. The flow through was collected after the first and second centrifuge steps. The collection consists of the concentrated heparin-biotin conjugate.

The ELISA data shows that there was an abundance of free biotin in the sample, with the level of free biotin making up approximately half of the total concentration of the sample. Due to this observation, the use of commercial heparin-biotin conjugate was abandoned in favour of a different heparin immobilization method.

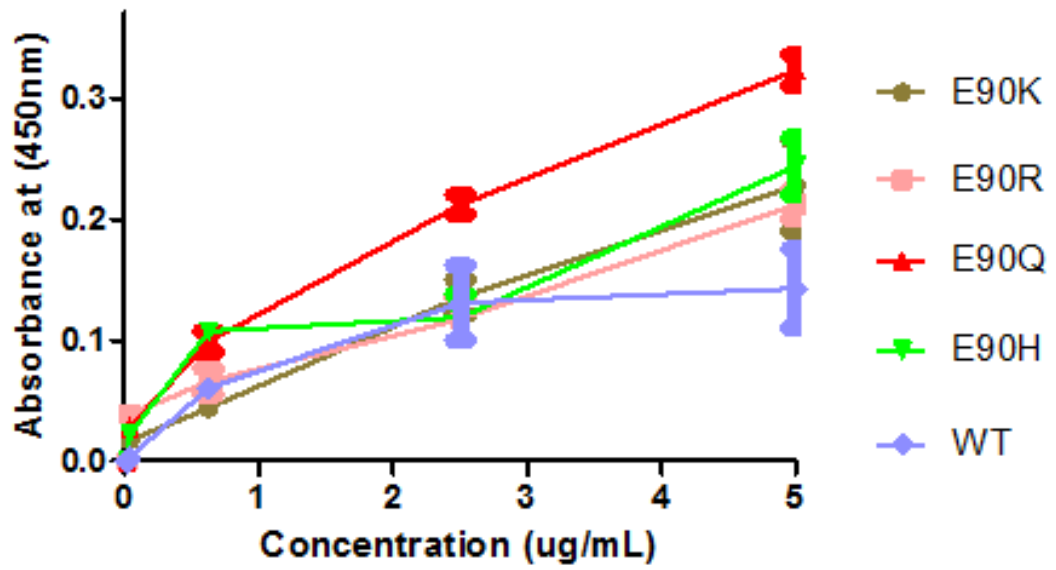
A different approach needed to be taken to immobilize heparin to an ELISA plate. Immobilization via passive absorption in carbonate buffer was tested using Microtest plates (Starsted), Maxisorb plates (eBioscience) and Minimal binding plates (Greiner). This approach was unsuccessful due to the binding of the protein of interest to the plate or the lack of binding detected at different concentrations of protein or heparin. Similarly in the use of heparin binding plates (BD Bioscience) the binding of the protein of interest to the plate prevented useful data from being obtained using this method.

The coating of the ELISA plate with a linker molecule was also investigated using poly-L-lysine. The positively charged polymer was immobilized to the plate via passive absorption and was used to capture heparin through electrostatic interaction. With sufficient blocking using 2% BSA in PBS, the method allowed the binding of the mutants to heparin to be determined despite the binding of the protein of interest to poly-L-lysine.

Although the method requires further optimization as increasing protein concentrations would be required to determine the protein binding saturation point and the inconsistency of the method prevented application in an inhibition ELISA. The data presented shows the protein binding to heparin with increasing protein concentrations (See Figure 29) (See Figure 30). The charge and the hydrophobic nature of the protein of interest made data collection a difficult process. The background observed is due to the binding of the protein of interest to poly-L-lysine. When the experimental controls were factored in (through background subtraction), the data indicates that the protein samples tested all bind to heparin.



**Figure 29:** SDS-PAGE showing HIC purified protein used in the ELISA experiment, (1) M= Molecular weight ladder, (2) Wild type, (3) E90R, (4) E90K, (5) E90Q and (6) E90H. Molecular weight of the protein of interest= 33 kDa. The protein concentrations were at 1 mg/ml for each sample and the total yield as follows: Wild type 6 mg in 6 ml, E90R 2 mg in 2 ml, E90K 3 mg in 3 ml, E90Q 4 mg in 4 ml and E90H 8 mg in 8 ml. The protein yield was obtained in the same manner as described in Figures 20-24, with the final 100% elution from the HIC purification making up the protein yield detailed above.



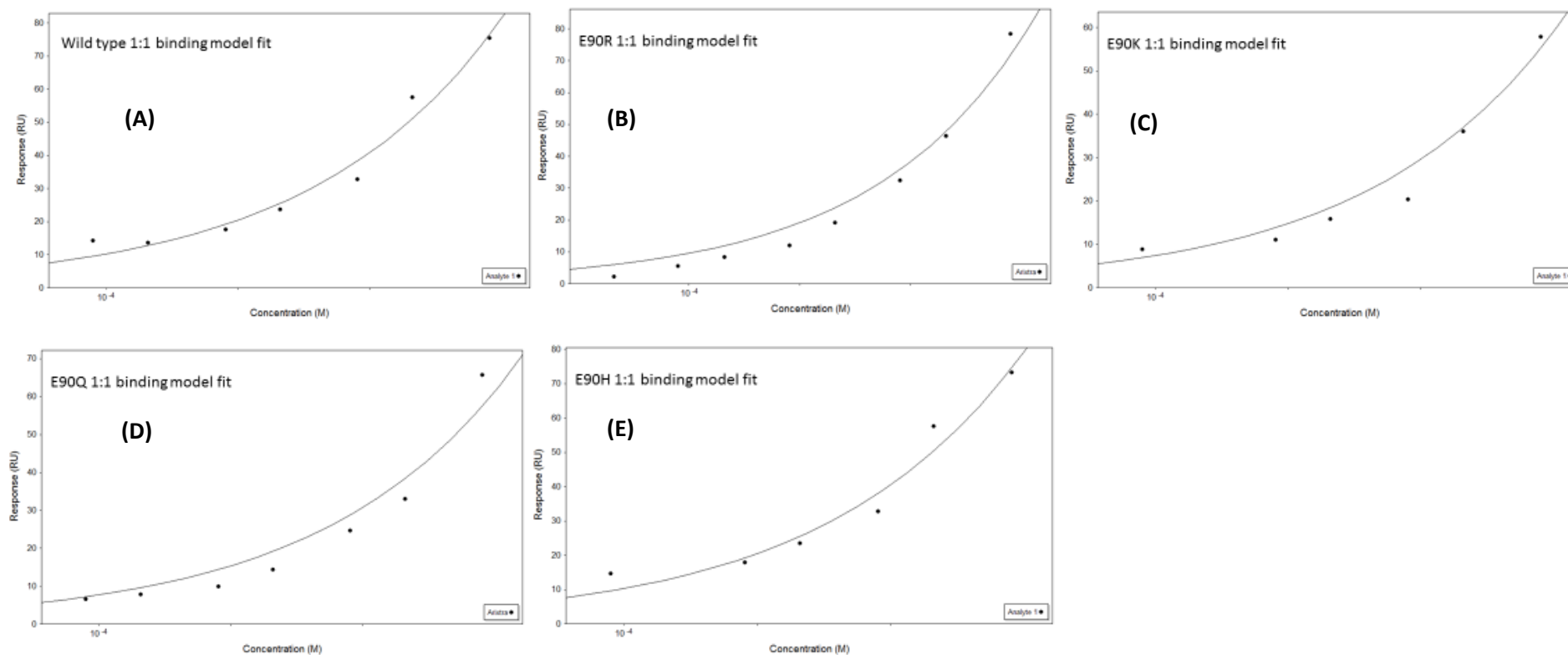
**Figure 30:** ELISA data of the wild type and mutant binding to heparin. The data indicated was blanked subtracted. The blank control wells consisted of wells coated in poly-L-lysine (absent heparin). The absorbance was measured at a wavelength of 450 nm and was plotted against the protein concentration in the corresponding well.

### **SPR analysis of wild type and mutants binding to Arixtra:**

The effect of mutation of position 90 in the interaction to Arixtra was investigated using SPR (see Table 2). In regards to the E90H mutant an additional experiment was carried out. In this experiment the level of protonation of the imidazole side chain was investigated through the binding of the E90H mutant to Arixtra at pH 7 and pH 5.6. At pH 7 the histidine of the E90H would act as a weak base and the positive charge may form a favourable interaction with the 3-O-sulphated position of the glucosamine residue. Performing the SPR experiment at pH 5.6 when theoretically the histidine residue would be fully protonated may therefore have a positive effect on binding (See appropriate section in the materials and methods chapter).

The data generated was processed by the Reichert SPR auto link software and was exported to Scrubber 2. The data was examined and was fitted to a 1:1 interaction binding model. However in order to fit the data, the reference subtraction involved subtracting the curve and the difference between the left and right channel, instead of just the difference in the data. This allowed the association and dissociation phases to be fit to the 1:1 interaction model (See Figure 31) (See Figure 33-37). However, it is very important to note that although the model fit the 1:1 interaction model, this does not accurately describe the binding data generated. Even though analysing the interaction in this way allows kinetics data to be generated, it is not an accurate representation of the binding data.

Upon fitting the data to a 1:1 binding model, the binding curves fit to the model with a residual standard deviation of less than 10 in each case (see Figure 28). However a 1:1 model describes a single binding event, inspection of the raw data would suggest this is not the case. Although the data generated fitted the model, it does not represent all the data points from the raw data (due to the way the data was processed as described above). The data was processed to fit the 1:1 interaction model although this does not represent the binding event when one considers the raw binding data. This is reinforced by Muñoz *et al.* in which they found that 3-OST-1 does not fit at 1:1 model (Muñoz, Xu et al. 2006).

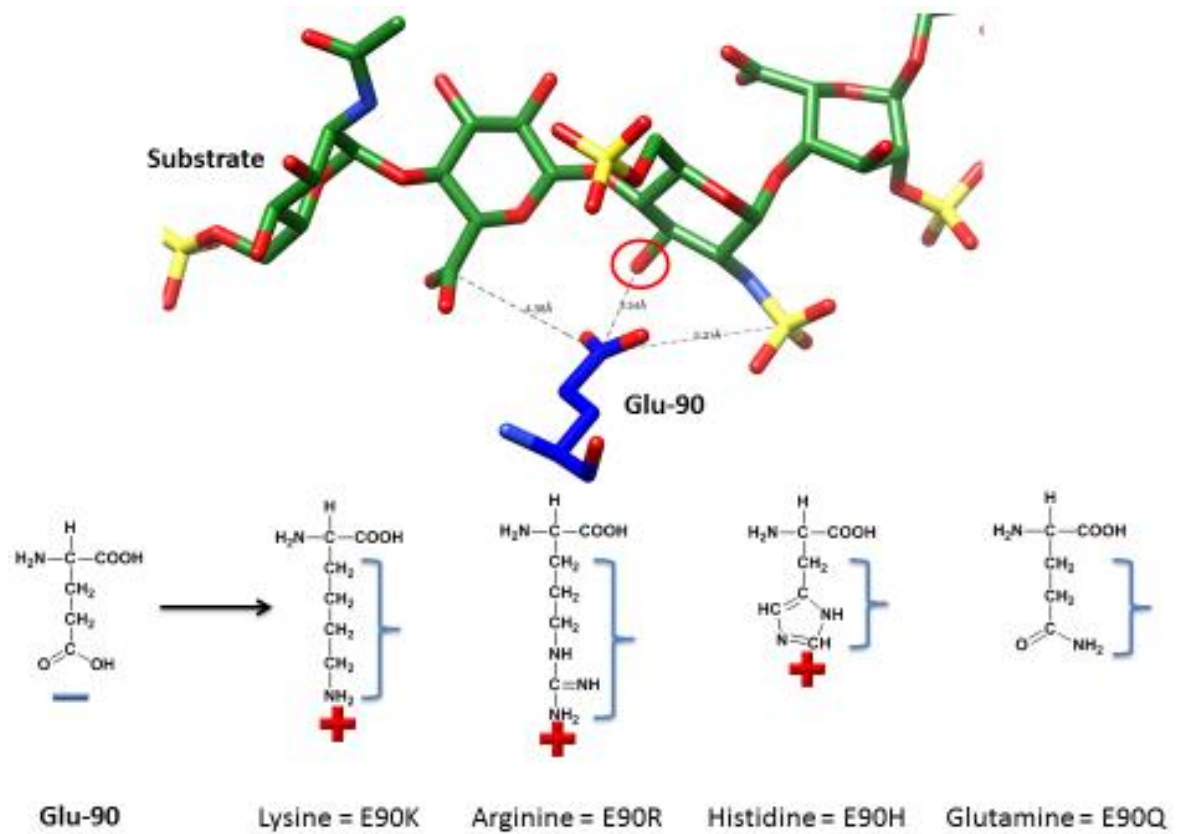


**Figure 31:** SPR binding curve fit to a 1:1 interaction model with immobilized protein as the ligand and Arixtra as the analyte. **(A)** Wild type versus Arixtra (Res SD : 4.219), **(B)** E90R versus Arixtra (Res SD: 2.425), **(C)** E90K versus Arixtra (Res SD: 3.413), **(D)** E90Q versus Arixtra (Res SD: 3.625) and **(E)** E90H versus Arixtra (Res SD: 9.876). The residual standard deviation (Res SD) describes the statistical estimate of the fit of the experimental data to the model, a value below 10 is considered a good fit.

**Table 2:** SPR affinity and kinetics data of the interaction of the protein samples to Arixtra:

	$K_{on}^a$ ( $M^{-1} s^{-1}$ )	$K_{off}^a$ ( $s^{-1}$ )	$K_{+1}^b$ ( $s^{-1}$ )	$K_{-1}^b$ ( $s^{-1}$ )	$R_{max}$	$\chi^2$	$K_D$ (Moles)
WT	0.158	0.0235	0.0208	1.62E-03	1.91E+03	55.6	11.5 mM
E90R	0.11	0.0204	0.102	0.0184	1.71E+03	20.6	33 mM
E90H	136	0.293	1.99E-03	8.92E-04	395	69.2	0.96 mM
E90Q	0.131	2.68E-03	3.28E-03	4.36E-06	1.47E+03	239	27 $\mu$ M
E90K	0.163	7.24E-04	0.0352	3.02E-05	8.45E+02	156	3.8 $\mu$ M

**Table 2:** SPR affinity and kinetic data of the interaction of 3-OST-1 wild type (WT), E90R, E90H, E90Q and E90K with Arixtra.  $K_{on}^a$  and  $K_{off}^a$ : the on and off rate constants of the binding step.  $K_{+1}^b$  and  $K_{-1}^b$ : the on and off rate constants of the conformational change. The data fit was for a two state interaction model (Muñoz, Xu et al. 2006).



**Figure 32:** Glu-90 was mutated to a lysine, arginine, histidine and a glutamine.

Highlighted above are the structures of the wild type and mutant residues with the length of the amino acid side chains indicated and the charges of the side chains highlighted by the red plus symbol and the blue negative symbol.

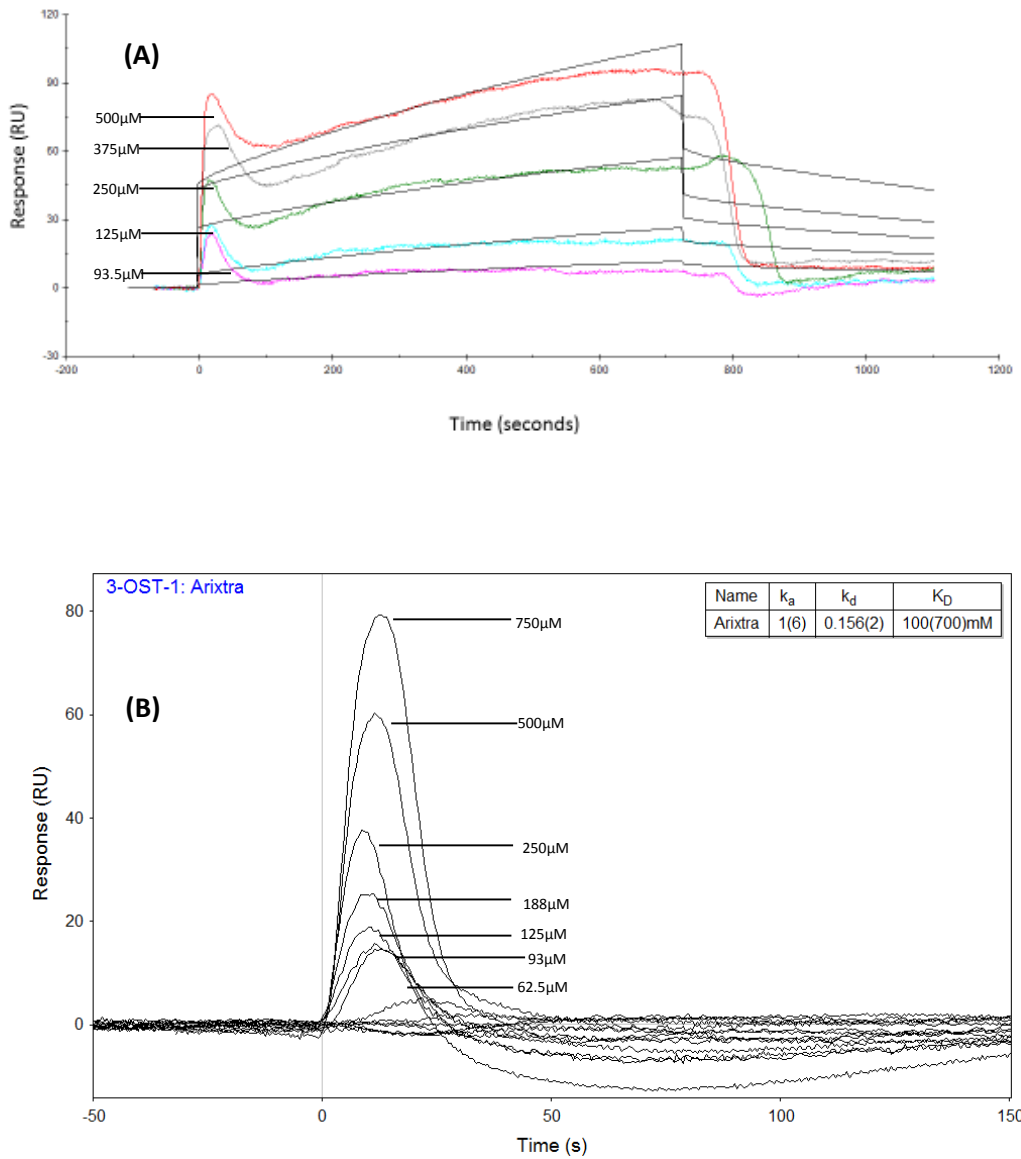
In order for binding curve data to reach equilibrium, the rate of ligand and analyte complex formation and break up must be equal. In order for binding curves to fit an equilibrium model, the curves are needed to have reach equilibrium and enough data must be available to fit correctly to the model. As one can see in figures 33-37, the binding curves generated do not reach equilibrium. The data generated in this study did not fit to an equilibrium mode. Even focusing on one small section of the curves did not allow the curves to fit an equilibrium model.

SPR analysis of 3-OST-1 has been shown to fit to a two state binding model that factor in conformational changes (Muñoz, Xu et al. 2006). Studies have indicated that a conformational change occurs in the C-terminus of 3-OST-1 upon binding to HS. This was shown in a study performed by Edavettal *et al*, where the relative susceptibility of lysine residues to acetylation by acetic anhydride was determined in order to ascertain the solvent accessibility of the lysine residues. By way of this approach conformational changes in 3-OST-1 upon binding to HS can be determined through the use of MALDI MS. Analysis of the ratio of peptides containing protonated lysine residues and  $^2\text{H}$ -labelled residues allowed a comparison to be made between unbound and bound 3-OST-1. This study indicates that a conformational change does occur due as shown by MALDI MS and the increase in the solvent accessibility that was observed in 3-OST-1 upon binding to HS (Edavettal, Lee et al. 2004).

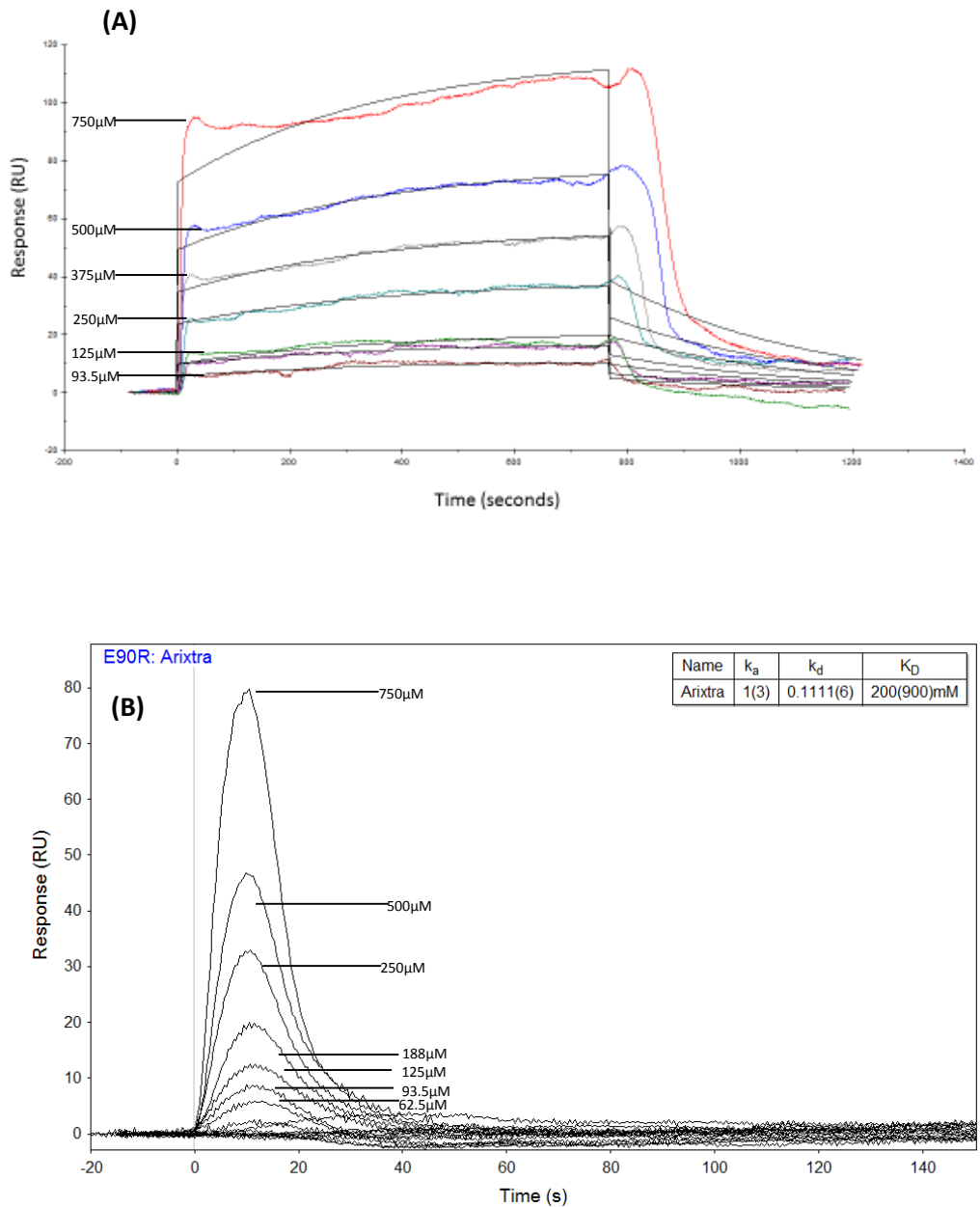
With this in mind the data was exported to BIAevaluation software and the data was fitted to a two state interaction as described by Muñoz *et al*.(Muñoz, Xu et al. 2006). The kinetics data generated is summarized in the Table 2. From this point forward the SPR data discussed will be to the data fitted to the two state model.

The results of the SPR analysis of the wild type and mutant binding to Arixtra highlight the difference between the samples (See Figure 33-37 and Table 2). The affinity of the mutants for Arixtra suggests the importance of positive charge at position 90 in the interaction with 3-OST-1 modified heparin, in particular the importance of amino acid side chain length. Visual inspection of the data fit can reveal the experimental data has a good fit to the initial association phase in each sample. After this initial phase there is a curve deviation observed in each sample. This effect is most prominent in the wild

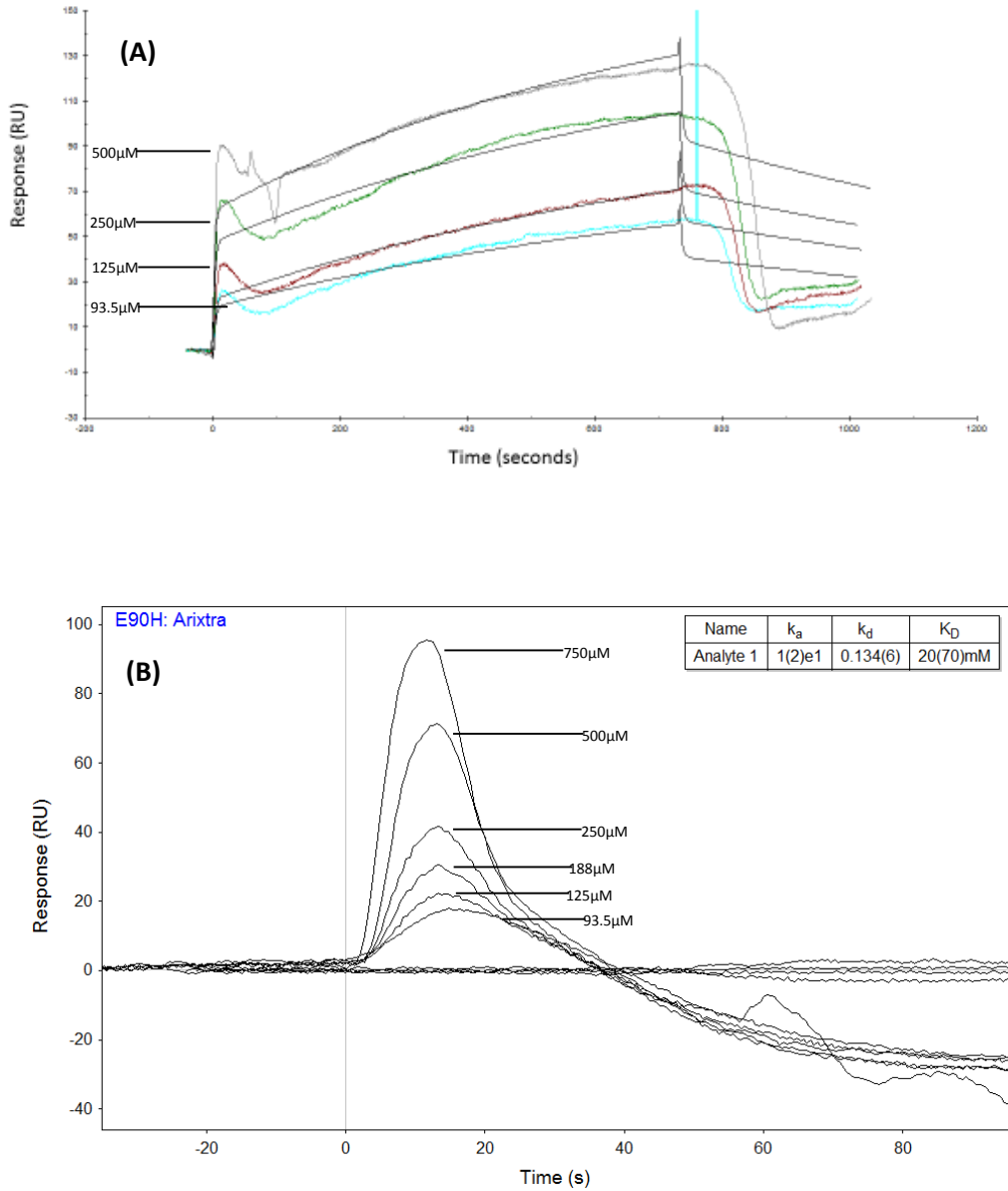
type and E90H samples. The final dissociation phase deviates from the fit in each sample. The  $\text{Chi}^2$  value reported relative to the  $R_{\text{max}}$  might suggest a mass transport effect may be affecting the goodness of the fit.



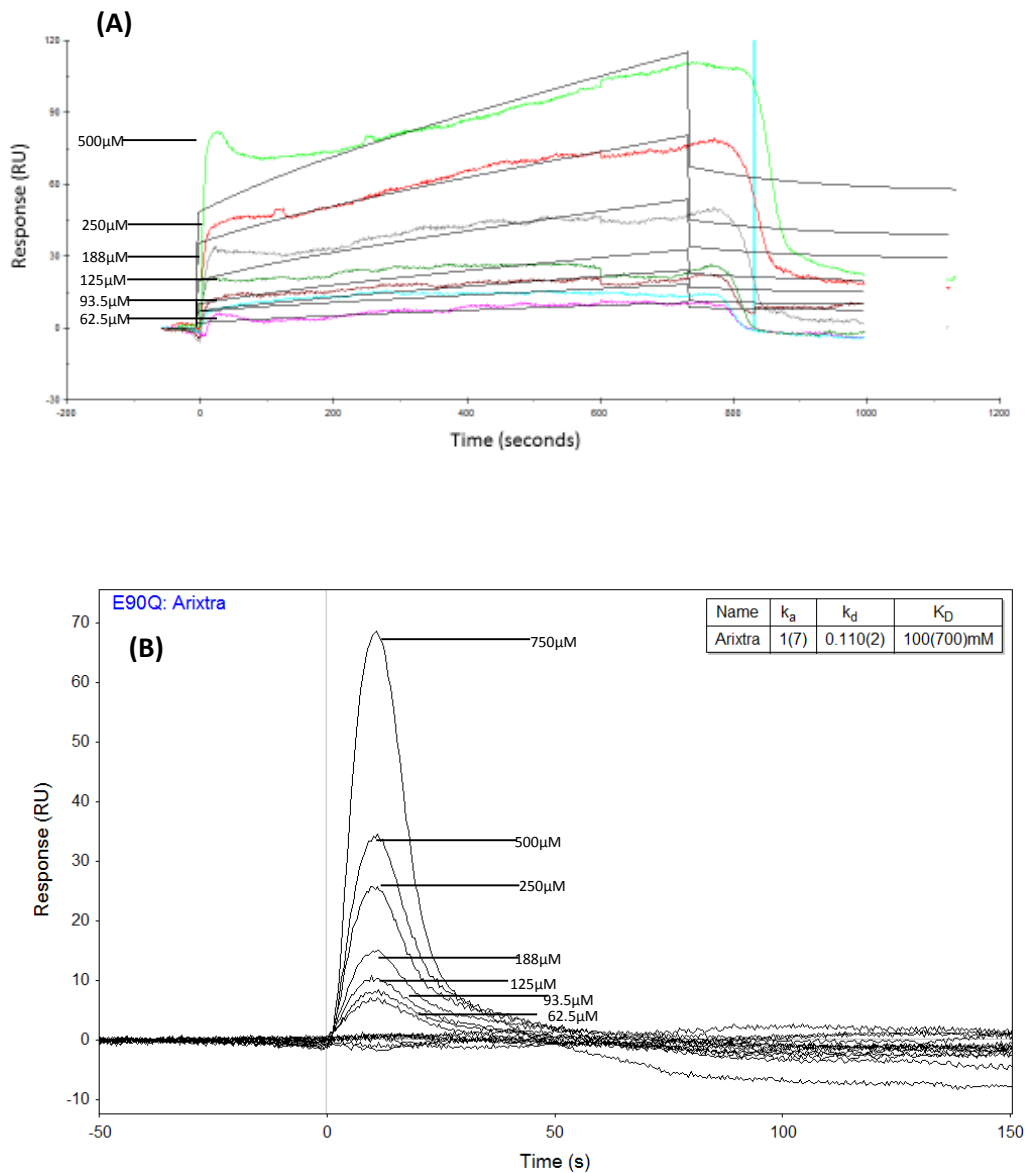
**Figure 33:** Sensograms of the interaction of wild type 3-OST-1 with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.



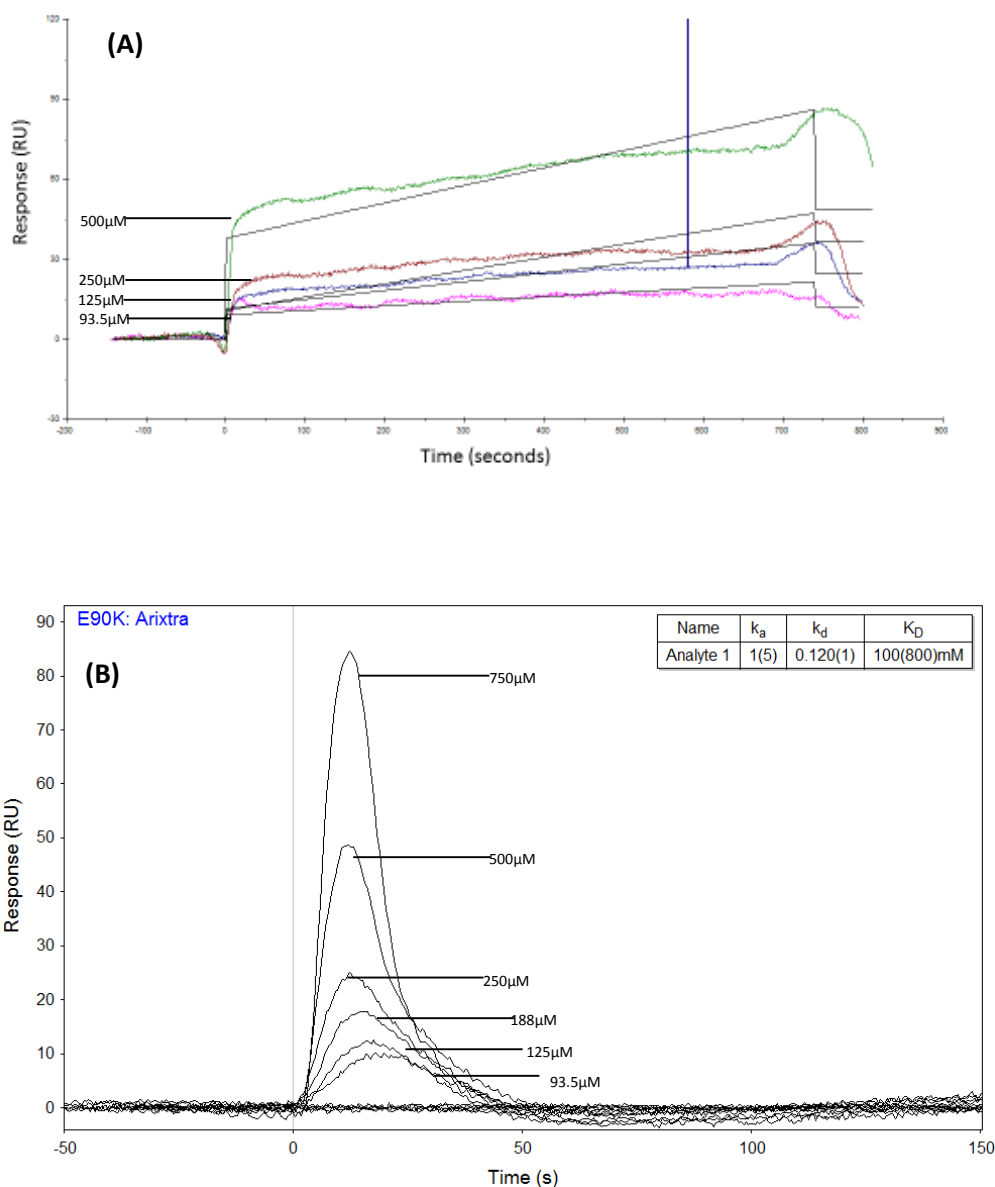
**Figure 34:** Sensograms of the interaction of E90R with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.



**Figure 35:** Sensograms of the interaction of E90H with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.



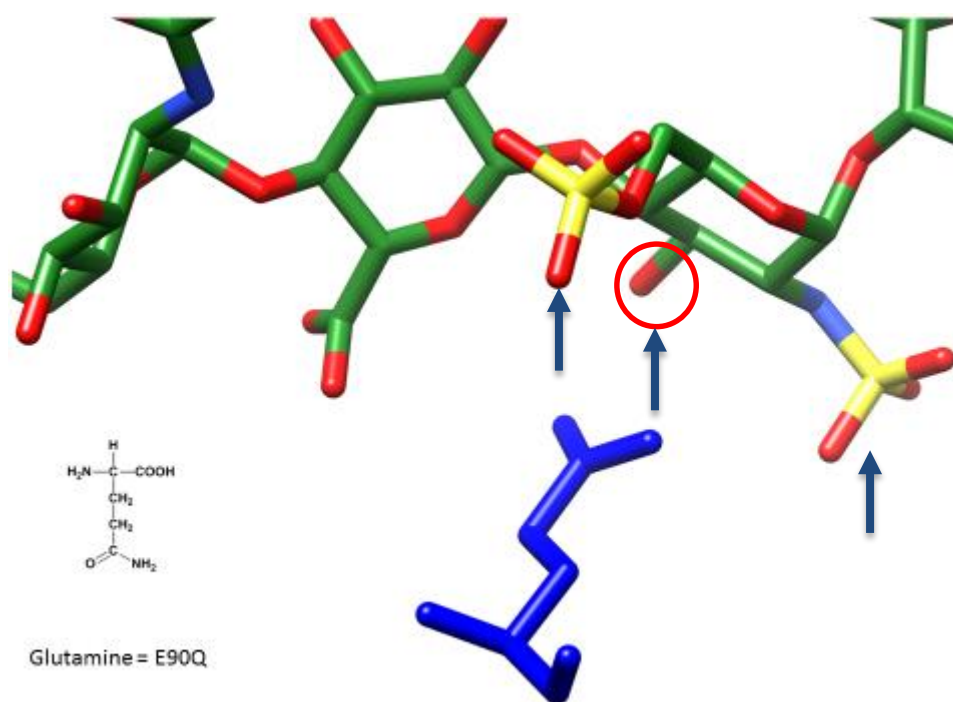
**Figure 36:** Sensograms of the interaction of E90Q with Arixtra (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.



**Figure 37:** Sensograms of the interaction of E90K with Arixtra. (A) Experimental data fit to a two state interaction model showing the model curve fit to the experimental curve (B) Kinetics data generated from the fit to a 1:1 interaction model.

The affinities of the wild type and the E90Q mutant would suggest that removing the negative charge of Glu-90 and replacing the amino acid with a glutamine in the E90Q mutant plays a significant role in the binding to 3-OST-1 modified heparin. One could suggest that the repulsive effect of the negative charge-to-charge interaction between the carboxyl group of Glu-90 and the 3-O-sulphated position of the binding sugar is significant. The absence of this carboxyl group in the E90Q mutant improves the affinity of the interaction relative to the wild type.

With a glutamine at position 90, the potential for a clash with the ligand should be considered. As structurally a glutamine is very similar to that of the glutamic acid in the wild type one could imagine that there would not be a clash. However it is very important to note that a glutamine could be present in different rotomers, one of which could lead to a clash with the ligand (see Figure 38). In addition, as with the possibility of a glutamine at position 90 being present in different rotomers, one must consider that with different rotomers, the ligand may be in a different conformation. This may cause the ligand to clash with the mutant.



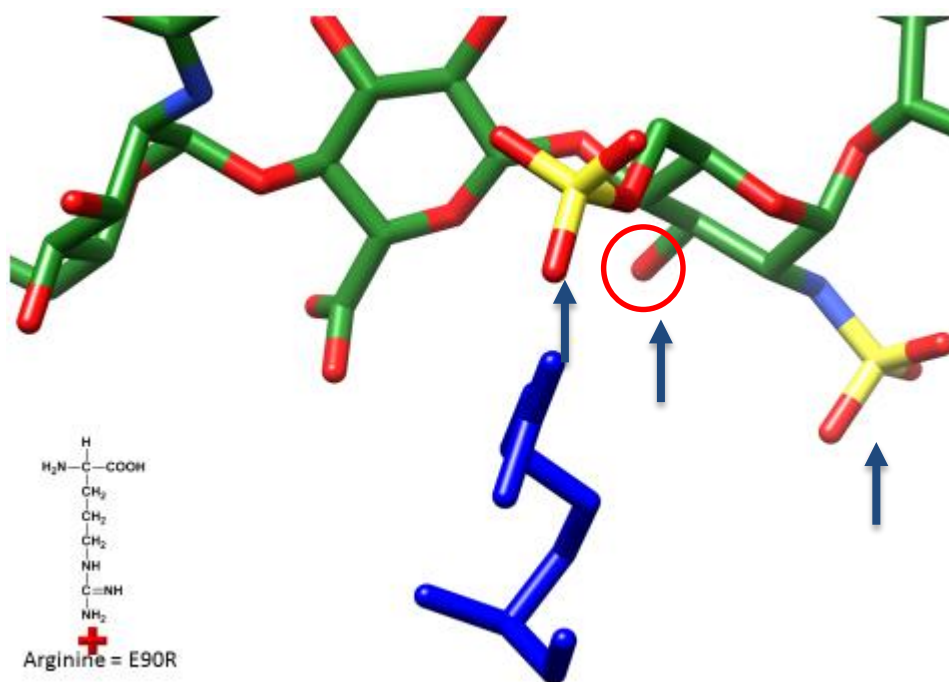
**Figure 38:** Position-90 substituted with a glutamine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotamer that the glutamine residues conforms to, a clash may occur. The positions that a glutamine substituted at position-90 may dash with are indicated by the blue arrows.

The addition of a guanidine side chain with the E90R mutant has a slight negative effect on the affinity relative to the wild type. The size of the side chain substitution with this mutant may explain the reduction in affinity. This could be explained by considering that of the mutants included in the study the E90R mutant has the largest side chain. The effect of this substitution on binding may cause a structural clash with the neighbouring residue Lys-123 or with the ligand.

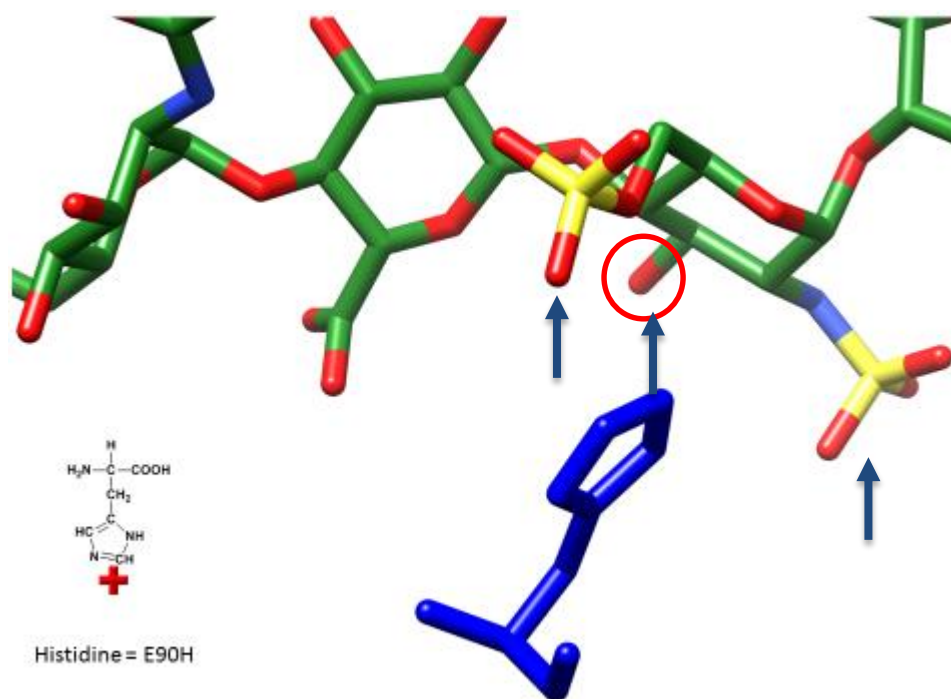
With an arginine at position 90, the potential for a clash with the ligand may be more likely. Structurally an arginine has a larger side chain than that of the glutamic acid in the wild type. This might increase the possibility of a clash with the ligand. Again it is very important to note that an arginine could be present in different rotomers, one of which could lead to a clash with the ligand (see Figure 39). This is also dependent on the possibility of the ligand being present in different conformations that may cause the ligand to clash with the mutant.

The importance of side chain length becomes more apparent with the comparison of the affinity of the E90H mutant with the E90R mutant. Reduction of the side chain length with the addition of the imidazole group in the E90H mutant appears to favour the interaction of the mutant with Arixtra. In the E90H mutant the imidazole group may result in less structural clashes compared to the guanidine group in the E90R mutant and may possibly explain the relative improvement in affinity.

The potential for a clash with the ligand with a histidine at position 90 compared to the other mutants may be less likely. As a histidine has the smallest side chain of the amino acids included in the study, it may contribute to a reduction in the chance for a clash with the ligand. Again it is very important to note that a histidine could be present in different rotomers. However there is a possibility that the mutant being present in different rotomers may not contribute to a ligand clash due to the amino acid side chain length (see Figure 40).



**Figure 39:** Position-90 substituted with an arginine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotamer that the arginine residues conforms to, a clash may occur. The positions that an arginine substituted at position-90 may dash with are indicated by the blue arrows.

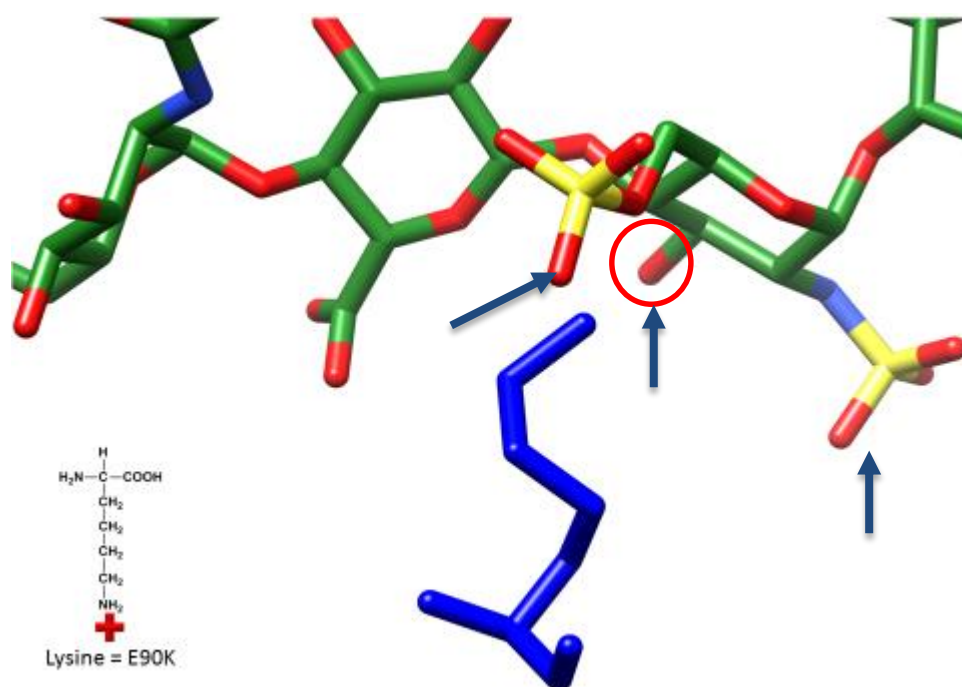


**Figure 40:** Position-90 substituted with a histidine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotamer that the histidine residues conforms to, a clash may occur. The positions that a histidine substituted at position-90 may dash with are indicated by the blue arrows.

With a lysine at position 90, the potential for a clash with the ligand should be considered. Structurally a lysine has the second largest side chain of the mutants in this study. This might increase the possibility of a clash with the ligand relative to the glutamine and histidine mutants. Again it is very important to note that a lysine could be present in different rotomers, one of which could lead to a clash with the ligand (see Figure 41). As with the other mutants this is also dependent on the possibility of the ligand being present in different conformations that may cause the ligand to clash with the mutant.

The effect of side chain length and positive charge at position 90 is most apparent with the E90K mutant. Relative to the other samples included in this study the E90K mutant proved to be the best binder. This may be due to the side chain substitution of this mutant. Considering the three positively charged residues of the mutants in this study, lysine may form the most structurally favourable interaction with the 3-O-sulphated position of the acceptor sugar. The side chain of the lysine mutant could possibly be of sufficient length to interact with the 3-O-sulphated position of the acceptor sugar while not forming any clashes with the neighbouring residues or the ligand.

In comparison to the E90Q mutant the introduction of the amino group in E90K enhances affinity to Arixtra. This is likely due to a charge-to-charge interaction between the positively charged amino group and the negatively charged 3-O-sulphated position of the acceptor sugar.



**Figure 41:** Position-90 substituted with a lysine residue and the 3-OH position of a glucosamine residue highlighted by the red circle (position sulphated by 3-OST-1). With the 3-OH position sulphated and depending on the rotomer that the lysine residues conforms to, a clash may occur. The positions that a lysine substituted at position-90 may dash with are indicated by the blue arrows.

The significance of positive charge in relation to side chain length appears to be of importance. The additional experiment performed using the E90H mutant at pH 5.6 where the imidazole of the E90H mutant at this pH would potentially be fully protonated did not show binding (Patronov, Dimitrov et al. 2012). However considering the affinity of the mutants E90H, E90K and E90R at pH 7 the lysine mutant has the second largest side chain and is in the middle in terms of positive charge. This might suggest that the lysine mutant provides the optimal side chain length while still providing favourable charge-to-charge interaction with the charged 3-O-sulphated position of the acceptor sugar.

The kinetics study revealed differences in association and dissociation rates of the samples tested. In the initial association phase the samples have similar association rates except the E90H mutant, being much slower. This may be due to a function of side chain length as the imidazole group of the E90H mutant is the shortest of the samples in this study. The initial dissociation rates of the samples vary with the rates of the E90Q and the E90K mutants being much faster relative to the other samples. After the conformational change the association rates also vary. The E90Q and the E90H, the least charges of the samples, have the fastest association rates. This suggests that after the conformational change that charge effects the rate of association. Except for E90R the dissociation rates after the conformational change are fast. One could suggest that as E90R has the largest amino acid side chain that this has an effect on the dissociation rate.

In comparison to the reported kinetics of the wild type and E90Q mutant the kinetics rates are slower in the study performed by Muñoz *et al.* (Muñoz, Xu et al. 2006). The association and dissociation phases are faster in this study, in particular the dissociation rates are much faster. This is likely due to the interaction of the samples included in this study with the enzymatic product. The 3-O-sulphated position of the acceptor sugar in Arixtra may therefore be responsible for the reduction in kinetic rates. In the case of these experiments Arixtra only occupies part of the protein active site, one could suggest the possibility that the affinity and the kinetics observed may have also been affected by the size of the ligand used in this study. Moon *et al.* has shown that Arg-268 plays a role in substrate activity which is located further along the

protein active site than that likely occupied by Arixtra when observing the crystal structure produced by Moon et al (Moon, Xu et al. 2012).

The SPR and kinetics data shown in this study provide insight into the interaction of 3-OST-1 and the enzymatic product. The study also provides insight into the importance of amino acid side chain length and charge at position 90.

## **Obstacles and Difficulties**

The goal of this project was achieved through extensive research and the difficulties faced required innovation to surpass. The difficulties faced can be divided in three parts:

1. The generation of single point mutants of 3-OST-1 required the extensive work on a number of different procedures from plasmid DNA isolation to SDM and transformation protocols.
2. The purification of the wild type and mutant expressions required the development of a new innovative strategy for the expression and purification of 3-OST-1. This approach required the development of the protein expression, cell lysis and protein purification protocols to overcome the poor protein solubility. Initially the charged and hydrophobic nature of the protein of interest made obtaining pure soluble protein difficult. The methods included in this study used the nature of the protein of interest to obtain pure soluble protein. This was in spite of the nature of the protein of interest being the initial cause of the poor protein solubility.
3. Finally the characterization of the ligand interaction with the wild type and mutants required development. In particular the ELISA format required a lot of work due to the charged and hydrophobic nature of the protein of interest. In addition the SPR work required optimization to overcome the fast on and off rates of the binding interaction so that kinetics data could be generated.

# **Chapter 4: Conclusions and Future Perspectives**

## **Conclusions**

In developing a reagent to not only bind to 3-OST-1 modified heparin but with greater affinity relative to the wildtype one could conclude a number of important factors. The strength of the positive charge and the length of the side chain of the amino acid substituted at position 90 are of particular importance. The carboxyl group of Glu-90 was theorized to have a repulsive effect on the binding of the wild type protein to a 3-OST-1 modified heparin molecule. The affinity of the wild type when compared to that of the E90Q mutant would suggest that this is the case. The results of the SPR experiments show an improvement in the affinity relative to the wild type for Arixtra with the E90H suggesting that side chain size may play a role in the increased affinity. The affinity of the E90R mutant relative to E90K may be due lysine being a weaker base and having a shorter side chain compared to E90R. This would suggest that amino acid side chain length in conjunction with positive charge appears to have dramatic effect on binding to Arixtra when considering the affinity of the E90K mutant.

The findings of the SPR experiments provide insight into substrate recognition in that the wild type 3-OST-1 is capable of binding to the 3-OST-1 sulphated product. In addition the results of the SPR experiments give insight into the interaction 3-O-sulphated heparin with positively charged amino acids substituted at site of the catalytic nucleophile.

## **Future Perspectives**

There are three main avenues of research that may be explored in future perspectives:

- Affinity purification of heparin.
- Heparin array screening.
- Computationally guided saturation mutagenesis of 3-OST-1.

### **Affinity column purification of heparin:**

Currently, in regards to drug development there is a lot of focus on the development of Low molecular weight heparin (LMWH). The production of LMWH is typically achieved through depolymerisation of unfractionated heparin. Methods like peroxidative cleavage, nitrous acid cleavage or chemical  $\beta$ -elimination process the unfractionated heparin into smaller lower molecular weights. This approach is harsh and produces unwanted artefacts (Higashi, Hosoyama et al. 2012). Enzymatic depolymerisation can be achieved through the use of heparin lyases. A drawback of this method is the cleavage of heparin within the antithrombin pentasaccharide sequence, thus hindering antithrombin activity (Xiao, Tappen et al. 2010).

Alternative methods of developing LMWH are being developed. Interestingly Li et al have developed an enzymatic ultrafiltration method of producing LMWH using heparin lyases and ultrafiltration. Using an enzymatic ultrafiltration reactor, unfractionated heparin was treated using heparin lyase II and was then filtered through a 10KDa membrane. The fractions were processed through size-exclusion chromatography and polyacrylamide electrophoresis and were analysed using NMR and disaccharide compositional analysis by LS-MS. The LMWH produced at an 80% yield was analysed by SPR for antithrombin binding affinity and was found to be comparable to that of Enoxaparin (Fu, Zhang et al. 2014).

The chemoenzymatic synthesis of heparin is an emerging method of producing biologically active LMWH. The method begins with heparosan, a polysaccharide found in *E. coli*. Through modification of the nitrous acid degraded starting structure to an elongated hexasaccharide. This construct is further modified to produce a construct

with GlcNAc and GlcNS at the nonreducing end of the respective secondary constructs. These constructs are then modified using heparin modifying enzymes to yield ultra-low molecular weight heparin (ULMWH) of 45 and 37% overall yield (Xu, Masuko et al. 2011).

Methods developed by Li et al and Xu et al could potentially benefit from an affinity purification step to enrich the heparin developed that contains the GlcUA-GlcNS3S6S disaccharide that is so important for antithrombin activity (Atha, Lormeau et al. 1985).

However the difficulty in studying heparin is in its microheterogeneity, 3-OST-1 however has evolved to recognise a specific structure in heparin (a glucosamine with a glucuronic acid in the non-reducing end) and the modification it provides is essential to anticoagulant activity. In addition SPR has shown that 3-OST-1 does not appear to interact with chondroitin sulphate C or hyaluronic acid, the binding of chondroitin sulphate A and dermatan sulphate was shown to much lower than that of heparin (Muñoz, Xu et al. 2006). The specificity of 3-OST-1 therefore makes it an interesting target.

Of the mutants generated of 3-OST-1, E90K provides a potential opportunity to be developed as an affinity purification reagent to target the GlcUA-GlcNS3S6S structure. There are GAG affinity purification methods available, for example Suto et al. have taken an alternative approach to the purification of GAGs. In this approach a hyaluronidase column was constructed and was used to purify newly synthesized oligosaccharides through tandem purification using HPLC (Suto, Kakizaki et al. 2014). In another approach, Hsieh et al have reported a method of GAG purification using poly-arginine coated nanodiamonds with analysis by MALDI-TOF mass spectrometry (Hsieh, Guo et al. 2013). Both methods employ affinity based interactions to process polysaccharides prior to downstream analysis. Heparin research would benefit from a similar approach.

In developing affinity purification systems there are a number of different formats for protein immobilization. For example, proteins can be immobilized on silicon nanowire through the interaction between the His-tag proteins and nickel ions chelated to

nitrilotriacetic acid coatings on a silicon nanowire. This allows for the immobilization of the protein with a defined orientation (Yi-Chi, Nathalie et al. 2010). This method has potential usefulness as a protein immobilization method and the his-tag provides a means of affinity purification. In addition, as in the study by Yi-Chi et al, the his-tag could be used to immobilize the E90K mutant in a defined orientation. There is also therefore the potential that the method could be adapted for use in affinity purification of heparin. The drawback of this method is the non-covalent method of protein immobilization which may not be suitable for wash steps necessary heparin for affinity purification.

In another approach a GST tag could be applied as an affinity purification system for heparin (Moon, Xu et al. 2012). Using the E90K mutant a GST tag could be engineered to define the orientation of the mutant and a glutathione column could be used as the solid surface for the reversible immobilization of the mutant. Heparin fractions could then potentially be purified through an affinity interaction with the immobilized mutant. However how would the heparin be recovered and how would it interact with the surface of the affinity column need to be considered. In addition the glutathione column may negatively interact with the heparin applied to the column, which in turn could affect the affinity purification of heparin.

In addition to the use of a his-tag or a GST tag, the use of a biotin tag could be employed for protein immobilization with a defined orientation (Viens, Harper et al. 2008). Engineering a biotin tag into the E90K mutant would provide an efficient means to immobilize the mutant to a streptavidin coated surface. The benefit of this method is the range of surfaces that can be coated with streptavidin. Additionally this method has potential to could be adapted for use in affinity purification of heparin.

In developing a heparin affinity purification method as a proof of concept, the affinity purified heparin fractions could be analysed by SPR using a solution competition assay with antithrombin. The eluted fractions from the affinity purification could be compared to the non-purified sample. Improvements in the  $IC_{50}$  values relative to the starting material would therefore suggest the success of the affinity purification. Another alternative would be to test the eluted fractions using a kit like the

Kinetichrome™ Heparin Anti-IIa kit. Should this approach be successful it would provide an interesting platform for further development.

**Computationally guided saturation mutagenesis:**

In using E90K as an affinity purification reagent there are other factors to consider such as the affinity of E90K for heparin of different sizes. The experiments carried out in this project involved using a pentasaccharide. Defining the affinity of the mutant for larger 3-OST-1 modified heparin structures would take into account the interaction of a larger 3-OST-1 modified heparin structure with important residues further along the protein active site (Moon, Xu et al. 2012). The use of Arixtra does not take into account the effect on mutant heparin affinity on incorporating the whole of the protein active site. Residues that are not accommodated by a pentasaccharide but may play a role in the affinity of the mutant for larger heparin structures should be investigated.

Another important aspect that needs further investigation are the residues that are involved in substrate dissociation. This would increase the understanding of the enzyme substrate interactions and defining the residues involved in dissociation would be useful for the potential affinity purification application of the project. The kinetic analysis performed through SPR show that the mutants have a very fast dissociation rate which may negatively impact an affinity purification application of the mutant.

Investigating the important residues in the protein active site and defining their roles can provide potentially useful information. To investigate this, molecular dynamic simulations could be carried out to potentially identify the residues involved in substrate dissociation. The computational study could provide potential targets for the development of a saturation mutagenesis library with the goal of reducing the dissociation rate while maintaining or enhancing the substrate affinity.

**Heparin array screening:**

Screening mutants against defined heparin molecules could potentially give greater understanding in relation to what size and composition of heparin the mutants bind to. Such screening results could potentially be very important for the development of an affinity purification method for heparin. In addition it could potentially provide understanding of the substrate recognition mechanism and could potentially provide information of substrate binding.

Gama et al reported the development of a chondroitin sulphate array in which structurally defined oligosaccharides were immobilized onto the array surface, the immobilization of the oligosaccharides was defined as was the sulphation pattern to give four different constructs. The array was used to investigate the binding of the growth factors midkine, brain-derived neurotrophic factor and fibroblast growth factor were the sulphation pattern influenced the binding of the growth factors (Gama, Tully et al. 2006).

Schwörer et al have reported the development of a heparin library of defined length and sulphation. The library was used to test the binding to BACE1 protease which is responsible for generating A $\beta$  peptides that accumulate in Alzheimer's disease. The data generated revealed that octa- and decasaccharides containing either GlcNAc6S-GlucA2S or GlcNAc6S-IdoA2S as lead inhibitors of BACE1 protease (Schwörer, Zubkova et al. 2013). Noti et al have reported the development of a heparin microarray to study the interactions of fibroblast growth factor 1, 2 and 4. The array consisted of heparin structures of varying sizes and sulphation patterns (Noti, de Paz et al. 2006).

Research has shown the benefits of using the array platform, its use in combination with SPR could potentially provide useful information such as on the substrate recognition mechanism and on the further applications of the E90K mutant.

# Appendix (A) Project Methods

## **Protein production in E. coli:**

### **General points to consider:**

Protein expression can be carried out in a number of different systems from cell-free, mammalian, yeast, insect to bacterial cell expression (Cuozzo and Soutter 2014).

Bacterial expression is the most prevalent means of protein expression with bacteria like *Bacillus subtilis* being used for protein expression (Vavrová, Muchová et al. 2010).

However the most common choice is *Escherichia coli* and there are a number of different points to consider when expressing protein in *E. coli* such as the strain used for expression.

### **Expression strains:**

The most common strains used in protein production are the BL21 and K12 strains (Phillips, VanBogelen et al. 1984). There are a number of *E. coli* strains available that provide key features for the expression of heterologous proteins. *E. coli* CodonPlus-RIL is a BL21 derived strain that enhances the expression of eukaryotic proteins that contain rare codons. Similarly the C41 strain also derived from BL21 is suitable for the expression of membrane proteins. The JM83 strain is a derivative of K-12 is used for the secretion of recombinant proteins into the periplasm. There are many different strains available that potentially overcome some of the problems associated with protein expression such as poor protein expression due to codon bias. The CodonPlus RIP, RP or RIPL strains provide the rare tRNAs necessary for the expression of proteins poorly expressed due to codon bias.

### **Expression vectors:**

In regards to protein expression in *E. coli* there are different plasmids available that are present at 15-60 (pMB1) or a few hundred copies per cell (pUC series) into which the DNA for the protein of interest can be cloned into. Maintenance of such plasmids

without selective pressure leads to plasmid loss especially in the case of genes that are toxic to the host (Summers 1998).

To overcome plasmid loss a number of strategies can be employed, supplementation of the culture media with appropriate antibiotics is common laboratory practice. There are however drawbacks to this approach due to loss of selective pressure upon plasmid degradation or inactivation. Alternatively genetic tools can be employed to cause cell death upon plasmid loss, plasmid R1 contains the *hok/sok (parB)* locus used for this purpose. Hok is a membrane damaging protein that is encoded on a translationally inactive transcript. Sok is an unstable antisense RNA that binds to the *hok* mRNA leader region. Sok decays rapidly and upon plasmid loss the *hok* transcript becomes active leading to cell death (Gerdes 1988).

A reduction in growth rate prior to induction of protein expression leads to reduced cell energy requirements of the host cell. This in turn leads to a reduction in metabolic stress and an improvement in plasmid stability. Cell cultures allowed to go into the stationary phase can result in the induction of proteases, additionally saturated cultures can lead to plasmid instability due to basal leakage when using the T7 expression system (Studier 2005).

The pET 28 a-c vector series is useful for the controlled expression of recombinant proteins. These vectors carry N- and C- terminal His-tags and can be employed in the expression of membrane proteins (Leviatan, Sawada et al. 2010).

### **Expression promoters:**

A common method of inducing protein expression is through the use of the T7 expression system under the control of T7 RNA polymerase (Studier 1991). Through the use of the *lacUV5* promoter present in the host chromosome, T7 RNA polymerase expression is controlled in *E. coli* strains like BL21. In expression vectors under the control of a T7 promoter, the T7 RNA polymerase recognises the T7 promoter upstream of the coding sequence of the target protein and protein expression is induced using an inducer like IPTG.

One of the limitations of the T7 expression system is the low levels of basal expression that can occur in the absence of an added inducer that can lead to the expression of target protein and host proteins leading to instability or mutations (Studier 1991).

The *T7lac* promoter can be used to reduce basal expression, the presence of the *lac* operator downstream of the T7 promoter allows for the *lac* repressor to bind and interfere with the binding of the T7 RNA polymerase, thus reducing basal expression. Induction of expression unblocks the *lacUV5* promoter and *T7lac* promoter sites allowing for targeted expression of the desired protein (Lopez 1998). In addition to the *T7lac* promoter there are other promoters available, the *trc* and *tac* promoters are induced in the presence of IPTG to give moderately high levels of protein expression (Brosius, Erfle et al. 1985). The *lac* promoter provides low to medium level protein expression and is induced in the presence of IPTG (Gronenborn 1976). The *tetA* promoter provides medium to high levels of protein expression in the presence of anhydrotetracycline giving tight regulation of expression independent of the metabolic state (Skerra 1994). The *araBAD* promoter with the addition of L-arabinose provides low to high levels of protein expression with low basal activity (Guzman et al, 1995). The *rhaP<sub>BAD</sub>* promoter under the control of L-rhamnose gives low to high levels of protein expression (Haldimann 1998).

Regions that flank promoters provide key roles in determining transcription efficiency, upstream elements of certain promoters are A+T rich which in turn interact with the  $\alpha$  subunit of RNA polymerase leading to increased transcriptional (Aiyar 1998).

### **Antibiotic resistance markers:**

The addition of antibiotics to cell cultures allow for the controlled growth of bacterial cultures when the appropriate antibiotic resistance marker is used. This is dependent on the expression vector used and whether or not the expression strain contains functional co-plasmids. There are a number of different antibiotics employed that typically function in causing cell death in cells that do not contain the desired expression vector.

Kanamycin functions through the induction of mistranslation during prokaryotic protein synthesis and the inhibition of translocation (Misumi and Tanaka 1980). Streptomycin functions in a similar manner in that it inhibits protein synthesis (Sharma, Cukras et al. 2007).

Studier *et al* has shown that increased resistance to kanamycin can occur in cell cultures that contain expression vectors that confer kanamycin resistance. This was potentially due to high concentrations of phosphates in rich media. The competent cells BL21 (DE3) were shown to grow in TB media (89mM phosphate) containing 50 µg/ml of kanamycin, cells were effectively killed at 200µg/ml of kanamycin. The same cells were shown not to grow in ZYB media (10 g/L NZ-amine, 5 g/L yeast extract, 5 g/L NaCl) (no added phosphate) with 25µg/ml of kanamycin (Studier 1991).

### **Potential problems to overcome:**

In protein expression there are a number of problems that can arise that may lead to reduced soluble protein yield to no detectible protein expression. There may be numerous causes to poor protein expression from inappropriate expression conditions to inclusion body formation.

### **Inclusion body formation:**

Inclusion bodies are intracellular deposits of misfolded protein aggregates that often form as a result of over-expression protein in *E.coli*. There are a number of approaches used to reduce the formation of inclusion bodies and to increase the expression of soluble protein expression.

As a consequence of over expression in *E. coli* protein can be distributed between the soluble fraction and the insoluble fraction as floccule-type inclusion bodies (Hart, Kallio et al. 1994). Umetsu *et al* have shown using FT-IR spectroscopy that inclusion bodies can be analysed for protein secondary structures. By analysing the relative amount of secondary structures in the FT-IR spectra the relative amount of secondary structures can be compared to the native structure. The group also showed that β2 microglobulin

can be solubilized from inclusion bodies in the presence of 2M L-arginine at room temperature. The effect is thought to be brought about through destabilizing hydrophobic interactions between native structures and non-native structures found in inclusion bodies (Umetsu, Tsumoto et al. 2005).

Green fluorescent protein expressed in *E. coli* which is commonly found in the insoluble fraction has been shown can be solubilized from the insoluble fraction in the presence of 2M arginine without denaturing the protein in question (Tsumoto, Umetsu et al. 2003) .

### **Obtaining soluble protein from inclusion bodies:**

Altering the concentration of inducer, reduction of induction temperature, changing expression cell lines and the use of chaperones are just some methods employed to increase the expression of soluble protein expression. Solubilization of inclusion bodies using denaturing agents and subsequent refolding is another method to increase the soluble fraction obtained from a cell culture. The concern arises with this method is that the protein of interested refolds correctly. Groups however have reported that biologically active proteins have been found inside inclusion bodies (García-Fruitós, Carrió et al. 2005) (Peternel, Jevševar et al. 2008). Using Fourier transform infrared spectroscopy biologically active protein can be detected within inclusion bodies (Ami, Natalello et al. 2006).

### **Protein folding and aggregation:**

High levels of eukaryotic protein production in *E. coli* can be limited due to a number of factors from difficulties associated with protein folding and mRNA translational efficiency to protein toxicity and protein degradation (Jana and Deb 2005). Through the use of osmolytes or heat shock proteins protein folding has been found to be enhanced (Kempf and Bremer 1998).

In addition native bacterial chaperones have been co-expressed or as fusion partners with target proteins to help prevent protein aggregation (Y., Jinmei. et al. 2003).

Overexpressed proteins produced in *E.coli* typically accumulate in the cytoplasm or the periplasmic space.

### **Improving soluble protein expression yield:**

Three are other potential solutions to poor protein expressions that can be employed to improve expression such as lowering induction temperature and inducer concentration. An alternative to protein expression in the cytoplasm is to express in the periplasm (Georgiou 2005). Expression in the periplasm facilitates correct folding and the formation of disulphide bonds in an environment that is less reducing than that of the cytoplasm. Through disulphide binding proteins like DsbC and DsbD the oxidation process is catalysed in the periplasm (Shokri, Sandén et al. 2003).

Alternatively thioreductase-deficient and glutathione reductase-deficient strains provide a means of overexpression proteins with disulphide bonds in the cytoplasm (Ritz, Lim et al. 2001).

Expression of soluble protein can be influenced by the hydrophobicity of the sequence and the sequence complexity. DePristo *et al* studied 95 mammalian proteins and it was discovered that adjoining hydrophobic residues and low complexity regions were linked to poor expression of soluble proteins (DePristo, Zilversmit et al. 2006).

### **Preventing protein aggregation:**

Through the stages of expression, purification or storage proteins aggregation can occur which can lead to issues with protein folding and loss of activity (Goldberg, R. et al. 1991). The addition of co-solvents can aid in protein folding and stability, charged co-solvents can interfere with protein electrostatic interactions (Timasheff, 1998). Chaotropic agents can be used to prevent intermolecular interactions through their interactions with peptide groups (Edwin *et al*, 2002). Kosmotropes like MgSO<sub>4</sub> can be used to stabilize intramolecular protein interactions through competition with intermolecular interactions that lead to protein aggregation (Neagu *et al*, 2001). Sugars and polyhydric alcohols optimize the number of strong water co-solvent interaction and force the co-solvent to be excluded from the protein surface. This leads to the

protein interacting more favourably with water thus stabilizing the state with the smallest surface area and a reducing protein aggregation (Timasheff 1998).

Polyhydric alcohols like poly ethylene glycol have been shown to aid in protein folding and to stabilize protein conformations. Proline is thought to form amphiphilic supramolecular assemblies allowing proline to act as a protein folding chaperone and to prevent protein aggregation by binding to folding intermediates (Samuel, Ganesh et al. 2000). Peternel *et al*, have shown that G-CSF inclusion bodies prepared at 25°C were found to be soluble in 0.2% N-lauoryl sacrcosine with almost 97% of the protein reported as being recovered (Peternel, Jevševar et al. 2008).

## **Generation of mutants:**

### **Plasmid Preparation:**

Plasmid preparations are typically carried out by any number of commercially available kits like those supplied by Qiagen. In the QIA Prep Spin Mini prep kit a number of buffers are provided. Cell cultures containing the desired plasmid DNA can be spun down and resuspended in buffer P1. Buffer P1 contains Tris, EDTA and RNAase A. EDTA protects the DNA from degradation by DNAses and the presence of EDTA weakens the cell envelope. The cells are lysed in the presence of the buffer P2 which contains SDS to solubilise the phospholipids and the proteins of the cell wall and in conjunction with sodium hydroxide leads to cell lysis. Buffer N3 neutralizes the cell lysate and adjusts the solution to high salt binding conditions. The result is the precipitation of denatured proteins, cellular debris and SDS. Buffer N3 also contains acetic acid which allows denatured DNA to renature. Buffer PB removes endonucleases and buffer PE removes excess salts, after which plasmid DNA can be eluted with distilled water. The concentration of the isolated plasmid DNA can be determined by Spectrophotometry or Agarose gel electrophoresis.

### **Spectrophotometer:**

Spectrophotometry is a means of quantitatively measuring the absorbance of a solution and can be used to correlate absorbance to monitor cell growth, determine the concentration of DNA or protein in a given sample. Prior to measuring the absorbance of a particular sample a blanking solution is used as a reference against which the absorbance of a particular sample can be calculated and compared to the blank reference spectrum. The absorbance of a solution is calculated using the equation  $\text{absorbance} = -\log(\text{intensity}_{\text{sample}}/\text{intensity}_{\text{blank}})$ . Using the calculated absorbance the Beer Lambert equation can be used to correlate the absorbance to concentration of a particular sample. Depending on the sample being analysed the absorbance of a sample can be calculated using specific wave length. The absorbance of a nucleic acid sample is calculated at an absorbance of 260-280nm. Samples

containing proteins are measured at an absorbance of 280nm and cell cultures can be monitored at an absorbance of 600nm.

When measuring protein concentrations the extinction co-efficient of the protein must be taken into account. The extinction co-efficient is related to the number of tyrosine, tryptophan and cysteine residues that make of the amino acid sequence of a protein being measured.

### **PCR and Site-Directed Mutagenesis:**

Polymerase Chain Reaction (PCR) is a method most commonly used in the amplification of target DNA. There are numerous applications of this laboratory technique from cloning to site directed mutagenesis. The process begins with a denaturation stage where the template DNA is denatured through heating a reaction mixture containing double stranded template DNA, primer DNA (single stranded DNA sequence to be amplified), nucleotides and DNA polymerase to around 95°C. Exposing the template DNA to such temperatures breaks the hydrogen bonds between the nucleotides, converting double stranded template DNA into single stranded DNA. The annealing stage follows this step with the lowering of the temperature of the reaction mixture to around 55°C. This allows the primer DNA to anneal to its complementary sequence in the template DNA. The final stage is the extension stage where the temperature is increased to around 72°C. At this extension stage a new double stranded DNA molecule is synthesised beginning at the 3' end of the primer DNA strand. DNA polymerase generates a new DNA strand by extending the primer DNA sequence in a 5' to 3' direction through the addition of complementary nucleotides. Over a number of cycles the steps are repeated in order to generate multiple copies of DNA containing the target sequence.

There are a number of DNA polymerases available such as Platinum *Pfx* DNA polymerase which is derived from *Thermococcus kodakaraensis*, *Pfu* DNA polymerase and *Pfu turbo* DNA polymerase derived from *Pyranose* species are utilized in the process of PCR. A thermocycler can be used to regulate and control the temperature

cycles required. For example a Veriti 96-well thermocycler includes six independent temperature blocks allowing for precise control over the temperature cycles.

Site-Directed Mutagenesis (SDM) is a process by which specific mutations can be introduced into a template DNA for the purpose of studying protein structure relationships and gene expression. The process utilizes PCR thermo-cycling to denature the DNA, anneal the primers and extend the mutant DNA. In designing primers for such a purpose a number of factors need to be considered, for example in a SDM application both forward and reverse primers need to contain the desired mutation. A primer sequence must be between 25-45 base pairs in length beyond which there is a greater risk of secondary structure formation. The melting temperature must be greater than equal to 78°C with the desired mutation being in the middle of the primer and the primer having a minimum GC content of 40%. The greater the GC content the higher the denaturation temperature required for the hydrogen bonds of the nucleotides to be broken.

After annealing of the primer sequence to the template DNA, the sequence is extended using DNA polymerase, thus incorporating the desired mutation into the template DNA.

After restriction enzyme digestion to remove parental DNA the product of the SDM reaction can be transformed into a competent cell line. The transformants generated can then be screened using colony PCR. Colony PCR provides a means to screen transformants for the presence of the desired plasmid. The colonies being tested can be added directly to the PCR reaction containing the appropriate primers (Packer, Lim et al. 2013).

### **Restriction Digestion:**

Restriction enzymes can be used to process PCR products or aid in DNA analysis. Restriction enzymes can be classified into three main types. Type 1 cut DNA at sites remote to the recognition sequence, type two restriction enzymes cut at sites within

the recognition sequence and finally type three restriction enzymes cut at sites near the recognition sequence (King and Murray 1994)

Dpn I is a type 2 restriction enzyme that is specific for methylated and hemimethylated DNA. The enzyme recognizes the restriction site 5' Gm6A↓TC 3'. Activity of this enzyme requires N6-methyladenine for cleavage to occur. Proceeding SDM DpnI digestion parental DNA can be targeted using this restriction enzyme as only parental DNA will be methylated. The process of generating the mutant DNA generates a mutant plasmid with staggered nicks which can be treated with DpnI restriction enzyme. Treatment of PCR products prior to transformation with this restriction enzyme greatly improves transformation efficiency.

Other restriction enzymes can be used to aid DNA analysis by agarose electrophoresis. Plasmid DNA can be super coiled plasmid and will travel faster than nicked or linear DNA of the same size. The use of an appropriate restriction enzyme can counteract this problem by rendering all DNA linear. Nco I is a restriction enzyme that recognises the restriction site 5' C<sup>^</sup>CATGG 3' and Eco RI is a restriction enzyme that recognises the restriction site 5' G<sup>^</sup>AATTC 3'. Both restriction enzymes can be used for this purpose.

### **Transformation:**

Transformation involves the use of chemically competent cells which have been treated in order to bring about competency. The process of chemical transformation involves the transfer of DNA into chemically competent cells through the process of heat shock. The introduction of a heat pulse into a medium containing DNA to be transferred and chemically competent cells causes the internalization of foreign DNA in the culture medium. The transformed cells are then grown on agar plates containing an antibiotic of choice to select for transformed cells. There are a number of different chemically competent calls available. E. coli BL21 (RIL) and E. coli BL21 (DE3) (RIL) are chemically competent cell lines chosen for their high level of protein expression. These competent cells encode for genes that encode for tRNA that recognise the arginine codons AGA and AGG, the cells also encode for tRNA that recognises the isoleucine

codon AUA and the leucine codon CUA. The competent cells contain extra genes that encode for tRNA that recognises the arginine codons AGA and AGG and the proline codon CCC, which allows for greater protein expression. Protein expression requiring the T7 promoter can be performed using this competent cell line. Typically protein expression involving recombinant proteins that are non-toxic in *E. coli* express higher levels of protein in this cell line compared to *E. coli* BL21 (DE3) pLysS-T1. *E. coli* BL21 (DE3) pLysS-T1 is a cell line that reduces protein degradation as it lacks the proteolytic activity of lon and ompT proteases. The cell line also contains the plasmid pLysS which inhibits T7 polymerase activity. *E. coli* HB2151 is a strain of chemically competent cells that allow for the recognition of the amber stop codon which can be used for the expression of single chain antibodies.

### **Agarose Gel Electrophoresis:**

Agarose gel electrophoresis (AGE) is a laboratory technique commonly used for size determination of DNA samples. Sample DNA being measured is passed through a highly cross-linked matrix aided through the use of an electric current. The negatively charged DNA sample being measured migrates towards the positive charged end of an AGE apparatus, separating according to size.

There are a number of factors effecting DNA migration through a gel matrix from size to conformation the ionic strength of the running buffer. The smaller the DNA samples the faster the migration through the gel matrix. Typically a 1% agarose gel is sufficient to resolve DNA up to 30kb. The conformation of the DNA sample also effects the migration of the DNA. In the case of plasmid DNA a super coiled plasmid will travel faster than nicked or linear DNA of the same size.

Essential to the utilization of AGE in determining the size of a DNA sample is the inclusion of a DNA ladder. For example, a 1kb DNA ladder provided by Fisher Bioreagents contains DNA fragments ranging from 300 to 10000 bp to which a DNA sample can be compared in determining the size of the sample. Another essential component is the inclusion of a DNA staining dye. SYBR safe DNA gel stain is an example that can be used to make DNA detectible using a UV transilluminator and is

beneficial in that it has a reduced mutagenicity compared to ethidium bromide. The addition of DNA loading dye is primarily used for the visual tracking of DNA migration on an agarose gel. The can contain dyes such as bromophenol blue or xylene cyanol for visualizing the DNA bands. Glycerol is typically added to aid in the loading of the dyes. EDTA is added typically to inhibit metal dependant nucleases which can degrade the DNA. The most common two loading and running buffers are Tris-acetate-EDTA (TAE) and Tris-borate-EDTA (TBE). TBE is typically used due to its greater conductive capacity than TAE, however TAE separates larger DNA fragments better than TBE. The addition of EDTA is to reduce intermolecular interactions.

## **Protein purification:**

### **Sonication:**

Sonication is a process by which cell cultures can be disintegrated prior to beginning the processes like protein purification as such experiments require the disruption of cell walls. There are a number of sonicators on the market for laboratory use such as the MSE Soniprep 150 Plus sonicators. In principal as with other sonicators this apparatus works through the transmission of high frequency vibrations in the form of sound waves produced by a piezoelectric transducer which pass through a cell culture medium. Gas bubbles are formed as a result of the alternating positive and negative pressures created by the sound waves. The gas bubbles disintegrate during the refraction phase caused by the sound wave. The pressure gradient generated results in the disruption of the cell walls of the cell culture medium through which the sound waves pass through. The resulting lysate can be spun down in a centrifuge and the supernatant used for further experiments such as protein purification.

### **Immobilized metal ion affinity chromatography:**

Immobilized metal affinity chromatography was developed to take advantage of the known affinity of transition metals for histidine. The DNA encoding for a protein of interest can be modified to contain an oligohistidine polypeptide that binds with high affinity to transition metals thus improving the performance of the IMAC purification.

Nickel column purification involves the use of Metal Ion Affinity Chromatography for the separation of Histidine-tagged proteins. The process involves using the selective affinity of His-tags for Ni<sup>2+</sup> ions. Typically His-tags are made up of six residues which is the optimal length for protein purification. Purification of His-tagged proteins can be carried out in different chromatography media like Ni sepharose. Ni sepharose is made up of highly cross linked agarose beads in which Ni<sup>2+</sup> ions have been chelated to the bead surface. For the purpose of purification a number of things are to be considered in preparing the media for protein purification. An alkaline to neutral pH of 7-8 is

typically used in the presence of low molar salts to prevent non-specific protein interactions. His-tagged proteins are typically eluted through the use of imidazole which competes with His-tags for Ni binding. Initial use of low concentrations of imidazole allows for controlled protein purification, with higher concentrations of imidazole being used for the elution of surface bound proteins. Ni ions are typically used for the purification of His tagged proteins. However  $\text{Co}^{2+}$  can also be used, although the binding strength is lower.

Automated purification can also be achieved through the use of systems such as an AKTA purifier in which the UV absorbance, pH, temperature and conductivity can be monitored. The conductivity of a solution is proportional to the solution's ion concentration therefore in the monitoring conductivity of a solution being purified measuring conductivity also provides a means of verification that the Ni ions essential to protein purification remain bound to the column. Temperature is also a useful factor to monitor in protein purification as conductivity is proportional to temperature. Monitoring the UV absorbance allows for greater control of protein column binding and elution as changes in UV levels correspond to protein binding and elution. For example after a protein has bound to a column elution of target protein using an elution buffer corresponds to a spike in UV absorbance readings. Depending on the scale of protein purification, small scale purification can be achieved using commercially available columns loaded with Ni Sepharose. Alternatively larger scale purifications can be carried out using His Trap FF columns which are pre-packed with Ni Sepharose in conjunction with an automated protein purification system such as an AKTA purifier.

#### **Co-purification of contaminants:**

When using an E. coli protein expression system co-purification of native E. coli proteins can occur especially when the native proteins are expressed at high levels (Cai, Moore et al. 2004). Ferric uptake regulator is a DNA-binding protein that controls the quantities of intracellular iron which can be over expressed in response to acid-shock stress and in turn it can be found to be co-purified during IMAC (Escobar, Pérez-Martín et al. 2000). YodA is a metal binding protein that plays a role in regulating

cadmium in the cells and can be over expressed in response to oxidative and acid stress where it can be co-purified during IMAC (Birch, O'Byrne et al. 2003). The metalloenzyme Cu/Zn-superoxide demutase protects the cell from oxygen toxicity but can be overexpressed under stress conditions leading to it being co-purified during IMAC (Battistoni, Folcarelli et al. 1996).

### **Hydrophobic interaction chromatography:**

Hydrophobic interaction chromatography is a method of protein purification that involves purification through the interactions of hydrophobic regions present in proteins and hydrophobic ligands. This interaction involves the exposure of hydrophobic regions to a polar solvent leading to a net increase in the entropy of the environment when the protein is introduced to the new environment. Essential to this purification method is the polarity of the solvent. The addition of anions to the solvent can increase the hydrophobicity. Potassium phosphate buffers provide greater hydrophobicity than that of sodium chloride (Påhlman, Rosengren et al. 1977). Hydrophobic interactions can also be weakened through the removal of lyotropes like potassium phosphate or through the introduction of organic solvents like ethanol (Fausnaugh and Regnier 1986).

The choice of hydrophobic ligands is very important to successful purification. There are a number of different media types available from phenyl, butyl, octyl, ether and isopropyl. Depending of the strength of the anion used and the hydrophobic nature of the protein of interest, different media types and anions need to be used to obtain effective purification of the target protein.

### **SDS-PAGE:**

Sodium dodecyl sulphate-polyacrylamide gel (SDS-PAGE) is a method of separating protein based on molecular weight. The acrylamide and bisacrylamide component polymerises in the presence of ammonium persulphate catalysed in the presence of TEMED. Increasing the amount of acrylamide decreases the size of the pores formed in

the cross-linked gel matrix. For the purpose of resolving larger proteins a lower percentage of acrylamide is typically used. Protein samples resolve when an electrical current is applied which causes protein samples to migrate according to their molecular weight. The treating of protein samples prior to applying them to the gel using SDS insures that the relative charge densities of the protein samples are equal. The gel matrix through which the protein samples pass through is divided into stacking and resolving gels. The stacking gel concentrates the protein samples into a tight band due to the lower percentage of acrylamide and the lower pH. The resolving gel has a higher concentration of acrylamide and is placed below the stacking gel resulting in a gel matrix of decreasing pore size. Once a gel has been run to completion the gel is stained in order to visualize the protein bands. Staining a gel with a fixing solution is recommended to prevent the diffusion of the protein bands in a gel, this can be achieved through the use of ethanol and acetic acid. A second staining solution is required to visualize the protein bands. This can be achieved through the use of a staining solution containing coomassie blue. Coomassie blue binds to proteins present in the gel through ionic interactions between sulphonic acid groups of the dye and amine groups in the protein. In determining the size of a protein using SDS-PAGE a protein ladder is required for comparison. EZ-run prestained Rec protein ladder for example contains a mixture of 10 highly purified recombinant proteins with molecular weights ranging from 10 to 170 kDa. A loading dye is used in order to visualize the migrating protein bands. Such dyes can contain SDS to equalize the charge density of the protein sample and a dye like bromophenol blue to track protein migration.

### **Western Blotting:**

Western Blotting is a laboratory technique typically used in the determination of the size of a protein. This can be achieved directly or indirectly through antibody detection. The direct detection method involves the use of a primary antibody that is fluorescently labelled to allow for the detection of the protein sample. The indirect method involves the use of primary and secondary antibodies that binds to the primary antibody and an enzyme like alkaline phosphatase conjugated to the secondary enzyme is used for the detection of the protein. To begin the process of performing a western blotting experiment, protein samples are separated using

electrophoresis techniques like SDS-PAGE. Protein samples are transferred to a polyacrylamide gel to a membrane through the use of an electrical current. After which the membrane is treated with a blocking solution to prevent non-specific binding. Proteins are transferred from a polyacrylamide gel through the process of electroforectic transfer. The process involves placing a polyacrylamide gel over a protein binding support membrane. An electrical current is used to transfer protein from a polyacrylamide gel to a protein binding support membrane. Efficient transfer is achieved through the use of low ionic strength buffers and low electrical current. Western blotting transfer buffer can contain things like Tris, glycine, SDS and methanol. Methanol is used to prevent the polyacrylamide gel from swelling during the transfer on the protein to the protein binding support membrane. Glycine is included in the transfer buffer as it has a lower conductivity than the protein which in turn has a lower conductivity than the chlorine ions in the gel. The result is the formation of a boundary which aids in the transfer of the protein to the transfer membrane. Once the protein has been transferred to the transfer membrane a blocking solution like 2% skimmed milk in PBS can be added to prevent non-specific binding. The transfer membrane is typically washed with buffers containing detergents like Tween between the addition of the primary and secondary antibodies to remove non-specifically bound material. The addition of primary antibodies like mouse pentahis antibody can be used to label proteins containing a histadine tag. A secondary antibody like goat anti-mouse antibody containing conjugated alkaline phosphatase will bind to the primary antibody and increase protein detection. A substrate can then be used to visualize the protein, for example BCIP/NBT will bind to alkaline phosphatase conjugated to the secondary antibody allowing the protein bands to be visualized. In determining the size of a protein using Western Blotting a protein ladder is required for comparison. EZ-run prestained Rec protein ladder for example contains a mixture of 10 highly purified recombinant proteins with molecular weights ranging from 10 to 170 kDa.

## **Enzymatic activity:**

### **DEAE Chromatography:**

DEAE chromatography uses the principal of Ion Exchange Chromatography. Ion exchange chromatography (IEX) involves using cation or anion exchange resins, were for example anion-exchange resins will be positively charged and will bind negatively charged anions. IEX separates molecules based on differences in there net surface charge. Each protein has a unique net surface charge in relation to pH. Each protein has a specific isoelectric point above which a protein will bind to a positively charged media and below which a protein will bind to a negatively charged media.

DEAE chromatography is typically used for the separation of proteins, it can be used however for the purification of carbohydrates (Edavettal, Lee et al. 2004). In a study performed by Edavettal *et al*, the sulphotransferase activity of 3-OST-1 was determined by incubating HS in a reaction mixture with among other materials, 3-OST-1 and radioactive PAPS. The sample was then measured using liquid scintillation counting to determine the activity of 3-OST-1 by assaying the modified HS using liquid scintillation counting. Before this could be performed the 3OST-1 and the radioactive PAPS needed to be removed from the reaction mixture. This was achieved using DEAE chromatography. As the catalytic domain of 3-OST-1 has an isoelectric point of 9.3 (as determined by [www.expasy.org](http://www.expasy.org)). 3-OST-1 can be removed from the reaction mixture using anion exchange chromatography. Any unbound material present in the reaction mixture when passed through the DEAE media was also discarded.

Increasing the ionic strength of the buffers by adding higher molar salts allowed for the elution of the bound HS, as  $\text{Na}^+$  and  $\text{Cl}^-$  compete with HS for the charged surface. Using such buffers therefore allowed for the elution of the HS for further analysis by liquid scintillation counting.

### **Liquid Scintillation Counting:**

Liquid Scintillation Counting (LSC) is a method of measuring fluorescent emissions derived from the interaction of  $\beta$ -particles emitted from radioactive decay and the solvent constituent of scintillation fluid. Scintillation fluid absorbs the energy emitted by radioactive decay. This absorbed energy is captured by phosphors which release this energy in the form of light. There is a direct relationship between light emitted and the radioactive decay. Typically aromatic organic compounds are used as the solvent constituent of scintillation fluid. For example toluene captures the energy emitted by  $\beta$ -particles interacting with the aromatic rings. The captured energy gets transferred between the aromatic rings of the solvent which consequently are captured by phosphors dissolved in the scintillation fluid. The phosphors react resulting in the emission of light, with the number of emissions per second being proportional to the number of radioactive emissions. In a liquid scintillation counter the light emitted is collected into three channels or counting windows. The lowest channel corresponds to  $^3\text{H}$  emissions and the highest channel corresponds to  $^{32}\text{P}$  emission and  $^{35}\text{S}$  emissions being detectable in all channels. The emissions of light are expressed as counts per minute averages. LSC can be used to measure the sulphotransferase activity of 3-OST-1. Using a form of PAPS that contains the radioactive isotope  $^{35}\text{S}$ , activity can be measured upon transferring the sulphate ( $\text{SO}_3$ ) containing the  $^{35}\text{S}$  to HS. The count per minute average determined by liquid scintillation analysis of HS will determine the activity of 3-OST-1 by way of measuring the number of radioactive isotopes through their proportional emission of  $\beta$ -particles that are read as emissions of light. These emissions of light are proportional to the amount of  $^{35}\text{S}$  isotope present, which is consequently proportional to the amount of sulphuryl groups (containing the  $^{35}\text{S}$  isotope) transferred to HS by 3-OST-1. Thus in this way LSC can be used to measure the sulphotransferase activity of 3-OST-1.

## **Mutant characterization:**

### **SPR:**

Surface Plasmon Resonance (SPR) is a method that can be used for studying binding interaction between a ligand and an analyte through the use of surface plasmon resonance to measure the association and dissociation rates in affinity and kinetics studies.

SPR occurs when a plane of polarized light hits a metal film under total internal reflection conditions, leading to the formation of surface plasmons (See Figure 42). Total internal reflection (TIR) occurs when a ray of light hits a boundary at an angle greater than the critical angle (the critical angle is the angle above which total internal reflection occurs) with respect to the normal (a vector perpendicular to the boundary). When light passes at such an angle and the refractive index is lower than the other side of the boundary all the light is reflected. In SPR under TIR, the photons generated by light hitting the sensor surface create an electrical field that extends beyond the sensor chip surface. This electrical field will interact with the electrons of the gold coated on the surface of the SPR chip. The photons hitting the sensor chip surface will transfer the energy to these electrons to form surface plasmons. Surface plasmons can be defined as electrons that oscillate at the interface between two materials. Under TIR the photons will generate an electrical field on the opposite side of the SPR chip surface and the plasmons generate a comparable field extends into the medium on either side of the SPR chip surface (called the evanescent wave). This allows for a measure in the change in surface plasmon resonance. Resonance occurs when the momentum of the incoming light is equal to the momentum of the plasmons. The velocity of the plasmons changes when composition of the medium changes, this results in a change in the angle of incident light in which resonance occurs. A resonance unit is a measure of this angle shift and can be used as a measure of analytes binding to ligands immobilized on the sensor chip surface (Di Primo and Lebars 2007).

In performing SPR experiments the ligand can be immobilized in a number of ways such as amine, thiol or (strept) avidin coupling. Sensor chips containing carboxyl groups can be modified by the addition of NHS and EDC to which the ligand will bind. Ethanolamine can be used to block the remaining active sites. Amine coupling interaction introduces N-hydroxysuccinimide esters into the sensor chip surface through modification of the carboxymethyl groups with and EDC. These esters react with the amine groups in the ligand to form covalent bonds.

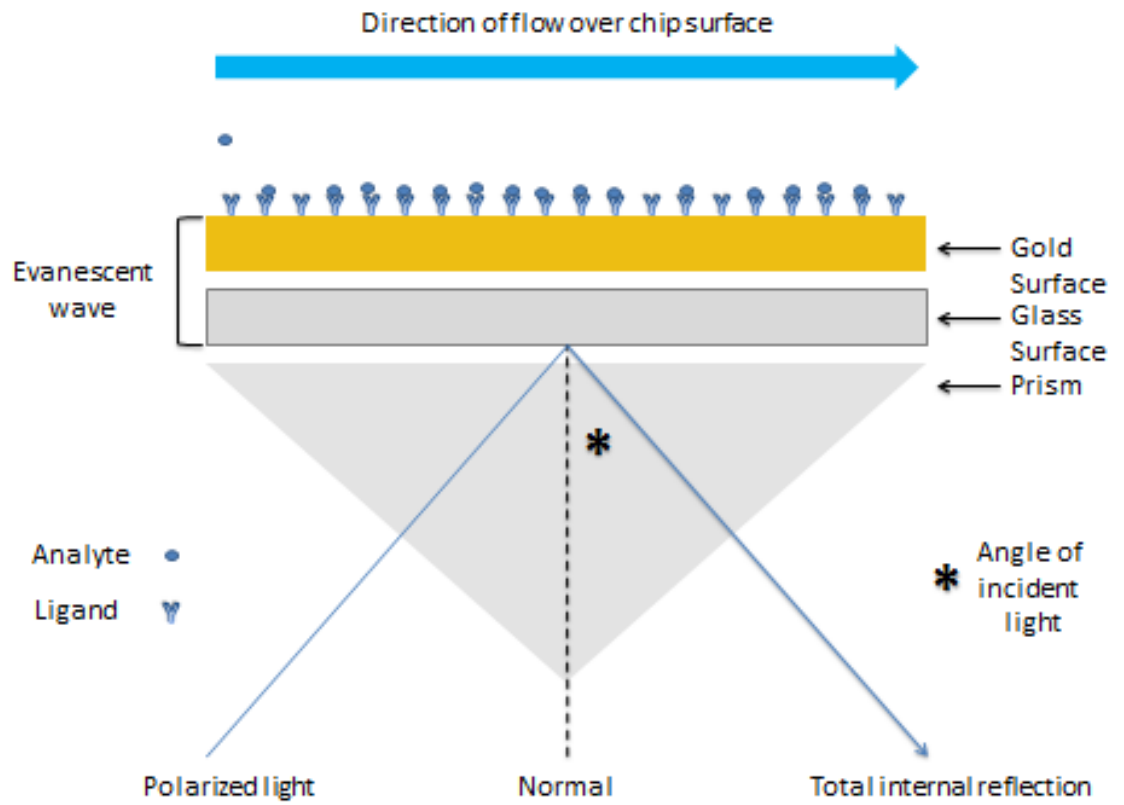
There are factors that can affect the performance of a SPR experiment. In SPR experiments the incident light and temperature are constant as heat can dissipate the evanescent wave. Changes in the SPR signal depend greatly on changes in the refractive index of the medium. Therefore when molecules bind to the sensor chip surface there is a change in the refractive index that results in a change in the resonance angle (Quinn, O'Neill et al. 2000).

SPR binding studies are typically performed as single-step or multi-step binding interactions. Single-step interactions are suitable for measuring concentrations and kinetic assays as the binding interactions are analyzed in a single step. Multi-step methods are useful for analyzing multi-molecular complex formation or incorporating analyte conformational changes in the binding study as this method involves a series of binding interactions. There are also indirect and direct binding interactions where the direct method involves measuring responses from the analyte itself. The indirect method however involves binding of the analyte to a competing molecule followed by determining the amount of free competing molecules remaining. This method is useful in kinetic studies and quantification.

The data generated by SPR is useful in kinetics and affinity studies, particularly in determining the affinity ( $k_D$ ) of an enzyme. Enzymatic affinity describes how strong an enzyme binds to its substrate. The lower the  $k_D$  the greater the affinity and the lower the concentration required to form an enzyme-ligand complex. For example an enzyme with a nM  $k_D$  has a greater affinity than an enzyme with a  $\mu$ M  $k_D$ . The association rate constant describes the rate of complex formation. The dissociation rate constant describes the rate of complex decay. Both the association and

dissociation rates of an enzyme are analysable using SPR. The steady state refers to a state at which association and dissociation are equal and the equilibrium binding response occurs when the same number of complexes forms as breaks down on the sensor surface are also analysable using SPR.  $\chi^2$  measures how closely the model fits the experimental data.

The important things to look out for in the shape a binding curve that the curve should be exponential and not parabolic or hyperbolic. The association phase should not be concave nor should the dissociation phase convex. A square shaped curve results from a fast  $k_d$ . Linear graphs are data sets that are limited by mass transport. Mass transport limitations occur when the binding rate of the analyte to the ligand is faster than its diffusion rate to the surface (Rich and Myszka 2008).



**Figure 42:** Illustration of key features in Surface Plasmon Resonance.

**ELISA:**

Enzyme linked immunosorbent assays are commonly used in molecular biology laboratories. An ELISA can take on three basic systems, a direct ELISA, an indirect ELISA and a sandwich ELISA. Direct ELISA's involve the immobilization of a test antigen to the ELISA plate followed by the addition of enzyme-conjugated antibody. A substrate system is added to detect the bound antibody. The antigen is bound to the plate via passive absorption and after an incubation period, unbound antigen can be washed away. An enzyme conjugated antibody can then be added to the plate that are specific for the antigen and again after incubation, free antibody is washed away. To the plate a substrate can be added that reacts with the enzyme conjugated to the antibody that is bound to the antigen. A colour reaction is allowed to develop which can be stopped by lowering the pH. The colour can be measured by a spectrophotometer allowing the binding of the conjugated antibody to the antigen to be assessed.

Indirect ELISA's differ from direct ELISA's in that antibodies specific for the antigen are themselves recognized by enzyme conjugated antibodies. The binding of the antibody to the antigen can therefore be indirectly measured by measuring the colour change of the substrate binding to the enzyme conjugate using a spectrophotometer. Indirect ELISA's allow for the testing of different antibodies to the antigen while requiring a single enzyme conjugated antibody to detect binding.

Direct sandwich ELISA's are characterized by the immobilization of the antibody to the surface of the plate which in turn capture the antigen. The captured antigen is then detected by an enzyme conjugated antibody, using a substrate the binding can be measured as in the direct and indirect ELISA's using a spectrophotometer. For this system to be used the antigen needs two epitopes for both capture of one antibody and detection by the other.

Indirect sandwich ELISA's entail the immobilization of an antibody to capture the antigen which is recognized by an additional antibody. An enzyme conjugated antibody then binds to the antibody bound to the antigen and then using a substrate the

binding can be measured using a spectrophotometer. This system allows for the use of different antibodies to detect the antigen.

Inhibition ELISA's can be employed to measure the binding of an antigen by its ability to inhibit an established pre-titrated system. A direct inhibition ELISA involves coating an antigen to an ELISA plate to act as a solid phase antigen. An enzyme conjugated antibody is then added in different concentrations of the test antigen to the plate. After incubation and washing a substrate is added to the plate and the colour change of the substrate binding to the enzyme conjugate antibody is measured using a spectrophotometer. Binding of the test antigen to the enzyme conjugated antibody will result in a decrease in the expected absorbance as measured by spectroscopy. The test antigen inhibits with the binding to the enzyme conjugated antibody to the plate bound antigen and competes with binding to the solid phase antigen. An antigen that does not inhibit will result in no change to the expected absorbance.

In an indirect inhibition ELISA the system involves using an antigen coated plate to which a liquid phase antigen is added in the presence of an antibody. The antigen competes with the binding of the antibody to the solid phase antigen. After incubation and wash steps an enzyme conjugated antibody is added to detect binding to the solid phase antigen. After another incubation and wash step a substrate is added and the colour change of the substrate binding to the enzyme conjugate is measured using a spectrophotometer. Inhibition will result in a decrease in the expected absorbance

# Appendix (B): Experimental Conditions

**Table 3:** Sonication parameters.

Percentage cycle	Duration (seconds)	Rest period (seconds)
50	30	30

**Table 4:** Enzyme activity reaction mixture.

Reagent	Volume ( $\mu$ l)
10% Triton X-100	2.5
1 M MES pH 7	2.5
1 M MgCl <sub>2</sub>	0.25
2 M MnCl <sub>2</sub>	0.25
10 mg/ml BSA	0.6
3-OST-1	10
Heparin 10mg/ml	1
1 $\times$ 10 <sup>7</sup> cpm [ <sup>35</sup> S] PAPS in 30 $\mu$ l of water	30
Water	37.5

**Table 5:** Site-directed mutagenesis primers.

Primer Name	Primer Sequence
<b>E90Q Forward Primer</b>	GCT GCA GCT GAA AAC CAA GTC CAT TTC TTT GAC
<b>E90Q Reverse Primer</b>	GTC AAA GAA ATG GAC TTG GTT TTC AGC TGC AGC
<b>E90H Forward Primer</b>	GCT GCA GCT GAA AAC CAT GTC CAT TTC TTT GAC
<b>E90H Reverse Primer</b>	GTC AAA GAA ATG GAC ATG GTT TTC AGC TGC AGC
<b>E90K Forward Primer</b>	GCT GCA GCT GAA AAC AAA GTC CAT TTC TTT GAC
<b>E90K Reverse Primer</b>	GTC AAA GAA ATG GAC TTT GTT TTC AGC TGC AGC
<b>E90R Reverse Primer</b>	GCT GCA GCT GAA AAC CGG GTC CAT TTC TTT GAC
<b>E90R Forward Primer</b>	GTC AAA GAA ATG GAC CCG GTT TTC AGC TGC AGC

**Table 6:** Site-directed mutagenesis primers features.

Primer Name	Melting temperature	Number of base pairs	G+C Content
<b>E90Q Forward Primer</b>	79.7°C	33	45.50%
<b>E90Q Reverse Primer</b>			
<b>E90H Forward Primer</b>	79.7°C	33	45.50%
<b>E90H Reverse Primer</b>			
<b>E90K Forward Primer</b>	78.3°C	33	42.00%
<b>E90K Reverse Primer</b>			
<b>E90R Reverse Primer</b>	82.4°C	33	52.00%
<b>E90R Forward Primer</b>			

The melting temperature was calculated using the following equation:

$$T_m = 81.5 + 0.41(\% \text{ GC}) - (675/N)$$

N = Length of primers

**Table 7:** Site-directed mutagenesis reaction mixture and thermocycling conditions.

(A) Site-directed mutagenesis reaction mixture for E90Q mutant:

Reagents *	Volume (μl)	Concentration
Millicule water	29.3	
2 mM dNTP	7.5	300 μM
1X Amplification buffer	5	
Template DNA	*'	50 ng
50 mM MgSO <sub>4</sub>	1	
Forward primer	1	
Reverse primer	1	
Platinum pfx DNA polymerase	0.4	

- Site-directed mutagenesis reaction mixture for E90H mutant:

Reagents *	Volume (μl)	Concentration
Millicule water	32.1	
Stock dNTP	1	
1X Amplification buffer	5	
1X Enhancer buffer	5	
Template DNA	*'	50 ng
50 mM MgSO <sub>4</sub>	1	
Forward primer	1	
Reverse primer	1	
Platinum pfx DNA polymerase	0.4	

(B) Site-directed mutagenesis reaction mixture for E90K mutant:

Reagents *	Volume (μl)	Concentration
Millicule water	26.3	
Stock dNTP	1	
1X Amplification buffer	5	
1X Enhancer buffer	5	
Template DNA	*'	50 ng
50 mM MgSO <sub>4</sub>	1	
Forward primer	1	
Reverse primer	1	
Platinum pfx DNA polymerase	0.4	

(C) Site-directed mutagenesis reaction mixture for E90R mutant:

Reagents *	Volume (μl)	Concentration
Millicule water	26.0	
Stock dNTP	1	
1X Phusion buffer	5	
5% DMSO	5	
Template DNA	*'	50 ng
50 mM MgSO <sub>4</sub>	1	
Forward primer	1	
Reverse primer	1	
Phusion high fidelity DNA polymerase	0.4	

\* Platinum pfx DNA polymerase kit used for E90Q, E90H and E90H. Phusion high fidelity DNA polymerase kit used for E90R.

\*' Calculated based on plasmid DNA concentration

**Table 8:** Thermocycling conditions for the generation of 3-OST-1 mutants:

The following thermocycling conditions were used in the generation of the 3-OST-1 mutants using the corresponding site-directed mutagenesis reaction mixture (see above).

(A) Thermocycling conditions for the generation of E90Q, E90H and E90K mutants:

Temperature	Time	Number of cycles	Stages
94°C	2 min	1	1
94°C	30 sec	20	2
55°C	30 sec		
68°C	6 min		
68°C	10 min	1	3
4°C	∞	∞	4

(B) Thermocycling conditions for the generation of E90R mutant:

Temperature	Time	Number of cycles	Stages
95°C	5 min	1	1
95°C	30 sec	30	2
55°C	30 sec		
68°C	7 min		
68°C	10 min	1	3
4°C	∞	∞	4

**Table 9:** Colony PCR conditions.(A) Colony PCR reaction mixture:

Reagents *	Volume (µl)	Concentration
Millicule water	32.2	
Stock dNTP	1	
1X Phusion buffer	5	
50 mM MgSO <sub>4</sub>	1	
T7 Forward primer	1	125 ng
T7 Reverse primer	1	125 ng
Phusion high fidelity DNA polymerase	0.5	

(B) Thermocycling conditions for Colony PCR :

Temperature	Time	Number of cycles	Stages
95°C	5 min	1	1
95°C	30 sec	35	2
58°C	1 min		
72°C	1 min		
72°C	10 min	1	3
4°C	∞	∞	4

**Table 10:** SDS-PAGE gel composition.

- Gel composition:

5% Stacking	
Reagents	Volume
40% polyacrylamide/bis-acrylamide	750 µl
1 M Tris pH 6.8	750 µl
10% APS	60 µl
10% SDS	60 µl
TEMED	9 µl
Water	4.38 ml

12% Resolving	
Reagents	Volume
40% polyacrylamide/bis-acrylamide	3.75 ml
1 M Tris pH 6.8	3.12 ml
10% APS	125 µl
10% SDS	125 µl
TEMED	7.5 µl
Water	5.3 ml

# References

- Aikawa, J.-i. and J. D. Esko (1999). "Molecular Cloning and Expression of a Third Member of the Heparan Sulfate/Heparin GlcNAcN-Deacetylase/ N-Sulfotransferase Family." *Journal of Biological Chemistry* 274(5): 2690-2695.
- Aiyar, S., E. Gourse, R. L. Ross, W. (1998). "Upstream A-tracts increase bacterial promoter activity through interactions with the RNA polymerase  $\alpha$  subunit." *Proceedings of the National Academy of Science U S A* 95(22): 14652-14659.
- Ami, D., A. Natalello, et al. (2006). "Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy." *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1764(4): 793-799.
- Aquino, R. S., M. S. Pereira, et al. (2010). "Heparins from porcine and bovine intestinal mucosa: Are they similar drugs?" *Thrombosis and Haemostasis* 103(5): 1005-1015.
- Atha, D. H., J. C. Lormeau, et al. (1985). "Contribution of monosaccharide residues in heparin binding to antithrombin III." *Biochemistry* 24(23): 6723-6729.
- Ayotte, L. and A. S. Perlin (1986). "N.m.r. spectroscopic observations related to the function of sulfate groups in heparin. Calcium binding vs. biological activity." *Carbohydrate Research* 145(2): 267-277.
- Ballerstadt, R. and J. Schultz, S. (2000). "A fluorescence affinity hollow fiber sensor for continuous transdermal glucose monitoring." *Anal Chem.* 72(17): 4185-4192.
- Barth, H., C. Schäfer, et al. (2003). "Cellular Binding of Hepatitis C Virus Envelope Glycoprotein E2 Requires Cell Surface Heparan Sulfate." *Journal of Biological Chemistry* 278(42): 41003-41012.
- Battistoni, A., S. Folcarelli, et al. (1996). "The Cu,Zn superoxide dismutase from *Escherichia coli* retains monomeric structure at high protein concentration. Evidence for altered subunit interaction in all the bacteriocupreins. ." *Biochemistry*:320(3): 713-716.
- Bethea, H. N., D. Xu, et al. (2008). "Redirecting the substrate specificity of heparan sulfate 2-O-sulfotransferase by structurally guided mutagenesis." *Proceedings of the National Academy of Sciences* 105(48): 18724-18729.
- Birch, R. M., C. O'Byrne, et al. (2003). "Enrichment of *Escherichia coli* proteins by column chromatography on reactive dye columns." *PROTEOMICS* 3(5): 764-776.
- Blossom, D. B., A. J. Kallen, et al. (2008). "Outbreak of Adverse Reactions Associated with Contaminated Heparin." *New England Journal of Medicine* 359(25): 2674-2684.
- Borders, C. L., J. A. Broadwater, et al. (1994). "A structural role for arginine in proteins: Multiple hydrogen bonds to backbone carbonyl oxygens." *Protein Science* 3(4): 541-548.
- Brosius, J., M. Erfle, et al. (1985). "Spacing of the -10 and -35 regions in the *tac* promoter. Effect on its *in vivo* activity." *Journal of Biological Chemistry* 260(6): 3539-3541.
- Cadwell, R. C. and G. F. Joyce (1992). "Randomization of genes by PCR mutagenesis." *Genome Research* 2(1): 28-33.
- Cai, Y., M. Moore, et al. (2004). "Genomic data for alternate production strategies. I. Identification of major contaminating species for Cobalt+2 immobilized metal affinity chromatography." *Biotechnology and Bioengineering* 88(1): 77-83.
- Caldwell, E. E. O., V. D. Nadkarni, et al. (1996). "Importance of specific amino acids in protein binding sites for heparin and heparan sulfate." *The International Journal of Biochemistry & Cell Biology* 28(2): 203-216.
- Carfí, A., S. H. Willis, et al. (2001). "Herpes Simplex Virus Glycoprotein D Bound to the Human Receptor HveA." *Molecular Cell* 8(1): 169-179.
- Carlsson, P., J. Presto, et al. (2008). "Heparin/Heparan Sulfate Biosynthesis: PROCESSIVE FORMATION OF N-SULFATED DOMAINS." *Journal of Biological Chemistry* 283(29): 20008-20014.

- Chen, L., K. Ichihara-Tanaka, et al. (2004). "Role of the Carboxyl-Terminal Region in the Activity of N-Acetylglucosamine 6-O-Sulfotransferase-1." *Journal of Biochemistry* 136(5): 659-664.
- Christendat, D., A. Yee, et al. (2000). "Structural proteomics of an archaeon." *Nature Structural Biology*. 7(10): 903-909.
- Chung, C. T. and R. H. Miller (1993). [43] Preparation and storage of competent *Escherichia coli* cells. *Methods in Enzymology*. W. Ray, Academic Press. Volume 218: 621-627.
- Conter, A., C. Gangneux, et al. (2001). "Survival of *Escherichia coli* during long-term starvation: effects of aeration, NaCl, and the *rpoS* and *osmC* gene products." *Research in Microbiology* 152(1): 17-26.
- Crawford, B. E., S. K. Olson, et al. (2001). "Cloning, Golgi Localization, and Enzyme Activity of the Full-length Heparin/Heparan Sulfate-Glucuronic Acid C5-epimerase." *Journal of Biological Chemistry* 276(24): 21538-21543.
- Cuozzo, J. W. and H. H. Soutter (2014). "Overview of Recent Progress in Protein-Expression Technologies for Small-Molecule Screening." *Journal of Biomolecular Screening*: 19(7): 1000-1013.
- DePristo, M., A., M. Zilversmit, M., et al. (2006). "On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins." *Gene* 1(378): 19-30.
- Desai, U. R., H. M. Wang, et al. (1993). "Oligosaccharide Composition of Heparin and Low-Molecular-Weight Heparins by Capillary Electrophoresis." *Analytical Biochemistry* 213(1): 120-127.
- Di Primo, C. and I. Lebars (2007). "Determination of refractive index increment ratios for protein-nucleic acid complexes by surface plasmon resonance." *Analytical Biochemistry* 368(2): 148-155.
- Donald, J. E., D. W. Kulp, et al. (2011). "Salt bridges: Geometrically specific, designable interactions." *Proteins: Structure, Function, and Bioinformatics* 79(3): 898-915.
- Duelli, A., E. Rönnberg, et al. (2009). "Mast Cell Differentiation and Activation Is Closely Linked to Expression of Genes Coding for the Serglycin Proteoglycan Core Protein and a Distinct Set of Chondroitin Sulfate and Heparin Sulfotransferases." *The Journal of Immunology* 183(11): 7073-7083.
- Eckert, K. A. and T. A. Kunkel (1991). "DNA polymerase fidelity and the polymerase chain reaction." *Genome Research* 1(1): 17-24.
- Edavettal, S. C., K. A. Lee, et al. (2004). "Crystal Structure and Mutational Analysis of Heparan Sulfate 3-O-Sulfotransferase Isoform 1." *Journal of Biological Chemistry* 279(24): 25789-25797.
- Escolar, L. a., J. Pérez-Martín, et al. (2000). "Evidence of an Unusually Long Operator for the Fur Repressor in the Aerobactin Promoter of *Escherichia coli*." *Journal of Biological Chemistry* 275(32): 24709-24714.
- Esko, J. D. and S. B. Selleck (2002). "ORDER OUT OF CHAOS: Assembly of Ligand Binding Sites in Heparan Sulfate1." *Annual Review of Biochemistry* 71(1): 435-471.
- Everberg, H., T. Leiding, et al. (2006). "Efficient and non-denaturing membrane solubilization combined with enrichment of membrane protein complexes by detergent/polymer aqueous two-phase partitioning for proteome analysis." *Journal of Chromatography A* 1122(1-2): 35-46.
- Fausnaugh, J., L. and F. Regnier, E. (1986). "Solute and mobile phase contributions to retention in hydrophobic interaction chromatography of proteins." *Journal of Chromatography*. 30(359): 131-146.
- Ferguson, B., W. and S. Datta (2011). "Role of Heparan Sulfate 2-O-Sulfotransferase in Prostate Cancer Cell Proliferation, Invasion, and Growth Factor Signaling." *Prostate Cancer:2011(2011): 1-14*.

- Feyzi, E., E. Trybala, et al. (1997). "Structural Requirement of Heparan Sulfate for Interaction with Herpes Simplex Virus Type 1 Virions and Isolated Glycoprotein C." *Journal of Biological Chemistry* 272(40): 24850-24857.
- Foxall, C., K. R. Holme, et al. (1995). "An Enzyme-Linked Immunosorbent Assay Using Biotinylated Heparan Sulfate to Evaluate the Interactions of Heparin-like Molecules and Basic Fibroblast Growth Factor." *Analytical Biochemistry* 231(2): 366-373.
- Fu, L., F. Zhang, et al. (2014). "Structure and Activity of a New Low-Molecular-Weight Heparin Produced by Enzymatic Ultrafiltration." *Journal of Pharmaceutical Sciences* 103(5): 1375-1383.
- Gama, C., I. , S. Tully, E. , et al. (2006). "Sulfation patterns of glycosaminoglycans encode molecular recognition and activity." *Nature Chemical Biology* 2(9): 467-473.
- García-Fruitós, E., M. M. Carrió, et al. (2005). "Folding of a misfolding-prone  $\beta$ -galactosidase in absence of DnaK." *Biotechnology and Bioengineering* 90(7): 869-875.
- Garner, J. A. (2003). "Herpes simplex virion entry into and intracellular transport within mammalian cells." *Advanced Drug Delivery Reviews* 55(11): 1497-1513.
- Georgiou, G. S., L. (2005). "Preparative expression of secreted proteins in bacteria: status report and future prospects." *Current Opinion in Biotechnology* 16(5): 538-545.
- Geraghty, R. J., C. Krummenacher, et al. (1998). "Entry of Alphaherpesviruses Mediated by Poliovirus Receptor-Related Protein 1 and Poliovirus Receptor." *Science* 280(5369): 1618-1620.
- Gerdes (1988). "The parB (hok/sok) Locus of Plasmid R1: A General Purpose Plasmid Stabilization System." *Nature Biotechnology* 6(12): 1402-1405.
- Ghosh, S., S. Rasheedi, et al. (2004). "Method for enhancing solubility of the expressed recombinant proteins in Escherichia coli." *Biotechniques* 37(3): 422-423.
- Gilligan, J., J. , P. Schuck, et al. (2002). "Mass spectrometry after capture and small-volume elution of analyte from a surface plasmon resonance biosensor." *Analytical Chemistry* 74(9): 2041-2047.
- Goldberg, M., R. R., et al. (1991). "A kinetic study of the competition between renaturation and aggregation during the refolding of denatured-reduced egg white lysozyme." *Biochemistry* 30(11): 2790-2797.
- Golovanov, A. P., G. M. Hautbergue, et al. (2004). "A Simple Method for Improving Protein Solubility and Long-Term Stability." *Journal of the American Chemical Society* 126(29): 8933-8939.
- Gorokhov, A., L. Perera, et al. (2000). "Heparan Sulfate Biosynthesis: A Theoretical Study of the Initial Sulfation Step by N-Deacetylase/N-Sulfotransferase." *Biophysical Journal* 79(6): 2909-2917.
- Gronenborn (1976). "Overproduction of phage lambda repressor under control of the lac promoter of Escherichia coli." *Molecular Genetics and Genomics* 148(3): 243-250.
- Guo, X. (2012). "Surface plasmon resonance based biosensor technique: A review." *Journal of Biophotonics* 5(7): 483-501.
- Guo, X., M. Condra, et al. (2003). "Determination of molecular weight of heparin by size exclusion chromatography with universal calibration." *Analytical Biochemistry* 312(1): 33-39.
- Guruprasad, K., B. V. B. Reddy, et al. (1990). "Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence." *Protein Engineering* 4(2): 155-161.
- Habuchi, H., M. Tanaka, et al. (2000). "The Occurrence of Three Isoforms of Heparan Sulfate 6-O-Sulfotransferase Having Different Specificities for Hexuronic Acid Adjacent to the Targeted N-Sulfoglucosamine." *Journal of Biological Chemistry* 275(4): 2859-2868.
- Haldimann, D., Wanner (1998). "Use of New Methods for Construction of Tightly Regulated Arabinose and Rhamnose Promoter Fusions in Studies of the Escherichia coli Phosphate Regulon." *Journal of Bacteriology* 180(5): 1277-1286.

- Hart, R., A., P. Kallio, T., et al. (1994). "Effect of biosynthetic manipulation of heme on insolubility of *Vitreoscilla* hemoglobin in *Escherichia coli*." *Applied and Environmental Microbiology*. 60(7): 2431-2437.
- Herold, B. C., D. WuDunn, et al. (1991). "Glycoprotein C of herpes simplex virus type 1 plays a principal role in the adsorption of virus to cells and in infectivity." *Journal of Virology* 65(3): 1090-1098.
- Higashi, K., S. Hosoyama, et al. (2012). "Photochemical Preparation of a Novel Low Molecular Weight Heparin." *Carbohydrate Polymers*. 67(2): 1737-1743.
- Hogrefe, H., H., J. Cline, et al. (2002). "Creating randomized amino acid libraries with the QuikChange Multi Site-Directed Mutagenesis Kit." *Biotechniques*. 33(5): 1164-1165.
- Hsieh, C.-C., J. Y. Guo, et al. (2013). "Quantitative Analysis of Oligosaccharides Derived from Sulfated Glycosaminoglycans by Nanodiamond-Based Affinity Purification and Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry." *Analytical Chemistry* 85(9): 4342-4349.
- IKAI, A. (1980). "Thermostability and Aliphatic Index of Globular Proteins." *Journal of Biochemistry* 88(6): 1895-1898.
- Jana, S. and J. K. Deb (2005). "Strategies for efficient production of heterologous proteins in *Escherichia coli*." *Applied Microbiology and Biotechnology* 67(3): 289-298.
- Johnsson, B., S. Löfås, et al. (1995). "Comparison of methods for immobilization to carboxymethyl dextran sensor surfaces by analysis of the specific activity of monoclonal antibodies." *Journal of Molecular Recognition*. 8(2): 125-131.
- Kakuta, Y., L. Li, et al. (2003). "Heparan sulphate N-sulphotransferase activity: reaction mechanism and substrate recognition." *Biochemistry Society Transactions* 31(2): 331-334.
- Kempf, B. and E. Bremer (1998). "Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments." *Archives of Microbiology* 170(5): 319-330.
- Khilko, S. N., M. Corr, et al. (1993). "Direct detection of major histocompatibility complex class I binding to antigenic peptides using surface plasmon resonance. Peptide immobilization and characterization of binding specificity." *Journal of Biological Chemistry* 268(21): 15425-15434.
- King, G. and N. Murray, E. (1994). "Restriction enzymes in cells, not eppendorfs." *Trends Microbiology*. 2(12): 465-469.
- Kleymann, G. (2003). "Novel agents and strategies to treat herpes simplex virus infections." *Expert Opinion on Investigational Drugs* 12(2): 165-183.
- Koerner, Megan M., Luis A. Palacio, et al. (2011). "Electrodynamics of Lipid Membrane Interactions in the Presence of Zwitterionic Buffers." *Biophysical Journal* 101(2): 362-369.
- Kreuger, J. and L. Kjellén (2012). "Heparan Sulfate Biosynthesis: Regulation and Variability." *Journal of Histochemistry & Cytochemistry* 60(12): 898-907.
- Kröger, D., A. Katerkamp, et al. (1998). "Surface investigations on the development of a direct optical immunosensor." *Biosensors and Bioelectronics* 13(10): 1141-1147.
- Kuczynska-Wisnik, D., S. Kędzierska, et al. (2002). "The *Escherichia coli* small heat-shock proteins IbpA and IbpB prevent the aggregation of endogenous proteins denatured in vivo during extreme heat shock." *Microbiology* 148(6): 1757-1765.
- Kyte, J. and R. F. Doolittle (1982). "A simple method for displaying the hydropathic character of a protein." *Journal of Molecular Biology* 157(1): 105-132.
- Landgraf, A. and H. Wolfes (1993). Taq Polymerase (EC 2.7.7.7). *Enzymes of Molecular Biology*. M. Burrell, Humana Press. 16: 31-58.
- Lawrence, R., T. Yabe, et al. (2007). "The principal neuronal gD-type 3-O-sulfotransferases and their products in central and peripheral nervous system tissues." *Matrix Biology* 26(6): 442-455.

- Leemhuis, H., H. J. Rozeboom, et al. (2003). "Conversion of Cyclodextrin Glycosyltransferase into a Starch Hydrolase by Directed Evolution: The Role of Alanine 230 in Acceptor Subsite +1†,‡." *Biochemistry* 42(24): 7518-7526.
- Lefèvre, F., M.-H. Rémy, et al. (1997). "Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function." *Nucleic Acids Research* 25(2): 447-448.
- Leviatan, S., K. Sawada, et al. (2010). "Combinatorial Method for Overexpression of Membrane Proteins in *Escherichia coli*." *Journal of Biological Chemistry* 285(31): 23548-23556.
- Li, W., D. J. D. Johnson, et al. (2004). "Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin." *Nature Structural and Molecular Biology* 11(9): 857-862.
- Lindahl, U., M. Kusche-Gullberg, et al. (1998). "Regulated Diversity of Heparan Sulfate." *Journal of Biological Chemistry* 273(39): 24979-24982.
- Linhardt, R., J. and N. Gunay, S. (1999). "Production and chemical processing of low molecular weight heparins." *Seminars in Thrombosis and Hemostasis* 25(3): 5-16.
- Liu, J. and R. Rosenberg, D. (2002). "Handbook of Glycosyltransferases and Their Related Genes." *Springer-Verlag Taniguchi, N and Fukuda, M. eds.*
- Liu, J., Z. Shriver, et al. (1999). "Heparan Sulfate d-Glucosaminy 3-O-Sulfotransferase-3A SulfatesN-Unsubstituted Glucosamine Residues." *Journal of Biological Chemistry* 274(53): 38155-38162.
- Liu, J., Z. Shriver, et al. (2002). "Characterization of a Heparan Sulfate Octasaccharide That Binds to Herpes Simplex Virus Type 1 Glycoprotein D." *Journal of Biological Chemistry* 277(36): 33456-33467.
- Liu, J. and S. C. Thorp (2002). "Cell surface heparan sulfate and its roles in assisting viral infections." *Medicinal Research Reviews* 22(1): 1-25.
- Lopez, P. J. G., Jean. Sousa, Rui. Dreyfus, Marc (1998). "On the mechanism of inhibition of phage T7 RNA polymerase by lac repressor." *Journal of Molecular Biology* 276(5): 861-865.
- Majumder, A., S. Basak, et al. (2001). "Effect of Osmolytes and Chaperone-like Action of P-protein on Folding of Nucleocapsid Protein of Chandipura Virus." *Journal of Biological Chemistry* 276(33): 30948-30955.
- Mattei, B., F. Cervone, et al. (2001). "The Interaction between Endopolygalacturonase from *Fusarium moniliforme* and PGIP from *Phaseolus vulgaris* Studied by Surface Plasmon Resonance and Mass Spectrometry." *Comparative and Functional Genomics* 2(6): 359-364.
- Matthews, B. (1996). "Structural and genetic analysis of the folding and function of T4 lysozyme." *Journal of the Federation of American Societies for Experimental Biology* 10(1): 35-41.
- Mettenleiter, T., C. (1989). "Glycoprotein gIII deletion mutants of pseudorabies virus are impaired in virus entry." *Virology* 171(2): 623-625.
- Misumi, M. and N. Tanaka (1980). "Mechanism of inhibition of translocation by kanamycin and viomycin: A comparative study with fusidic acid." *Biochemical and Biophysical Research Communications* 92(2): 647-654.
- Mochizuki, H., K. Yoshida, et al. (2008). "Tetrasulfated Disaccharide Unit in Heparan Sulfate: ENZYMATIC FORMATION AND TISSUE DISTRIBUTION." *Journal of Biological Chemistry* 283(45): 31237-31245.
- Mogk, A., M. P. Mayer, et al. (2002). "Mechanisms of Protein Folding: Molecular Chaperones and Their Application in Biotechnology." *Journal of Chemical Bbiology*: 3(9): 807-814.
- Moon, A. F., S. C. Edavettal, et al. (2004). "Structural Analysis of the Sulfotransferase (3-O-Sulfotransferase Isoform 3) Involved in the Biosynthesis of an Entry Receptor for Herpes Simplex Virus 1." *Journal of Biological Chemistry* 279(43): 45185-45193.

- Moon, A. F., Y. Xu, et al. (2012). "Dissecting the substrate recognition of 3-O-sulfotransferase for the biosynthesis of anticoagulant heparin." Proceedings of the National Academy of Sciences: 109(14): 5265-5270.
- Muñoz, E., D. Xu, et al. (2006). "Affinity, Kinetic, and Structural Study of the Interaction of 3-O-Sulfotransferase Isoform 1 with Heparan Sulfate<sup>†</sup>." Biochemistry 45(16): 5122-5128.
- Murray, V., J. Chen, et al. (2010). "Preparation of Very-High-Yield Recombinant Proteins Using Novel High-Cell-Density Bacterial Expression Methods." Cold Spring Harbor Protocols 2010(8): pdb.prot5475.
- Nath, D. and M. Rao (2001). "Artificial chaperone mediated refolding of xylanase from an alkalophilic thermophilic *Bacillus* sp. ." European Journal of Biochemistry 268(20): 5471-5478.
- Neagu, A., M. Neagu, et al. (2001). "Fluctuations and the Hofmeister Effect." Biophysical Journal 81(3): 1285-1294.
- Nedelkov, D. and R. W. Nelson (2003). "Surface plasmon resonance mass spectrometry: recent progress and outlooks." Trends in Biotechnology 21(7): 301-305.
- Negishi, M., L. G. Pedersen, et al. (2001). "Structure and Function of Sulfotransferases." Archives of Biochemistry and Biophysics 390(2): 149-157.
- Noti, C., J. L. de Paz, et al. (2006). "Preparation and Use of Microarrays Containing Synthetic Heparin Oligosaccharides for the Rapid Analysis of Heparin-Protein Interactions." Chemistry – A European Journal 12(34): 8664-8686.
- O'Donnell, C. D., M. Kovacs, et al. (2010). "Expanding the role of 3-O sulfated heparan sulfate in herpes simplex virus type-1 entry." Virology 397(2): 389-398.
- O'Donnell, C. D., V. Tiwari, et al. (2006). "A role for heparan sulfate 3-O-sulfotransferase isoform 2 in herpes simplex virus type 1 entry and spread." Virology 346(2): 452-459.
- O'Shannessy, D. J., M. Brigham-Burke, et al. (1992). "Immobilization chemistries suitable for use in the BIAcore surface plasmon resonance detector." Analytical Biochemistry 205(1): 132-136.
- Ostermeier, M. and S. Benkovic, J. (2001). "Construction of hybrid gene libraries involving the circular permutation of DNA." Biotechnology Letters 23: 303-310.
- Packeiser, H., C. Lim, et al. (2013). "An Extremely Simple and Effective Colony PCR Procedure for Bacteria, Yeasts, and Microalgae." Applied Biochemistry and Biotechnology 169(2): 695-700.
- Packer, N. H., C.-W. von der Lieth, et al. (2008). "Frontiers in glycomics: Bioinformatics and biomarkers in disease An NIH White Paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11–13, 2006)." PROTEOMICS 8(1): 8-20.
- Påhlman, J. Rosengren, et al. (1977). "Hydrophobic interaction chromatography on uncharged Sepharose derivatives. Effects of neutral salts on the adsorption of proteins." Journal of Chromatography. 21(131): 99-108.
- Patronov, A., I. Dimitrov, et al. (2012). "Peptide binding to HLA-DP proteins at pH 5.0 and pH 7.0: a quantitative molecular docking study." Bio Med Central Structural Biology: 21(20): 12-20.
- Peternel, Š., S. Jevšvar, et al. (2008). "New properties of inclusion bodies with implications for biotechnology " Biotechnology Applications in Biochemistry: 49(4): 229-246.
- Petitou, M., J.-P. Herault, et al. (1999). "Synthesis of thrombin-inhibiting heparin mimetics without side effects." Nature 398(6726): 417-422.
- Phillips, T., A., R. VanBogelen, A., et al. (1984). "lon gene product of *Escherichia coli* is a heat-shock protein." Journal of Bacteriology. 159(1): 283-287.
- Pinhal, M. A. S., B. Smith, et al. (2001). "Enzyme interactions in heparan sulfate biosynthesis: Uronosyl 5-epimerase and 2-O-sulfotransferase interact in vivo." Proceedings of the National Academy of Sciences 98(23): 12984-12989.

- Prasad, S., P. B. Khadare, et al. (2011). "Effect of Chemical Chaperones in Improving the Solubility of Recombinant Proteins in *Escherichia coli*." *Applied and Environmental Microbiology* 77(13): 4603-4609.
- Quinn, J. G., S. O'Neill, et al. (2000). "Development and Application of Surface Plasmon Resonance-Based Biosensors for the Detection of Cell–Ligand Interactions." *Analytical Biochemistry* 281(2): 135-143.
- Rich, R. L. and D. G. Myszka (2008). "Survey of the year 2007 commercial optical biosensor literature." *Journal of Molecular Recognition* 21(6): 355-400.
- Ritz, D., J. Lim, et al. (2001). "Conversion of a Peroxiredoxin into a Disulfide Reductase by a Triplet Repeat Expansion." *Science* 294(5540): 158-160.
- Samuel, D., G. Ganesh, et al. (2000). "Proline inhibits aggregation during protein refolding." *Protein Science* 9(2): 344-352.
- Schlieker, C., B. Bukau, et al. (2002). "Prevention and reversion of protein aggregation by molecular chaperones in the *E. coli* cytosol: implications for their applicability in biotechnology." *Journal of Biotechnology* 96(1): 13-21.
- Schwörer, R., O. V. Zubkova, et al. (2013). "Synthesis of a Targeted Library of Heparan Sulfate Hexa- to Dodecasaccharides as Inhibitors of  $\beta$ -Secretase: Potential Therapeutics for Alzheimer's Disease." *Chemistry – A European Journal* 19(21): 6817-6823.
- Seo, Y., A. Andaya, et al. (2012). "Preparation, Separation, and Conformational Analysis of Differentially Sulfated Heparin Octasaccharide Isomers Using Ion Mobility Mass Spectrometry." *Analytical Chemistry* 84(5): 2416-2423.
- Shafikhani, S., R. Siegel, A., et al. (1997). "Generation of large libraries of random mutants in *Bacillus subtilis* by PCR-based plasmid multimerization." *Biotechniques* 23(2): 304-310.
- Sharma, D., A. R. Cukras, et al. (2007). "Mutational Analysis of S12 Protein and Implications for the Accuracy of Decoding by the Ribosome." *Journal of Molecular Biology* 374(4): 1065-1076.
- Shokri, A., A. Sandén, et al. (2003). "Cell and process design for targeting of recombinant protein into the culture medium of *Escherichia coli*." *Applied Microbiology and Biotechnology* 60(6): 654-664.
- Shukla, D., J. Liu, et al. (1999). "A Novel Role for 3-O-Sulfated Heparan Sulfate in Herpes Simplex Virus 1 Entry." *Cell* 99(1): 13-22.
- Shworak, N. W., J. Liu, et al. (1997). "Molecular Cloning and Expression of Mouse and Human cDNAs Encoding Heparan Sulfate d-Glucosaminyl 3-O-Sulfotransferase." *Journal of Biological Chemistry* 272(44): 28008-28019.
- Shworak, N. W., J. Liu, et al. (1999). "Multiple Isoforms of Heparan Sulfate d-Glucosaminyl 3-O-Sulfotransferase: ISOLATION, CHARACTERIZATION, AND EXPRESSION OF HUMAN cDNAs AND IDENTIFICATION OF DISTINCT GENOMIC LOCI." *Journal of Biological Chemistry* 274(8): 5170-5184.
- Sieber, V., C. A. Martinez, et al. (2001). "Libraries of hybrid proteins from distantly related sequences." *Nature Biotechnology* 19(5): 456-460.
- Skerra (1994). "Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in *Escherichia coli*." *Gene* 151(1): 131-135.
- Smith, M., A. and M. Bidochka, J. (1998). "Bacterial fitness and plasmid loss: the importance of culture conditions and plasmid size." *Canadian Journal of Microbiology*. 44(4): 351-355.
- Spear, P. G., R. J. Eisenberg, et al. (2000). "Three Classes of Cell Surface Receptors for Alpha herpesvirus Entry." *Virology* 275(1): 1-8.
- Studier (2005). "Protein production by auto-induction in high density shaking cultures." *Protein Expression and Purification*. 41(1): 207-234.
- Studier, F. W. (1991). "Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system." *Journal of Molecular Biology* 219(1): 37-44.
- Sugahara, K. and H. Kitagawa (2002). "Heparin and Heparan Sulfate Biosynthesis." *IUBMB Life* 54(4): 163-175.

- Summers, D. (1998). "Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability." *Molecular Microbiology* 29(5): 1137-1145.
- Suto, S., I. Kakizaki, et al. (2014). "One set system for the synthesis and purification of glycosaminoglycan oligosaccharides reconstructed using a hyaluronidase-immobilized column." *Biopolymers* 101(3): 189-196.
- Tao, H., W. Liu, et al. (2010). "Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS." *Biotechniques*. 48(1): 61-64.
- Thakkar, S., N. Nanaware-Kharade, et al. (2014). "Affinity improvement of a therapeutic antibody to methamphetamine and amphetamine through structure-based antibody engineering." *Scientific Reports* 4(3673): 1000-1036.
- Thévenot, D., R. , K. Toth, et al. (2001). "Electrochemical biosensors: recommended definitions and classification." *Biosensors and Bioelectronics* 16(1): 121-131.
- Timasheff, S., N. (1998). "Control of protein stability and reactions by weakly interacting cosolvents: the simplicity of the complicated." *Advances in Protein Chemistry*. 51(1): 355-432.
- Tkachenko, E., J. M. Rhodes, et al. (2005). "Syndecans." *Circulation Research* 96(5): 488-500.
- Toida, T., H. Yoshida, et al. (1997). "Structural differences and the presence of unsubstituted amino groups in heparan sulphates from different tissues and species." *Biochemistry Journal*. 322(2): 499-506.
- Trybala, E., T. Bergström, et al. (1994). "Localization of a functional site on herpes simplex virus type 1 glycoprotein C involved in binding to cell surface heparan sulphate." *Journal of General Virology* 75(4): 743-752.
- Tsumoto, K., M. Umetsu, et al. (2003). "Solubilization of active green fluorescent protein from insoluble particles by guanidine and arginine." *Biochemical and Biophysical Research Communications* 312(4): 1383-1386.
- Tuncbag, N., A. Gursoy, et al. (2009). "Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy." *Bioinformatics* 25(12): 1513-1520.
- Turnbull, J., E. and J. Gallagher, T. (1991). "Distribution of iduronate 2-sulphate residues in heparan sulphate. Evidence for an ordered polymeric structure." *Biochemistry Journal*. 273(3): 553-559.
- Turnbull, J., A. Powell, et al. (2001). "Heparan sulfate: decoding a dynamic multifunctional cell regulator." *Trends in Cell Biology* 11(2): 75-82.
- Umetsu, M., K. Tsumoto, et al. (2005). "Nondenaturing solubilization of 82 microglobulin from inclusion bodies by L-arginine." *Biochemical and Biophysical Research Communications* 328(1): 189-197.
- Upadhyay, A., K. , A. Murmu, et al. (2012). "Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*." *PLoS One*. 7(3): 1061-1071.
- Vavrová, L., K. Muchová, et al. (2010). "Comparison of different *Bacillus subtilis* expression systems." *Research in Microbiology* 161(9): 791-797.
- Viens, A., F. Harper, et al. (2008). "Use of Protein Biotinylation In Vivo for Immunoelectron Microscopic Localization of a Specific Protein Isoform." *Journal of Histochemistry & Cytochemistry* 56(10): 911-919.
- Volpi, N. and F. Maccari (2006). "Electrophoretic approaches to the analysis of complex polysaccharides." *Journal of Chromatography B* 834(1-2): 1-13.
- Wang, J., S. Zhang, et al. (2007). "PCR-based strategy for construction of multi-site-saturation mutagenic expression library." *Journal of Microbiological Methods* 71(3): 225-230.
- Weinshilboum, R. M., D. M. Otterness, et al. (1997). "Sulfation and sulfotransferases 1: Sulfotransferase molecular biology: cDNAs and genes." *The FASEB Journal* 11(1): 3-14.
- Whitbeck, J. C., M. I. Muggeridge, et al. (1999). "The Major Neutralizing Antigenic Site on Herpes Simplex Virus Glycoprotein D Overlaps a Receptor-Binding Domain." *Journal of Virology* 73(12): 9879-9890.

- Xia, G., J. Chen, et al. (2002). "Heparan Sulfate 3-O-Sulfotransferase Isoform 5 Generates Both an Antithrombin-binding Site and an Entry Receptor for Herpes Simplex Virus, Type 1." *Journal of Biological Chemistry* 277(40): 37912-37919.
- Xiao, Z., B. R. Tappen, et al. (2010). "Heparin Mapping Using Heparin Lyases and the Generation of a Novel Low Molecular Weight Heparin." *Journal of Medicinal Chemistry* 54(2): 603-610.
- Xie, G. and S. N. Timasheff (1997). "Temperature dependence of the preferential interactions of ribonuclease a in aqueous co-solvent systems: Thermodynamic analysis." *Protein Science* 6(1): 222-232.
- Xu, D., A. F. Moon, et al. (2008). "Engineering sulfotransferases to modify heparan sulfate." *Nature Chemical Biology* 4(3): 200-202.
- Xu, Y., S. Masuko, et al. (2011). "Chemoenzymatic Synthesis of Homogeneous Ultralow Molecular Weight Heparins." *Science* 334(6055): 498-501.
- Y., C., S. Jinmei., et al. (2003). "DnaK and DnaJ facilitated the folding process and reduced inclusion body formation of magnesium transporter CorA overexpressed in Escherichia coli." *Protein Expression and Purification* 32(2): 221-231.
- Yang, B., K. Solakyildirim, et al. (2011). "Hyphenated techniques for the analysis of heparin and heparan sulfate." *Analytical and Bioanalytical Chemistry* 399(2): 541-557.
- Yazdanparast, R. and R. Khodarahmi (2005). "The combined effects of two anti-aggregatory agents, alpha-cyclodextrin and Ca<sup>2+</sup>, on the refolding process of denatured alpha-amylase." *Biotechnology Applied Biochemistry* 41(2): 157-162.
- Yi-Chi, C. L., R. Nathalie, et al. (2010). "Specific and reversible immobilization of histidine-tagged proteins on functionalized silicon nanowires." *Nanotechnology* 21(24): 957-1088.
- Zhao, H., L. Giver, et al. (1998). "Molecular evolution by staggered extension process (StEP) in vitro recombination." *Nature Biotechnology* 16(3): 258-261.
- Ziegler, A. and J. Zaia (2006). "Size-exclusion chromatography of heparin oligosaccharides at high and low pressure." *Journal of Chromatography B* 837(1-2): 76-86.