# THE DEVELOPMENT OF AFFINITY AQUEOUS TWO PHASE SYSTEMS AS A METHOD OF PURIFYING THERAPEUTIC PROTEIN

JENNIFER ELIZABETH LINES

Master of Philosophy

ASTON UNIVERSITY

February 2015

## Thesis summary

Aston University, The Development of Affinity Aqueous Two Phase Systems as a Method of Purifying Therapeutic Proteins, Jennifer Elizabeth Lines, Master of Philosophy, 2015.

Aqueous two-phase systems (ATPS) are formed when two immiscible and physico-chemically disparate polymers, such as poly(ethylene) glycol (PEG) and dextran (DEX), are mixed in water. Full explanations of the precise physical and chemical mechanisms of phase separation and macromolecule partitioning have yet to be described causing the prediction of macromolecule behaviour to be difficult. A systematic analysis of multiple 20% w/v PEG and DEX combinations was performed at room temperature in order to determine which formed two distinct phases with both purified green fluorescent protein (GFPuv) and *E.coli* Tuner (DE3) cell lysate containing GFPuv. All layers were checked for the presence of protein, DNA and RNA using fluorescence, Bradford assay and ethidium bromide staining of 1% agarose gels. The formation of two distinct layers occurred in 45% of combinations and preferentially when both PEG and DEX were of a higher molecular weight. All ATPS formed showed complete partitioning of genomic material to the lower phase and 60% showed complete partitioning of proteins and cell debris to the bottom phase or interface. It was also determined by fluorescence that 100% of GFPuv partitioned to the lower phase in all ATPS. As determined from previous screening, select ATPS combinations were formed using PEGylated glutathione and GST-fused GFPuv. Experimentation progressed onto using cell lysates containing histidine-tagged GFPuv and PEGylated nitrilotriacetic acid (NTA). Protein partitioning was analysed using Bradford assay and SDS-PAGE. PEG phase partitioning of GFPuv can be seen when 1% of overall PEG is functionalised. These results show that the partitioning of GFPuv is not dependent upon the presence of cell debris and will allow for quicker selection of PEG and DEX molecular weight combinations. The results will also allow further development of the system into a generic affinity purification method using PEGylated NTA with coordinated copper ions as a ligand for histidine tagged proteins.

Key Words or Phrases: ATPS; Affinity Purification; PEG; Dextran; Protein Purification;

## Dedication

First and foremost, I would like to dedicate this thesis to my family. Without their love, support, generosity and also their enthusiasm for education I would surely have not got this far. I also dedicate this work to Rosie, Jess and Coco, for making me keep normal hours, offering copious cuddles and always listening without judgement, thank you.

## Acknowledgments

I would like to extend my sincere thanks to my supervisor, Dr Anna Hine, for her knowledge, support and encouragement throughout my doctoral studies. I also acknowledge Professor Steve Brocchini, UCL, for his help and support throughout the project, in particular for sharing his PEGylation expertise. I would also like to thank Dr Andrew Sutherland, Aston University, for his advice in all aspects of chemistry. This project was jointly funded by the BBSRC and Polytherics Ltd through a BBSRC Industrial CASE Studentship.

For Mike,

never forgotten.

# Table of Contents

# Table of Figures

# Table of Tables

## Table of Schemes

.

## Abbreviations

| | |
|---|---|
| $^{1}$H NMR | Proton nuclear magnetic resonance |
| Af-ATPS | Affinity aqueous two-phase system |
| BSA | Bovine serum albumin |
| CIP | Cleaning in place |
| DNA | Deoxyribonucleic acid |
| EDTA | Ethylenediaminetetraacetic acid |
| EMA | European Medicines Agency |
| G6PDH | Glucose-6-phosphate dehydrogenase |
| GSH | Glutathione |
| GST | Glutathione *S*-transferase |
| HEL | Hen egg white lysozyme |
| IDA | Iminodiacetic acid |
| IEX | Ion exchange chromatography |
| IgG | Immunoglobulin G |
| IPTG | Isopropyl β-D-1-thiogalactopyranoside |
| mABs | Monoclonal antibodies |
| mPEG | Methoxy poly(ethylene) glycol |
| Ni-NTA | Nickle nitrilotriacetic acid |
| PBS | Phosphate buffered saline |
| PEG | Poly(ethylene) glycol |
| PrP | Prion proteins |
| SDS | Sodium dodecyl sulphate |
| SEC | Size-exclusion chromatography |
| TEMED | Tetramethylethylenediamine |
| UV | Ultraviolet |
| ATPS | Aqueous two phase system |
| CCC | Counter current chromatography |
| DEX | Dextran |

| | |
|---|---|
| E. *coli* | Escherischia *coli* |
| Fc region | Fragment crystallisable region |
| GFPuv | Green fluorescent protein |
| GSH | Glutathione |
| HSA | Human serum albumin |
| $His_x$ | Histidine tag (x denotes the number of histidine residues) |
| IDA | Iminodiacetic acid |
| IFN | Interferon |
| IMAC | Immobilised metal ion affinity column |
| LB | Lysongeny broth |
| MBP | Maltose binding protein |
| mPEGmal | Methoxy poly(ethylene) glycol maleimide |
| NTA | Nitrilotriacetic acid |
| PCR | Polymerase chain reaction |
| PEGylated | A molecule that has been conjugated to a PEG molecule |
| RNA | Ribonucleic acid |
| SDS-PAGE | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| TAE | Tris-acetate-EDTA |
| Tris | Tris(hydroxymethyl)aminomethane |
| VLP | Virus like particles |

# Chapter 1. Introduction

Therapeutically relevant proteins are required in increasing quantity for a wide number of applications ranging from new drug discovery, recombinant protein production and high-throughput screening on a research level to the development and production of proteins of pharmaceutical importance on an industrial scale (Saraswat *et al.,* 2013). Whilst many of the shortcomings of the upstream processes have been addressed already in recent years, with areas such as cell culture being improved to the point that recombinant protein yields have increased almost a thousand fold to tens of grams per litre (Gottschalk, 2008). With such high yields now achievable the purification process can be considered a highly labour intensive part of the process at the same time as incurring high production costs, in some cases the cost of downstream processing can range between 50 % and 80 % of the total cost of manufacture (Lowe, 2001). This has resulted in traditional downstream processing techniques, such as affinity chromatography, being considered too costly, as well as time consuming and not capable of processing these higher quantities within a reasonable time frame.

At present, a common method for large scale purification of therapeutically relevant proteins after harvesting, by using a method such as centrifugation or sterile filtration, is a number of affinity column chromatography steps followed by a viral filtration method. This can then be followed by an ultrafiltration / diafiltration stage in order to obtain a product that is concentrated in the required final buffer and pure enough for therapeutic use (Trexler-Schmidt *et al.,* 2009). There must also be a large enough quantity of the purified protein in both industrial and research scale productions in order to test the efficacy, structural integrity and resulting activity of the protein. There are various constraints associated with these methods, such as the high cost of resins and buffers, which have a limiting effect on the volumes able to be processed. In addition, high volume columns can incur problems such as hysteresis and edge-effects where

different flow rates are achieved across the column due to uneven resin packing. Resin compression can also reduce porosity, and therefore decrease the pressure across the column (P. A. J. Rosa *et al.,* 2010). These problems are also observed at laboratory bench scale, leaving a gap for investigation into a cost effective, generic affinity purification process that does not utilise column chromatography.

Accordingly, a systematic review of the literature was performed to determine whether there was a gap in common knowledge that may be filled by further laboratory based research into the area. The aim of the following review is to identify whether the proposed development of affinity based aqueous two phase systems (ATPS) for the purification of therapeutic proteins is not only conceptually possible, but also laboratory feasible and commercially practicable.

The databases searched were ISI Web of Knowledge, Science Direct and Google Scholar using the following search terms – aqueous two phase*, ATPS, PEG, dextran, DEX, phase separation, protein purification, large scale, affinity ATPS, protein purification medicine, protein purification biotechnology and affinity purification. Articles that were published after 1990 were preferentially chosen with the exception of pertinent original research relevant to the area such as works by the author Albertsson which are relevant to the original development of ATPS.

## 1.1 Current Purification Methods

The methods employed to purify proteins can be broadly separated into two different types – affinity methods and chromatography methods. In this section both types of methods will be investigated into their current use in protein purification. The first group of methods described here are affinity based protein purification protocols. One of the preferred methods for the purification of therapeutically relevant proteins is affinity column chromatography due to its high resolving power, capacity and recovery (Waites, 2001). There are four main types of system which can be employed - polyhistidine tags that bind to metal ions; enzyme fusions that bind to immobilised ligands; antibody-antigen interactions and bacterial receptors that bind to serum proteins (Nilsson *et al.,* 1997).

Polyhistidine tags were first demonstrated in 1975 (Chaga, 2001) and have since become known as immobilised metal ion affinity chromatography (IMAC). IMAC is possible as the imidazole groups of the polyhistidine tag on either the C or N terminal of the protein will bind selectively to the metal ions available, such as nickel, copper and zinc. The metal ion is coordinated to the substrate via a linker and metal ion coordinating group such as iminodiacetic acid (IDA) or nitrilotriacetic acid (NTA). The protein is then eluted by lowering the pH, causing protonation of the imidazole groups, or introducing competition for the binding sites using imidazole. High protein recovery and purity make this a popular method, especially at research level however, the purity can be affected by non-specific binding of proteins containing external histidines in the tertiary structure. The IMAC methodology has become common practice when purifying proteins that aggregate into inclusion bodies as purification can occur under denaturing conditions and the protein can subsequently be refolded before elution (Rogl *et al.,* 1998). In addition to this, IMAC can be modified and used to purify certain antibodies by first using a recombinant his-tagged antigen that binds to the substrate, which can then bind the

bio-active antibodies to be purified (Casey *et al.,* 1995). However, there are disadvantages to this purification method since the metal chelates used to bind the metal ions to the substrate can be used for site specific cleavage of proteins, which potentially reduces the yield (Gaberc-Porekar & Menart, 2001). Another problem that can occur is the leaching of the metal ions, of which nickel compounds are known to be carcinogens, rendering the products unsuitable for pharmaceutical use (Kasprzak *et al.,* 2003). Ni-NTA resins are commercially available through a number of companies such as Qiagen who provide instructions to regenerate the resin enabling more than one purification step with each aliquot. This allows for the reduction in costs for the purification step on a laboratory bench scale, nevertheless, the regeneration and reuse of the Ni-NTA resin may not be suitable for industrial purification due to the potential for cross contamination. An additional drawback for the up scaling of this process is the high cost of the metal ion as well as the resin used to pack the column. Such large volumes would be required that not only would the costs be large, but other effects can occur such as edge hysteresis reducing the efficiency of purification. His-tag purification of recombinant proteins, whilst inherently useful for research level purposes, may not be a first port of call for industrial use due to the requirement of a final protein with no additional C- or N- terminal products. A protease able to digest the his-tag away from the required protein would be necessary which adds an additional digestion and purification step to the process increasing not only the time necessary but also the cost of manufacture.

Recombinant proteins that are fused to enzymes such as Glutathione-S-transferase (GST) are also successfully purified by using an immobilised ligand, in this case glutathione to which GST has a high specificity and affinity (H.M. Chen *et al.,* 1999). GST vectors have now been designed to allow the fusion to occur at either the N- or C-terminal of the desired protein; this is particularly useful as some proteins have terminals that may be involved in protein folding (Aatsinki & Rajaniemi, 2005). Once the GST-Protein fusion has bound to the column it can be eluted competitively using reduced glutathione, but this can be disadvantageous since reduced glutathione can affect the

disulfide bonds contained within the protein that is being purified, preventing it from folding properly (Chakravarthi *et al.,* 2006). When the GST-protein fusion is made, a cleavage site is engineered in to allow for the easy removal of the GST. The site specific protease is often itself developed with an affinity tag fusion in order to allow a further step of affinity purification to occur, and therefore a pure product (Arnau *et al.,* 2006). GST-protein fusions have also been shown to be able to be renatured if inclusion bodies are formed (Nguyen *et al.,* 1996). GST fusion proteins have been shown to be able to be produced in large batch quantities (Korf, U *et al,* 2005), however, the drawbacks to this method arise when harvesting the fusion protein from solution. Similarly to Ni-NTA and his-tag purification, a large quantity of immobilising substrate is required which can prove costly and have a reduced efficiency.

Maltose binding protein (MBP) is a bacterial protein that is used in a similar fashion to GST for affinity purification. MBP specifically binds maltose or maltodextrins in the periplasmic space in order for the transport and catalysis of these sugars to occur across the cytoplasmic membrane (Shuman, 1982). Once the MBP fusion, constructed with a protease cleavage site to allow for removal, has bound onto the cross linked amylose resin, competitive elution can occur by submitting the column to high concentrations of maltose (Srinivasan & Bell, 1998). This method is particularly useful as the conditions used are mild and because the MBP does not contain any cysteine residues which could interfere with the fused protein's disulfide bonds and therefore folding abilities. In fact, MBP has been shown to be effective at promoting the solubility of the protein that it is fused to (Kapust & Waugh, 1999).

Protein A contains a binding site for the Fc region of mammalian immunoglobulin G (IgG) and, when bound to a substrate, has been established as a method of affinity purification within both the pharmaceutical industry and at research level, especially during the production of antibodies and specifically monoclonal antibodies (mAbs). It is widely acknowledged for its high resolving power resulting in elutions with a high degree of purity although additional polishing purification steps may be necessary to remove

any impurities such as leached Protein A to render the product suitable for use (Liu *et al.,* 2010). Commercially available Protein A resins such as Mabselect available from GE Healthcare Life Sciences are able to withstand tough conditions, especially when using harsh cleaning-in-place (CIP) methods, and the ability to harvest mAbs from clarified lysate directly. These commercially available systems range from small, laboratory bench scale up to industrial scale allowing for effective upscaling of production. (http://www.gelifesciences.com/gehcls_images/GELS/Related%20Content/Files/1314823637792/litdoc11001159_20130905164352.PDF, Accessed 17/01/15). The long usable lifetime of the resin makes this method attractive when a reduction in costs is important; however the narrow scope of its application is limited to the purification of IgG only therefore it cannot be a generic process. The limitation with all of these affinity column chromatography methods is the high cost of resin as well as a limit to the volume of supernatant that can be applied at that point in time. This causes these methodologies to still create a bottleneck in the downstream processing of therapeutic proteins (Porath, J. 1992).

Other chromatographic methods that can be used to purify proteins are available but lack one or more of the essential characteristics mentioned previously. Ion-exchange chromatography (IEX) can be used to purify proteins with a known overall charge, which can interact with a covalently bonded anionic or cationic functional group. IEX is often employed when a protein is produced without an affinity tag and high resolutions are achievable with the correct pH binding buffer for the overall charge of the protein. To allow for the removal of cellular deoxyribonucleic acid (DNA), anion exchange chromatography is preferentially selected after a lysate clearing protocol for removing cellular debris and contaminants. The protein can then be eluted from the column by changing the ionic concentration or altering the pH of the mobile phase. Two main disadvantages to this method are that only a small volume can be loaded onto the column initially, and a concentration step must be performed in order to utilise the molecule once it has been eluted. This can lead to high concentrations of residual salts

in the solution rendering the protein unsuitable for therapeutic use without an additional buffer exchange step (Aguilar, 2004). Additional costs and time are incurred as the lysate clearing step is required prior to IEX being used in order for the method to be effective.

Gel filtration, or size exclusion chromatography (SEC), is a chromatographic purification method that filters proteins based on size as they pass through porous beds at different speeds. Smaller proteins travel through the beads and therefore spend a longer amount of time on the column than a larger protein that does not pass within the resin. To establish an accurate time course of proteins eluted due to size a long column and slow flow rates are necessary which once again increase the time and costs required for purification. In comparison to other methods previously described, SEC has a low resolution which can be widely affected by the choice of resin composition and pH of the mobile phase. SEC can only be employed when a pure and undiluted product is not required as there is an absence of specificity. This method can however be useful when employed after IEX as a polishing step due to its ability to desalt protein containing buffers. In order to desalt these buffers, water is applied to the SEC resin, that has been pre-equilibrated with water, allowing for proteins to pass quickly around the resin beads whilst the salts are delayed flowing through pores present in the resin beads (Hagel, 1998).

Chromatography methods are not always employed for protein purification. For example, acetone can be used to precipitate proteins from a solution, but any residual acetone may cause unwanted modifications in the protein (Simpson & Beynon, 2010). Another non chromatographic method is the use of magnetic beads which can be placed directly into the solution. These beads are composed of iron coated in a polymer that has a specific affinity ligand on the surface, and once bound with the protein they can be removed from the original solution using a magnet. This method negates the need for columns and centrifugation, which reduces the time and cost involved (Smith, 2005).

## 1.2 Aqueous Two-Phase Systems

Within organic chemistry liquid – liquid extraction, more generally referred to as solvent extraction, is a common methodology employed for the partitioning of a desired compound from any unwanted compounds using two liquid, immiscible solvents. Examples of compounds that are separated using solvent extraction can include aromatic compounds from within crude oil fractions and vitamins required from aqueous solutions (Gunt, 2015). When two organic solvents or an organic solvent and water are used to form the two phases for particle partitioning, the shear forces produced across the interface are high. These forces render a solvent based extraction redundant when being considered in the biotechnology industry for the partitioning of molecules such as proteins which need to maintain the essential secondary and tertiary structures in order to maintain function. To overcome this, the organic solvents can be replaced with aqueous solutions that are immiscible and therefore the forces experienced by the proteins are reduced when being partitioned across the interface.

In 1896, Beiherinck first observed the formation of Aqueous Two-Phase Systems (ATPS) when he combined liquid solutions of gelatin and agar; this was also the case when he mixed solutions of gelatin and starch. These observations have been the foundations of over a century of research into the mechanisms of phase separation and macromolecule partitioning phenomena, of which a full explanation is presently elusive. It is understood that ATPS are formed when two immiscible and disparate polymers, such as poly (ethylene) glycol (PEG) and dextran (DEX), or a polymer and a salt solution, are mixed in water. Due to the differing natures of the polymers, for example molecular weight and external polarities, two phases develop as there is a slight difference in water density between the layers (Albertsson, 1971). PEG and DEX ATPS were first described by Albertsson in 1958 and have since been the subject of much investigation, with

systems being described for the partitioning of many materials such as DNA (Barbosa *et al.,* 2008), proteins (Rosa *et al.,* 2007) and entire cells (Edahiro *et al.,* 2005).

The biotechnology industry has been increasing research in the ATPS area with a high interest in counter current chromatography (CCC) development due to the ability to have continuous separation. This is advantageous at an industrial scale as a reduction in time and material wastage, usually associated with batch purification, is seen. However, for the purposes of the researcher, CCC is likely to be costly and excessive for their needs. Therefore, a generic batch ATPS formed using a stationary separation methodology is beneficial on a smaller scale.

## 1.3 Affinity-Based Aqueous Two-Phase Systems

Considering how effectively affinity purification is used in other applications, it is surprising that there are not more reports of it being used in conjunction with ATPS. In 2010, Barbosa *et al.*, reported affinity purification of plasmid DNA, by using a DNA binding fusion protein that was affinity bound using a $His_6$ tag to a PEG-iminodiacetic acid (IDA)-$Cu^{2+}$. This dual affinity directs the complex to the PEG top phase out of a crude cell lysate ready for a second step in order to further purify the plasmid DNA away from the affinity complex. This PEG-IDA-$Cu^{2+}$ system has also been explored for the partitioning of erythrocytes which have metal binding sites on the cell surface (Goubran Botros *et al.,* 1991).

Affinity ATPS (Af-ATPS) have also been used to purify proteins by a number of methods. One of these was the partitioning of Immunoglobulin G (IgG) using PEG functionalised with glutaric, amino and benzyl groups. The authors determined that the most effective ligand was diglutaric acid (PEG-COOH) with an extraction of 97 % of IgG into the PEG phase (Azevedo *et al.,* 2009). This method also allowed the authors to achieve a high percentage of 94 % purity of IgG from other proteins. An affinity tagged green fluorescent protein (GFPuv) has also been partitioned in an ATPS from a clarified E.*coli* cell lysate using a micellar system. This system used decyl β-D-glucopyranoside ($C_{10}G_1$) as a surfactant which formed micelles and had affinity for the novel tag attached to the GFPuv (Mazzola *et al.,* 2006). Another example is shown by Fong *et al.,* in 2002 where an Fv antibody fragment was conjugated with a thermo separating polymer in order to partition hen egg white lysozyme (HEL). The authors showed that this method removed 80 % of the HEL from the solution. In order to partition glucose-6-phosphate dehydrogenase (G6PDH) and hexokinase to the top phase of a PEG / hydroxypropyl starch system, Xu and co-workers added free triazine dyes as affinity ligands (Xu *et al.,* 2003). It was found that the dye / enzyme complexes preferentially partitioned to the top

PEG layer and enabled the authors to further develop this system into a two stage extraction process for G6PDH.

Human acute myeloid lymphoma cells that are $CD34^+$ can also be separated from human T lymphoma cells using Af-ATPS. This is due to the antibody anti-CD34 being conjugated to a temperature sensitive polymer which was added to the PEG / DEX ATPS and then the polymer / cell complex was precipitated out of the top PEG layer by heating to 32 ºC (Kumar *et al.,* 2001).

If further developed into a generic, rather than a specific process, Af-ATPS has the potential to revolutionise the downstream processing that is commonplace in research and industry. Some of the advantages are that the ATPS environment, including the interface, is very mild and therefore does not denature the product, resulting in higher yields (Albertsson *et al.,* 1990). It can also be significantly more cost effective than column chromatography as alternative polymers to DEX, such as hydroxypropyl starch, have been studied and could potentially be scaled up to be a cheaper and more time-efficient processing method for large volumes of product (Tjerneld *et al.,* 1986).

## 1.4 Up Scaling

The ability to upscale protein purification methods has been investigated for many years. There is currently a bottle neck present in the downstream process of purification as large volumes of protein are able to be produced in large fermentation vats. Properties that make ATPS an attractive methodology for upscaling are the speed at which purification can be performed, attracting lower costs than chromatographic methods and promising high yields and purities (Kaul,R.H, 2011). When compared to Protein A purification, affinity ATPS is shown very favourably as having the ability to not only continuously purify mAbs but also a reduction in costs and environmental impact (Rosa, P.A.J. *et al.,* 2011).

It has been suggested by Gupta and co-workers that the importance for ATPS development on a large scale is increasing in industry and research due to the high cost and poorer yields of existing downstream processing techniques (Gupta *et al.,* 1999). Even though the up-scaling of ATPS is reported to be simple and could address the issues associated with the techniques being used at the moment, there are surprisingly few examples in the literature demonstrating any research into this area. It was recently shown that PEG / salt ATPS could be formed within flexible blood bags, allowing for the partitioning of human serum albumin (HSA) (Garza-Madrid *et al.,* 2010). The authors found that there was no impact on the partitioning behaviour of HSA caused by forming the ATPS in blood bags, and recovered 85% of the HSA in the top layer. They also demonstrated that the ATPS had a short equilibrium time of less than 6 minutes in the absence of centrifugation. Although using a PEG-salt system has shown to be effective in this case, salt is not always an ideal phase component due to increasing the interfacial tension between the two layers. The high salt environment can also be considered to be harsh for many therapeutic proteins causing them to denature before they can be purified for use.

In addition to the aforementioned studies, virus like particles (VLP) have been partitioned using a 500 ml multistage ATPS based on a centrifugal partitioning chromatic device (Effio *et al.,* 2015). This multistage system allowed for improved partitioning of the VLP from host cell proteins (69.1% purity when analysed using HPLC), albeit with a reduced yield of 40.1% VLP recovery when compared to the single stage ATPS where 100% VLP recovery was detected using HPLC. This methodology was not as successful as the previous blood bag example when comparing yields and purities but also required specialist equipment in order to produce continuous purification. This would increase the cost of each purification for a reduced yield and purity of therapeutic protein.

Furthermore, there has been research into the recovery of prion proteins (PrP) from crude feedstocks that are derived from homogenated bovine brain tissue; the presence of which has been associated with transmissible spongiform encephalopathies such as BSE (Walker *et al.,* 1996). An ATPS that contained a thermo-separating polymer has been shown to work at a volume of 500 L and a total recovery of 71 % of recombinant cutinase (Kepka *et al.,* 2003). The feasibility of up-scaling was further shown using a detergent based ATPS at a volume of 2000 L (Selber *et al.,* 2004), suggesting that it will be possible with other systems. Further investigation into suitable phase forming polymers can be done in order to reduce the costs associated. A number of examples have been suggested for use instead of dextran, which carries a higher cost than PEG, such as PVA, starch derivatives and ethyl hydroxyls. The ability to recycle many of the constituent polymers was also demonstrated when using the thermos regulated polymers (Li & Cao, 2010).

More recently, the scaling up of an ATPS system by 850x has been described by Ruiz-Ruiz *et al.,* in 2013. The authors describe the production and purification of ß-phycoerythrin from lab bench scale to pilot plant facilities. In this instance, however, an amalgamation of strategies were implemented in order to purify the protein including isoelectric precipitation, ATPS and ultrafiltration. The final yield was reported to be 54% with a final purity of 4.1 under optimal conditions. This demonstrates that there is a need

for further research into a generic large scale ATPS to fill this gap present in the literature.

## 1.5 Project Outline

In conclusion, the need to generate a generic Af-ATPS is becoming ever more pressing due to the high costs, long processing times and low yields of current solid/liquid affinity chromatography strategies. Therefore, the development of a new methodology would be of great benefit to both research and industry. Accordingly, the following plan of experimentation was devised.

- Initial formation of ATPS.
    - Differing combinations of PEG and DEX molecular weights were assessed to determine which formed two distinct phases.
    - The effect on ATPS formation and protein partitioning of both the settling and centrifugation methods were compared to find the optimal methodology.
    - The behaviours of purified GFPuv and GFPuv containing cell lysate in PEG / DEX ATPS were also analysed to determine the influence of cell debris upon protein partitioning within ATPS.
- ATPS optimisation of phase compositions.
    - The concentration and molecular weight of PEG and DEX was systematically analysed to find the optimal phase compositions for both distinct phase formation and protein partitioning.
- PEG-Glutathione based affinity to the upper PEG phase of ATPS.
    - PEGylation of glutathione (GSH) was performed at Polytherics for use within the novel affinity ATPS.
    - A glutathione s-transferase (GST) tagged protein was engineered as an exemplar for use within the novel affinity ATPS.

- PEG-Nitrilotriacetic acid-$Cu^{2+}$ based affinity to the upper PEG phase of ATPS.
    - PEGylation of Nitrilotriacetic acid (NTA) was performed at both UCL and Aston University for analysis of its suitability for use within the novel affinity ATPS.
    - A Histidine tagged ($His_6$) protein was engineered as an exemplar for use within the novel affinity ATPS.

# Chapter 2. Materials and Methods

This chapter will provide all materials and methods utilised in the development of aqueous two phase systems as a method of purifying therapeutic proteins. Unless otherwise stated, all experiments were performed in triplicate in order to show repeatability. When analysing the results produced in this study, quantitative statistical analysis was not performed due to the formation of two phases being visual and replicated being identical. quantitative or descriptive statistical analysis was not performed.

## 2.1 Media Recipes

### (2.1.1) Luria Bertani Broth

Luria Bertani (LB) Broth Powder (Sigma-Aldrich) at a concentration of 20 g/L was dissolved into $H_2O$ and sterilised by autoclaving at 121°C for 20 minutes. If a selective culture was required sterile Ampicillin solution (2.2.3) or sterile Kanamycin solution (2.2.4) was added to cooled media (<50°C) to a final concentration of 50 µg/ml.

### (2.1.2) Luria Bertani Agar

Luria Bertani (LB) Broth Powder (Sigma-Aldrich) was dissolved at a concentration of 20 g/L in $H_2O$ with the addition of Bacto Agar (DIFCO) at 20 g/L and sterilised by autoclaving at 121°C for 20 minutes. If selective media was required sterile Ampicillin solution (2.2.3) or sterile Kanamycin solution (2.2.4) was added to cooled media (<50°C) to a final concentration of 50 µg/ml.

## 2.2 Buffer Recipes

### (2.2.1) Tris-Acetate-EDTA Buffer (TAE)

50 x stock solution was made containing 2 M Tris-acetate and 0.05 M EDTA at pH 8.0. 1 x TAE was prepared by dilution in double distilled $H_2O$.

### (2.2.2) Loading Buffer

30% (v / v) Glycerol, 0.025% (w / v) Xylene cyanol and double distilled $H_2O$.

### (2.2.3) Ampicillin Solution

Stock solutions were prepared using double distilled $H_2O$ and Ampicillin Sodium Salt (Sigma-Aldrich). The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.4) Kanamycin Solution

Stock solutions were prepared using double distilled $H_2O$ and Kanamycin Sulfate (Sigma-Aldrich). The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.5) Phosphate Buffered Saline (PBS)

PBS was prepared by dissolving one PBS tablet (Invitrogen) into 200 ml distilled $H_2O$ yielding 0.01 M phosphate buffer, 0.0027 M potassium chloride and 0.137 M sodium chloride, pH 7.4, at 25 °C. This was then sterilised by autoclaving at 121°C for 20 minutes.

### (2.2.6) Calcium Chloride ($CaCl_2$) Solution

$CaCl_2$ stock solution (50mM) was made by combining double distilled $H_2O$ with $CaCl_2{\cdot}2H_2O$ followed by a filter sterilisation step using a 0.2 µm syringe filter (Nalgene).

### (2.2.7) Bradford Reagent

Bradford Reagent was obtained from Bio-Rad Laboratories Inc., and prepared by diluting 1 part Bradford Reagent to 4 parts distilled $H_2O$. This solution as then filtered through filter paper before use in methods (2.7.14) and (2.3.15).

### (2.2.8) Bovine Serum Albumin (BSA) Stock Solution

A stock solution of BSA at a concentration of 1mg/ml was prepared in double distilled water and aliquots frozen at -20°C to be used to prepare standard curves for the Bradford assay.

### (2.2.9) Lysis Buffer

A stock solution was prepared containing 50 mM $NaH_2PO_4$ (Sigma-Aldrich), 300 mM NaCl (Sigma-Aldrich), 10 mM imidazole (Sigma-Aldrich) and the pH adjusted to 8.0 using NaOH. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.10) Wash Buffer

A stock solution was prepared containing 50 mM $NaH_2PO_4$ (Sigma-Aldrich), 300 mM NaCl (Sigma-Aldrich), 20 mM imidazole (Sigma-Aldrich) and the pH adjusted to 8.0 using NaOH. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.11) Elution Buffer

A stock solution was prepared containing 50 mM $NaH_2PO_4$ (Sigma-Aldrich), 300 mM NaCl (Sigma-Aldrich), 250 mM imidazole (Sigma-Aldrich) and the pH adjusted to 8.0 using NaOH. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.12) Isopropyl β-D-1-thiogalactopyranoside (IPTG)

1 M stock solutions were prepared using double distilled water and then filter sterilised using a 0.2 µm syringe filter (Nalgene) and aliquots stored at -20°C.

### (2.2.13) PEGylation Reaction Buffer

This buffer was prepared using 50 mM $NaH_2PO_4$ (Sigma-Aldrich), 10 mM EDTA (Sigma-Aldrich) and the pH adjusted to 7.8. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.14) Sodium Phosphate Buffer for PEG Maleimide Conjugation

This buffer was prepared using 50 mM $NaH_2PO_4$ (Sigma-Aldrich), 10 mM EDTA (Sigma-Aldrich) and the pH adjusted to 6.0. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.15) ATPS Phosphate Buffer

This buffer was prepared using 10 mM $NaH_2PO_4$ (Sigma-Aldrich) and the pH adjusted to 8.0. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.16) Sodium Acetate Buffer pH 4.0

Solution A was prepared by diluting 11.55 ml glacial acetic acid (Sigma-Aldrich) per litre of $H_2O$. Solution B was prepared by dissolving 27.2 g sodium acetate (Sigma-Aldrich) per litre of $H_2O$. An aliquot of solution A (41 ml) was added to 9 ml of solution B to produce a 0.1 M buffer solution that is adjusted to pH 4.0.

### (2.2.17) Tris pH 9.0 Conjugation Buffer

This buffer was prepared using 20 mM Tris (Sigma-Aldrich) and the pH adjusted to 9.0. The solution was then filter sterilised using a 0.2 µm syringe filter (Nalgene).

### (2.2.18) 10 x Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS PAGE) Running Buffer

A 1 litre stock solution of 10 x Running Buffer was prepared using 30 g Tris, 144 g Glycine and 10 g SDS in distilled $H_2O$. When necessary this was diluted in distilled $H_2O$ to produce a working 1 x buffer solution.

### (2.2.19) 4 x Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS PAGE) Sample Buffer

The protein sample buffer was produced by combining 20 ml 0.5 M Tris (pH 6.8), 20 ml glycerol, 2 g SDS, 20 mg Bromophenol Blue and 2 ml β-mercaptoethanol.

## 2.3 Plasmid DNA

### (2.3.1) pGFPuv

pGFPuv was obtained from Clontech Laboratories Inc. Oligonucleotides described in section (2.4.1) and (2.4.2) were used to amplify the gene encoding Green Fluorescent Protein (GFPuv) out of this plasmid with the required restriction enzyme sites for insertion into fusion plasmids.

### (2.3.2) pGFPuv1KL4

pGFPuv1KL4 was very kindly donated by Dr David Nagel, Aston University. Oligonucleotides described in section (2.4.1) and (2.4.2) were used to amplify the gene encoding Green Fluorescent Protein (GFPuv) out of this plasmid with the required restriction enzyme sites for insertion into fusion plasmids.

## 2.4 Oligonucleotides

### (2.4.1) GFPuv Smal Blunt

5’ CCCGGGTTATGAGTAAAGGAGAAGAAC 3’ obtained from Eurofins MWG Operon. Double distilled $H_2O$ was added and the solution to obtain a final concentration of 100 pmoles / µl and mixed by vortex for 10seconds. The solution was then mixed on a flat rollerbed for 10 minutes before being stored at -20 ºC.

### (2.4.2) GFPuv EcoRI

5’ GAATTCATTATTTGTAGAGC 3’ obtained from Eurofins MWG Operon. Double distilled $H_2O$ was added and the solution to obtain a final concentration of 100 pmoles / µl and mixed by vortex for 10seconds. The solution was then mixed on a flat rollerbed for 10 minutes before being stored at -20 ºC.

## 2.5 Proteins

### (2.5.1) Purified His-tagged Interferon-α-2a (IFN α-2a)

Purified $His_8$-tagged IFN α-2a was supplied by Polytherics Ltd at a concentration of 0.5 mg / ml in a buffer containing 20 mM Tris HCl, 100 mM NaCl and 10% glycerol. This protein was used in experiments as supplied.

## 2.6 Electrophoresis

Electrophoresis is a common methodology used in molecular biology to separate small charged molecules. In this body of work, two electrophoresis methods have been employed; agarose gel electrophoresis to separate different length strands of Deoxyribose Nucleic Acid (DNA) and sodium dodecyl sulfate poly-acrylamide gel electrophoresis (SDS PAGE) to separate proteins of different molecular weights. Agarose gel electrophoresis was used to not only separate strands of DNA for isolation purposes, but also to identify whether genetic material from *E.coli* used to produce the target proteins was present in the different ATPS layers. SDS PAGE was used to determine successful production of the target protein conjugated to its fusion partner protein (e.g. glutathione *s*-transferase) as well as the location of proteins within the different ATPS layers.

### (2.6.1) Agarose Gel Electrophoresis

Agarose gels were made by dissolving molecular biology grade agarose in 1 x TAE Buffer (2.2.1) and stained using ethidium bromide at a concentration of 0.5 µg / ml. The DNA and loading buffer (2.2.2) were then pipetted into the wells at the top of the gel and a current passed through at 10 V / cm until the dye front had travelled two-thirds of the gel length. The DNA could then be visualised by exposure to UV light and photographed using a UV transilluminator (UVP Products).

### (2.6.2) SDS Poly-Acrylamide Gel Electrophoresis (PAGE)

To make a 10% resolving gel (5 ml), 1.25 ml of 40% acrylamide / bisacrylamide solution (Sigma-Aldrich) was combined with 1.25 ml of 50% glycerol, 1.25 ml of 4 x Tris pH 8.8, 1.15 ml of distilled $H_2O$, 50 µl of 10% SDS and 50 µl of 10% APS. This solution was mixed well with 4.5 µl TEMED, overlaid with 0.5 ml 100% isopropanol and allowed to set within the sealed glass plates. A 5% stacking gel (2 ml) was made using 0.25 ml of 40% acrylamide /

bisacrylamide solution (Sigma-Aldrich) combined with 0.5 ml of 50% glycerol, 0.5 ml of 4 x Tris pH 6.5, 0.71 ml of distilled $H_2O$, 20 µl of 10% SDS and 20 µl of 10% APS. This solution was mixed well with 2 µl TEMED and after removal of the 100% isopropanol, poured above the resolving gel up to approximately 2 cm below the rim of the lower glass plate. The comb was inserted into the gel which was allowed to set. When set, the gels are transferred into the gel tank containing 1 x running buffer (2.2.18). Samples to be run on the gel were heated at 95°C for 5 minutes in 4 x sample buffer (2.2.19) and cooled to room temperature. Once loaded onto the gel with an appropriate molecular weight marker, the gel was run at 120 V until the dye front had moved into the resolving gel when the voltage was increased to 150 V. When ready the gels were carefully disassembled from the glass plates ready for coomassie staining or western blotting.

## 2.7 Polymer Solutions

Careful consideration was taken when choosing polymers for this study. Not only was it important to ensure that two phases would be formed when the polymers were mixed in water, but biocompatibility had to be considered to allow for this protein purification method to be suitable for the production of therapeutic proteins. Poly(ethylene) glycol and dextran were both chosen due to their use as excipients in the pharmaceutical industry to provide long term stabilisation and bulking properties (Alpar, EK *et al.,* 2004). Dextran is a polysaccharide made from many glucose monomers and is regularly used medically as a volume expander in hypovolaemia (Bulger, E., 2011). PEG and Dextran are both approved by the EU and FDA for use in both food (http://www.food.gov.uk/science/additives/enumberlist accessed 30/07/16) and medicines (Alconcel, S., *et al.,* 2011)

### (2.7.1) Poly(ethylene) Glycol

Poly(ethylene) Glycol (PEG) of varying molecular weights was obtained from Sigma-Aldrich and dissolved into double distilled water at the required concentration. PEG molecular weights used were as follows: 300 Da, 600 Da, 1000 Da, 1500 Da, 2000 Da, 4000 Da, 5000 Da, 10000 Da and 20000 Da.

### (2.7.2) Dextran

Dextran of varying molecular weights was obtained from Sigma-Aldrich and dissolved into double distilled water at the required concentration. Dextran molecular weights used were as follows: 6 kDa, 40 kDa, 70 kDa, 100 kDa, 200 kDa, 500 kDa and 2000 kDa.

## 2.8 General Techniques

### (2.8.1) Transformation of *E.coli* Tuner (DE3) Cells using $CaCl_2$ Method

A 30 ml starter culture was prepared by selecting a single *E.coli* Tuner colony from a fresh overnight streak plate and incubated overnight at 37°C in a shaking incubator. The cells were then harvested by centrifugation at 4000 rpm for 5 minutes after which the supernatant was discarded and any excess media drained away. The pellet was re-suspended in 6 ml of ice cold 50 mM $CaCl_2$ (2.2.6) by gentle vortexing and then kept on ice for 10 minutes. Centrifugation was repeated and the pellet again re-suspended in 6 ml 50 mM $CaCl_2$ and left to stand on ice for 10 minutes after which the centrifugation was repeated a third time. The pellet was then re-suspended in 1.2 ml of 50 mM $CaCl_2$ and kept on ice for 20 minutes, during which time 1.5 ml microfuge tubes were chilled on ice in preparation for the transformation. 5 ng of DNA to be transformed was prepared in 1.5 ml microfuge tubes on ice and 100 µl of *E.coli* Tuner cells were added. The tubes were mixed well and left on ice for 30 minutes, after which they were placed in a water bath at 37°C for 30 seconds to 1 minute and then returned to ice for a further 2 minutes. 0.5 ml LB Broth (2.1.1) was then added to the tubes which were then incubated in a 37°C water bath for 40 minutes. 100 µl and 200 µl aliquots of transformed cells were then plated out onto selective LB Agar (2.1.2) and incubated overnight at 37°C.

### (2.8.2) Expression of Protein in *E.coli* Tuner (DE3) Cells

A single colony that had been freshly transformed with the desired plasmid DNA was selected and used to inoculate a 30 ml selective LB Broth starter culture (2.1.1) and was then incubated overnight in a 37°C shaking incubator. 5 ml of this overnight culture was then used to inoculate a 200 ml selective LB Broth and incubated at 37°C in a shaking incubator until an optical density of between 0.45 and 0.55 was achieved when analysed at a wavelength of 600 nm. At this point 1 mM IPTG (2.2.12) was added, in order to induce the expression of the

desired protein, and the culture then incubated at 30°C in a shaking incubator until the desired amount of protein has been produced and the cells harvested by centrifugation.

### (2.8.3) His-Tag Purification of GFPuv

His-tag purification of GFPuv was performed using Ni-NTA resin (QIAgen) and the protocol supplied by the manufacturer for purification from *E.coli* cell lysates under native conditions.

### (2.8.4) Initial formation of ATPS

ATPS were formed by combining 0.5 ml of the required dextran solution and 0.5 ml of the required PEG solution in a 1.5 ml microfuge tube. The sample to be examined (100 µl) was then added at a known protein concentration (where relevant) and the system mixed by vortexing. The systems were then either left to settle for 30 minutes or centrifuged at 2000 g for 2 minutes at room temperature. Systems were then checked for a visible meniscus and those with a meniscus present were further analysed for protein and genomic material-partitioning behaviours.

### (2.8.5) PEG M1a PEGylation of GSH

PEGylation was performed by first combining 1.4 equivalent moles of GSH with 1 equivalent mole of PEG M1a (Polytherics) in 3 ml of PEGylation reaction buffer as described in section (2.2.13) and left to react at room temperature for 18 hours and then freeze-dried overnight. Acetone precipitation of the PEGylated GSH was performed at -80°C on the residue to remove any un-reacted GSH. At each stage of acetone precipitation, filtration was also performed in order to remove any residual buffer salts. $^{1}H$ NMR spectroscopy was performed on 10 mg of precipitate in $D_2O$ to determine the product obtained as described (2.8.17).

For 5 kDa PEGylated GSH

$^1$H NMR (400 MHz, $D_2O$) δ 8.0 2H (d), 7.8 2H (d), 4.5 2H (d), 3.8 - 3.3 498H (methylene envelope), 3.25 3H (s), 3.1 1H (dd), 2.9 3H (m), 2.5 2H (m), 2.1 2H (m).

For 10 kDa PEGylated GSH

$^1$H NMR (400 MHz, $D_2O$) δ 8.0 2H (d), 7.8 2H (d), 4.5 2H (d), 3.8 - 3.3 640H (methylene envelope), 3.25 3H (s), 3.1 1H (dd), 2.9 3H (m), 2.5 2H (t), 2.1 2H (q).

### (2.8.6) Methoxy PEG Maleimide PEGylation of lysine NTA

PEGylation was performed by first combining 5 equivalent moles of lysine NTA (Sigma, 14580) with 1 equivalent mole of methoxy PEG maleimide (Jenkem USA Ltd.) in 5 ml of Sodium Phosphate buffer pH 6.0 (2.2.14) and left to react at room temperature for 24 hours. The conjugates were then separated from any unreacted lysine NTA and simultaneously exchanged into Sodium Acetate buffer pH 4.0 (2.2.16) using a PD10 desalting column (GE Healthcare) in preparation for copper coordination. Absorbance at $A^{280}$ of each 1 ml elution from the PD10 column was measured using a quartz cuvette and the appropriate elutions were pooled. $^1$H NMR spectroscopy was performed on 10 mg of conjugate in $D_2O$ to determine the product obtained as described (2.8.17).

For Lysine NTA

$^1$H NMR (400 MHz, $D_2O$) δ 3.8 5H (m), 3.0 2H (t), 2.0 – 1.8 2H (m), 1.75 – 1.5 4H (m).

For 10 kDa Methoxy PEG Maleimide PEG

$^1$H NMR (400 MHz, $D_2O$) δ 6.8 1H (s), 3.9 1H (m), 3.8 - 3.4 720H (methylene envelope), 3.3 3H (s),

### (2.8.7) Conjugation of methoxy PEG amine to B1 Linker

Prior to conjugation with the B1 linker, azeotropic drying of methoxy PEG amine was performed using 5 ml toluene under vacuum. Conjugation was then performed by combining 4 equivalent moles of carboxylic DB-*bis*-sulfide linker (B1, obtained from Prof. Steve Brocchini, UCL) with 1 equivalent mole of Methoxy PEG amine (Jenkem USA Ltd.) in 5 ml Anhydrous Dichloromethane avoiding contact with air as much as possible. The flask was purged with Argon and, once sealed with parafilm, left to stir for 48 hours at room temperature. After the incubation period, the Anhydrous Dichloromethane was removed using rotary evaporation and the resulting precipitate dissolved into 5 ml acetone. This solution was filtered through cotton wool into a pre-weighed 50 ml falcon tube and placed onto dry ice for 5 minutes to precipitate the PEG conjugate. Centrifugation was performed at 4000 rpm for 30 minutes at -9°C and the supernatant discarded. The acetone precipitation and centrifugation steps were repeated twice more in order to remove any unreacted linker. After the final centrifugation step and the supernatant discarded, a needle was used to puncture the cap of the falcon tube and placed into a desiccator overnight. $^1$H NMR spectroscopy was performed on 10 mg of precipitate in $D_2O$ to determine the product obtained as described (2.8.17).

For 10 kDa PEG B1 Sulfide

$^1$H NMR (400 MHz, $D_2O$) δ 7.7 2H (d), 7.4 2H (d), 7.0 4H (s), 3.8 - 3.4 848H (methylene envelope), 3.3 3H (s), 3.1 1H (s), 2.2 3H (s), 1.3 – 1.1 3H (s).

### (2.8.8) Oxidation of mPEG B1

1 molar equivalent of mPEG B1 produced in method (2.8.7) was dissolved into 2.5 ml methanol in a round bottomed flask before the addition of 6 molar equivalents of Oxone (Sigma, 228036), 2.5 ml $H_2O$, a magnetic stirrer and left overnight at room temperature. The insoluble Oxone was filtered away through cotton wool and the filtrate collected. The $H_2O$ and methanol was evaporated away using rotary evaporation and the residual precipitate dissolved into 5 ml acetone. This solution was filtered through cotton wool into a pre-weighed 50 ml falcon tube and placed onto dry ice for 5 minutes to precipitate the PEG conjugate. Centrifugation was performed at 4000 rpm for 30 minutes at -9°C and the supernatant discarded. The acetone precipitation and centrifugation steps were repeated twice more in order to remove any unwanted compounds. After the final centrifugation step and the supernatant discarded, a needle was used to puncture the cap of the falcon tube and placed into a desiccator overnight. $^1$H NMR spectroscopy was performed on 10 mg of precipitate in $D_2O$ to determine the product obtained as described (2.8.17).

For 10 kDa PEG B1 Sulfone

$^1$H NMR (400 MHz, $D_2O$) δ 7.7 2H (d), 7.5 2H (d), 7.4 4H (m), 3.9 - 3.5 583H (methylene envelope), 3.45 3H (s), 2.4 2H (s).

### (2.8.9) Conjugation of mPEG B1 Sulfone to lysine NTA

PEGylation was performed by first combining 5 equivalent moles of lysine NTA (Sigma, 14580) with 1 equivalent mole of mPEG B1 Sulfone (2.8.8) in 5 ml of Tris buffer pH 9.0 (2.2.17) and left to react at 37°C overnight. The conjugates were then separated from any unreacted lysine NTA and simultaneously exchanged into Sodium Acetate buffer pH 4.0 (2.2.16) using a PD10 desalting column (GE Healthcare) in preparation for copper coordination. Absorbance at $A^{280}$ of each 1 ml elution from the PD10 column was measured using a quartz cuvette and

the appropriate elutions pooled. $^{1}$H NMR spectroscopy was performed on 10 mg of conjugate in $D_2O$ to determine the product obtained as described (2.8.17).

For 10 kDa PEG B1 lysine NTA

$^{1}$H NMR (400 MHz, $D_2O$) δ 7.5 2H (d), 7.35 2H (d), 3.9 - 3.4 1264H (methylene envelope), 3.3 3H (s), 3.0 4H (m), 2.0 – 1.75 4H (m), 1.75 – 1.35 8H (m).

### (2.8.10) Conjugation of amine PEG amine to B1 Linker

Prior to conjugation with the B1 linker, azeotropic drying of amine PEG amine was performed using 5 ml toluene under vacuum. Conjugation was then performed by combining 8 equivalent moles of carboxylic DB-*bis*-sulfide linker (B1, obtained from Prof. Steve Brocchini, UCL) with 1 equivalent mole of amine PEG amine (Jenkem USA Ltd.) in 5 ml Anhydrous Dichloromethane avoiding contact with air as much as possible. The flask was purged with Argon and, once sealed with parafilm, left to stir for 48 hours at room temperature. After the incubation period, the Anhydrous Dichloromethane was removed using rotary evaporation and the resulting precipitate dissolved into 5 ml acetone. This solution was filtered through cotton wool into a pre-weighed 50 ml falcon tube and placed onto dry ice for 5 minutes to precipitate the PEG conjugate. Centrifugation was performed at 4000 rpm for 30 minutes at -9°C and the supernatant discarded. The acetone precipitation and centrifugation steps were repeated twice more in order to remove any unreacted linker. After the final centrifugation step and the supernatant discarded, a needle was used to puncture the cap of the falcon tube and placed into a desiccator overnight. $^{1}$H NMR spectroscopy was performed on 10 mg of precipitate in $D_2O$ to determine the product obtained as described (2.8.17).

### (2.8.11) Oxidation of PEG *Bis*-B1

1 molar equivalent of PEG *Bis*-B1 produced in method (2.8.10) was dissolved into 2.5 ml methanol in a round bottomed flask before the addition of 12 molar equivalents of Oxone (Sigma, 228036), 2.5 ml $H_2O$, a magnetic stirrer and left overnight at room temperature. The insoluble Oxone was filtered away through cotton wool and the filtrate collected. The $H_2O$ and methanol was evaporated away using rotary evaporation and the residual precipitate dissolved into 5ml acetone. This solution was filtered through cotton wool into a pre-weighed 50 ml falcon tube and placed onto dry ice for 5 minutes to precipitate the PEG conjugate. Centrifugation was performed at 4000 rpm for 30 minutes at -9°C and the supernatant discarded. The acetone precipitation and centrifugation steps were repeated twice more in order to remove any unwanted compounds. After the final centrifugation step and the supernatant discarded, a needle was used to puncture the cap of the falcon tube and placed into a desiccator overnight. $^1$H NMR spectroscopy was performed on 10 mg of precipitate in $D_2O$ to determine the product obtained as described (2.8.17).

### (2.8.12) Conjugation of PEG *Bis*-B1 Sulfone to lysine NTA

PEGylation was performed by first combining 10 equivalent moles of lysine NTA (Sigma, 14580) with 1 equivalent mole of mPEG *bis*-B1 Sulfone (2.8.11) in 5 ml of Tris buffer pH 9.0 (2.2.17) and left to react at 37°C overnight. The conjugates were then separated from any unreacted lysine NTA and simultaneously exchanged into Sodium Acetate buffer pH 4.0 (2.2.16) using a PD10 desalting column (GE Healthcare) in preparation for copper coordination. Absorbance at $A^{280}$ of each 1 ml elution from the PD10 column was measured using a quartz cuvette and the appropriate elutions pooled. $^1$H NMR spectroscopy was performed on 10 mg of conjugate in $D_2O$ to determine the product obtained as described (2.8.17).

### (2.8.13) Coordination of Copper Ions to PEGylated lysine NTA compounds

Copper Sulfate ($CuSO_4$) was added to pooled fractions of the PEGylated NTA compounds that were buffer exchanged into Sodium Acetate buffer pH 4.0 (2.2.16). Differing molar equivalents were added to the different PEGylated lysine NTA compounds available (5 molar equivalents of $CuSO_4$ for mPEG Maleimide, 10 molar equivalents of $CuSO_4$ for mPEG B1 and 20 molar equivalents of $CuSO_4$ for PEG *bis*-B1 lysine NTA). The coordination reactions were left at room temperature overnight and buffer exchanged into ATPS phosphate buffer pH 8.0 (2.2.15) in preparation for use in ATPS.

### (2.8.14) Bradford assay Standard Method

Bradford reagent (Bio-Rad Laboratories Inc.) was diluted 1 part dye to 4 parts distilled $H_2O$ and filtered through filter paper. Bovine Serum Albumin (BSA) standards were produced in triplicate with concentrations ranging between 0.05 mg / ml to 0.5 mg / ml. 10 µl of the sample to be analysed was mixed with 200 µl of diluted Bradford reagent, comprising coomassie brilliant blue dye, phosphoric acid and methanol, followed by a 5 minute incubation at room temperature. A spectrophotometer was used to read the absorbance of the samples at a wavelength of 595 nm as described by Bradford, M.M., 1976.

### (2.8.15) Bradford assay Microplate Method

Bradford reagent (Bio-Rad Laboratories Inc.) was diluted 1 part dye to 4 parts distilled $H_2O$ and filtered through filter paper. Bovine Serum Albumin (BSA) standards were produced in triplicate with concentrations ranging between 8 µg / ml to 80 µg / ml. 160 µl of the sample to be analysed was mixed with 40 µl of diluted Bradford, comprising coomassie brilliant blue dye, phosphoric acid and methanol, followed by a 5 minute incubation at room temperature. A spectrophotometer was used to read the absorbance of the samples at a wavelength of 595 nm as described by Bradford, M.M., 1976.

### (2.8.16) ATPS Formation with Affinity Conjugates

ATPS (1.5 g systems) were formed by weighing out the required molecular weight dextran and molecular weight PEG polymers into a 1.5 ml microfuge tube. To this was added the appropriate volume of known concentration functionalised PEG and 100 µl of the sample to be examined. Finally ATPS phosphate buffer (2.2.15) was added to produce a final system weight of 1.5 g. The systems were mixed by vortexing and were then left to settle for 30 minutes at room temperature. Systems were then checked for a visible meniscus and those with a meniscus present were further analysed for protein partitioning behaviours.

### (2.8.17) $^{1}$H NMR Spectroscopy

$^{1}$H NMR spectroscopy was carried out using a Bruker 400 spectrometer. Chemical shifts are measured in ppm, and multiplicity denoted as follows: s = singlet, d = doublet, dd = double doublet, t = triplet, q = quartet, m = multiplet.

## 2.9 Enzyme Techniques

### (2.9.1) SmaI, Antarctic Phosphatase and EcoRI Digestion Protocol

One unit of SmaI was used to digest 1 µg DNA in 1 hour at 25 ºC in a total reaction volume of 50 µl. This was followed by Antarctic Phosphatase and then subsequent heat inactivation at for 20 mins at 65 ºC. The product was purified directly from the enzyme buffer mixture using a GeneClean purification kit which utilises GLASSMILK® nucleic acid binding matrix. The purified product was subsequently digested using one unit of EcoRI (New England BioLabs Inc.) to digest 1 µg DNA in 1 hour at 25 ºC in a total reaction volume of 50 µl, followed by heat inactivation at 65 ºC for 20 mins.

### (2.9.2) BamHI Digestion Protocol

BamHI (New England BioLabs Inc.) was used at 1/10$^{th}$ of the concentration recommended by the manufacturers and for only a 10 minute incubation period at 37 ºC to produce incomplete digestion. As heat inactivation is not effective on BamHI, the product was quickly purified using directly from the enzyme buffer mixture using a GeneClean purification kit which utilises GLASSMILK® nucleic acid binding matrix.

### (2.9.3) BamHI, Mung Bean Nuclease and EcoRI Digestion Protocol

One unit of BamHI (New England BioLabs Inc.) was used initially for one hour at 37 ºC in a total reaction volume of 50 µl, followed by the addition of 4.5 units of Mung Bean Nuclease (New England BioLabs Inc.) and incubated at 30°C for 30 minutes. After purification of the product using a GeneClean purification kit, one unit of EcoRI (New England BioLabs Inc.) was used to digest the product in a total reaction volume of 50 µl at 37 ºC, followed by heat inactivation for 20 mins at 65 ºC.

### (2.9.4) T4 DNA Ligase Ligation Protocol

One unit of T4 DNA Ligase (New England BioLabs Inc.) was used with the supplied buffer. Reactions were left overnight at 16°C in a total reaction volume of 20 µl.

### (2.9.5) Polymerase Chain Reaction (PCR) Protocol

PCR was performed using Pfu Polymerase (New England BioLabs Inc.) or Taq Polymerase (New England BioLabs Inc.) at 75 ºC in 1 x standard reaction buffer. Primers used in this instance are described in section 2.4.

# Chapter 3. Behaviour of *E.coli* cell lysates in poly(ethylene glycol) (PEG) and dextran (DEX) aqueous two phase systems (ATPS)

## 3.1 Introduction

ATPS can be formed using varying concentrations of polymer solutions that are sufficiently disparate to produce a difference in water concentration around the polymer chains (Albertsson, 1986.). In this work poly(ethylene glycol) (PEG) and dextran (DEX) systems were systematically analysed to find the optimal concentrations and molecular weights of these two polymers to first form an interface and then two distinct phases to give an ATPS. Secondly, cell lysate was added to the resulting ATPS. The PEG phase was the upper or top phase, and DEX phase was the lower phase of the ATPS. Optimisation of the ATPS was conducted in an effort to partition the cell lysate components to the lower DEX phase. Protein detection was performed using the Bradford assay and visually determined using the location of fluorescence from green fluorescent protein (GFPuv) contained within the cell lysate. The location within the ATPS of cellular DNA and RNA was analysed using ethidium bromide staining of agarose gels obtained after electrophoresis.

## 3.2 Formation of ATPS

To determine which combinations of PEG and Dextran would form an ATPS, a range of molecular weight PEG and DEX, in 10% (w/v), 20% (w/v) and 25% (w/v) concentrations were combined as described in section (2.8.4). These systems (Tables 3.1 A-C) were allowed to settle for 30 minutes at room temperature.

| A | PEG MW | | | | | | |
|---|---|---|---|---|---|---|---|
| Dextran MW | 300 | 600 | 1000 | 1500 | 2000 | 4000 | 20000 |
| 6 | | | | | | | |
| 40 | | | | | | | |
| 70 | | | | | | | |
| 100 | | | | | | | |
| 200 | | | | | | | |
| 500 | | | | | | | |
| 2000 | | | | | | | |

| B | PEG MW | | | | | | |
|---|---|---|---|---|---|---|---|
| Dextran MW | 300 | 600 | 1000 | 1500 | 2000 | 4000 | 20000 |
| 6 | | | | | | | |
| 40 | | | | | | | |
| 70 | | | | | | | |
| 100 | | | | | | | |
| 200 | | | | | | | |
| 500 | | | | | | | |
| 2000 | | | | | | | |

| C | PEG MW | | | | | | |
|---|---|---|---|---|---|---|---|
| Dextran MW | 300 | 600 | 1000 | 1500 | 2000 | 4000 | 20000 |
| 6 | | | | | | | |
| 40 | | | | | | | |
| 70 | | | | | | | |
| 100 | | | | | | | |
| 200 | | | | | | | |
| 500 | | | | | | | |
| 2000 | | | | | | | |

**Table 3.1: Formation of ATPS using various concentrations of PEG and dextran solutions.** ATPS were formed by settling, as described (2.8.4). All Experiments were performed in triplicate. Grey shading indicates the successful formation of two phases. **A,** 10% (w/v) polymer solutions; **B**, 20% (v) polymer solutions; **C**, 25% (w/v) polymer solutions.

As can be seen from Table 3.1 A-C, there was a similar trend in all cases, where combinations of higher molecular weight PEG and DEX solutions successfully formed ATPS (there was a visible meniscus), whilst polymer mixtures of lower MW did not,

although it is noticeable that higher concentrations of polymer prompted successful ATPS formation with the polymer solutions of lower MW. However, the 25% (w/v) polymer solutions had sufficiently high viscosity to make accurate pipetting problematic. Since 10% (w/v) polymer solutions gave fewer options for ATPS formation, 20% (w/v) polymer solutions were chosen for use in future studies.

## 3.3 Protein assays in the presence of PEG and/or Dextran polymer solutions

Prior to analysing partitioning behaviour in 20% ATPS, it was necessary to determine whether the Bradford assay would function in the presence of ATPS components and/or be affected by those components. Accordingly, Bovine Serum Albumin (BSA) standard solutions were measured by the Bradford assay (2.8.14) in the presence or absence of 10 µl of 20% DEX 100 and 20% PEG 2000 (an 80 fold dilution).

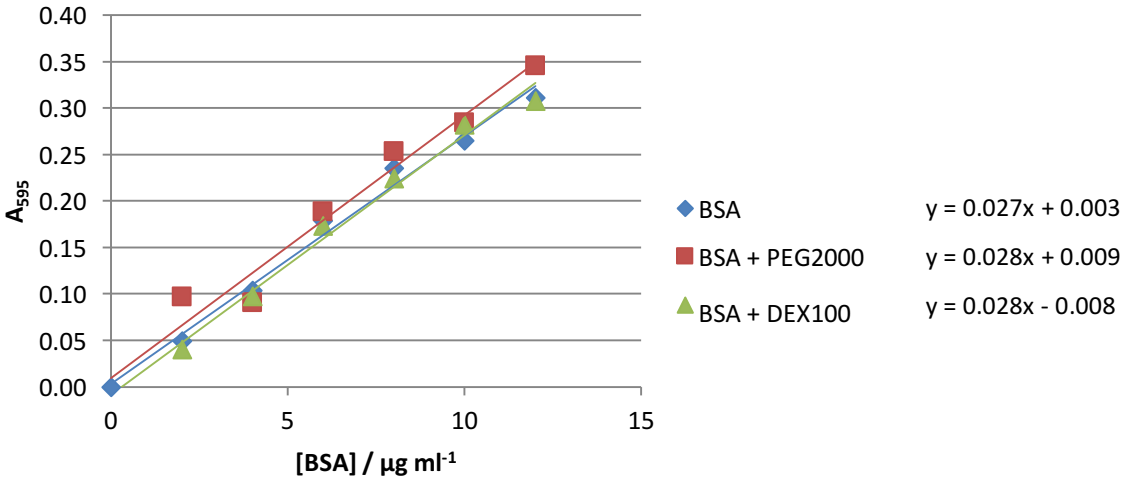



**Figure 3.1: Comparison of Bradford assay performance in the presence / absence ATPS components.** BSA stock solutions were quantified as described (2.8.1) with or without 0.25% (w/v) PEG 2000 or 0.25% (w/v) DEX 100, where 0.25% represents an 80-fold dilution of the standard ATPS component concentration.

Figure 3.1 confirms the earlier finding of Barbosa *et al* (2009) that these ATPS reagents have only a very minor influence on Bradford assay results at the diluted concentrations required for the assay. Further analysis of each of the individual polymer solutions used in making the various ATPS combinations showed a similarly low absorbance in the presence of Bradford reagent (Table 3.2).

| Polymer | Optical Density (595nm) |
|---|---|
| PEG 1500 | 0.016 |
| PEG 2000 | 0.022 |
| PEG 4000 | 0.018 |
| PEG 20000 | 0.020 |
| Dextran 40 | 0.009 |
| Dextran 70 | 0.012 |
| Dextran 100 | 0.010 |
| Dextran 200 | 0.012 |
| Dextran 500 | 0.007 |
| Dextran 2000 | 0.009 |

**Table 3.2: Optical densities obtained by Bradford analysis of various 80-fold diluted polymer solutions (0.25%, w/v).** Bradford analyses of each diluted solution were performed as described (2.8.14).

## 3.4 ATPS performance in the presence of cell lysates: settling versus centrifugation

Many examples are described in the literature of ATPS formed by either centrifugation and or by settling, however studies comparing the two methods are not available. To determine which method is superior for the PEG-DEX system, it was necessary to investigate the difference in the partitioning behaviour between systems that were left to settle and systems that were centrifuged. Since the ultimate aim of this study is to purify proteins directly from cell lysates, ATPS performance was initially assessed using cleared *E. coli* cell lysates. To facilitate evaluation, lysates containing expressed GFPuv were studied.

### 3.4.1 Comparison of ATPS formation from 20% polymer solutions in the presence of cell lysate, either by settling or centrifugation.

It had previously been demonstrated that 20% solutions of PEG and DEX form an ATPS using various combinations of high MW polymers (Table 3.1 B). ATPS formation might be affected by the presence of *E. coli* lysate and/or by the experimental protocol adopted. The ATPS formation experiment (section 3.2) with 20% (w/v) polymer solutions was therefore repeated in the presence of 100 µl *E. coli* lysate containing GFPuv, by both settling and centrifugation methodologies (Table 3.3).

| A | PEG MW | | | | | | |
|---|---|---|---|---|---|---|---|
| Dextran MW | 300 | 600 | 1000 | 1500 | 2000 | 4000 | 20000 |
| 6 | | | | | | | |
| 40 | | | | | | | |
| 70 | | | | | | | |
| 100 | | | | | | | |
| 200 | | | | | | | |
| 500 | | | | | | | |
| 2000 | | | | | | | |

| B | PEG MW | | | | | | |
|---|---|---|---|---|---|---|---|
| Dextran M | 300 | 600 | 1000 | 1500 | 2000 | 4000 | 20000 |
| 6 | | | | | | | |
| 40 | | | | | | | |
| 70 | | | | | | | |
| 100 | | | | | | | |
| 200 | | | | | | | |
| 500 | | | | | | | |
| 2000 | | | | | | | |

**Table 3.3: Comparison of methodologies for formation of ATPS containing *E. coli* cell lysate.** ATPS were formed either by settling (**A**) or by centrifugation (**B**). Each ATPS contained cleared *E.coli* Cell lysate (100µl) containing expressed GFPuv. Total protein concentration within the lysate was 730 µg/ml, as determined by the Bradford method (2.8.14). Green colouration indicates successful ATPS formation.

It is interesting to note that addition of cell lysate to the ATPS reduced the overall number of polymer combinations that formed an ATPS (Table 3.1B and Table 3.3 A-B). Specifically, combinations with the lower molecular weight polymers failed to form ATPS in the presence of cell lysate. However, results shown in Table 3.3 suggest that there is little difference between the two methodologies in terms of the successful formation of an ATPS. In fact, by comparing Tables 3.3 A and B, it can be seen that only one

additional ATPS, namely PEG 2000 and DEX 200 was formed when settling rather than centrifugation methodology was employed.

### 3.4.2 Comparison of ATPS partitioning performance from 20% polymer solutions in the presence of cell lysate, either by settling or centrifugation.

Those systems which had successfully formed ATPS in section 3.4.1 were further analysed to examine partitioning performance in terms of both protein and nucleic acid separation. Simple observation of such systems demonstrated that all visually-detectable GFPuv had partitioned to the lower, DEX layer in all cases. However, GFPuv partitioning behaviour is not necessarily indicative of total cell protein behaviour. Therefore, both the top (PEG) and bottom (DEX) layers of the formed ATPS, shown in green in Table 3.3, were analysed for total protein concentration using the Bradford assay (2.8.14; Figure 3.2).

A standard curve for the Bradford assay was produced to allow for the determination of protein concentration when an optical density reading was obtained for a sample (Figure 3.2).

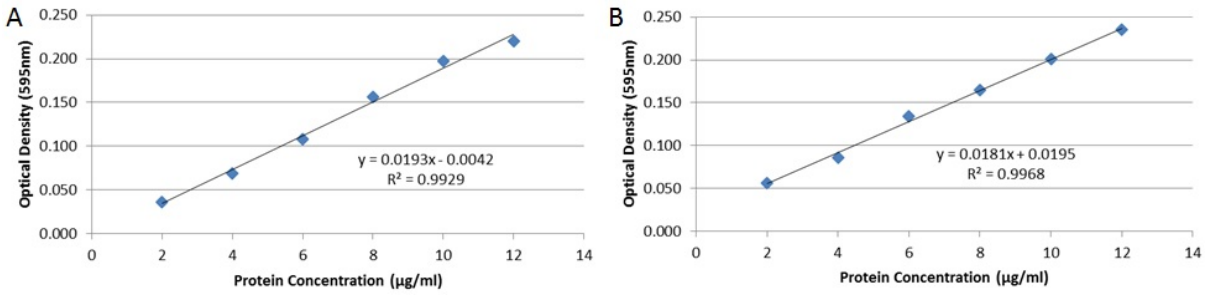



**Figure 3.2: Bradford assay standard curve for settled (A) and centrifuged (B) systems.** The standard was produced using known concentrations of BSA prepared from stock solution as described in section (2.2.8) of “Materials and Methods”

The optical densities recorded for each phase were analysed using the equations obtained from the Bradford assay standard curves shown above in Figure 3.2. An additional value for polymer interference, as determined from Table 3.2, was subtracted from the appropriate optical density so that the percentage of protein partitioning in that layer could be estimated (Figure 3.3).

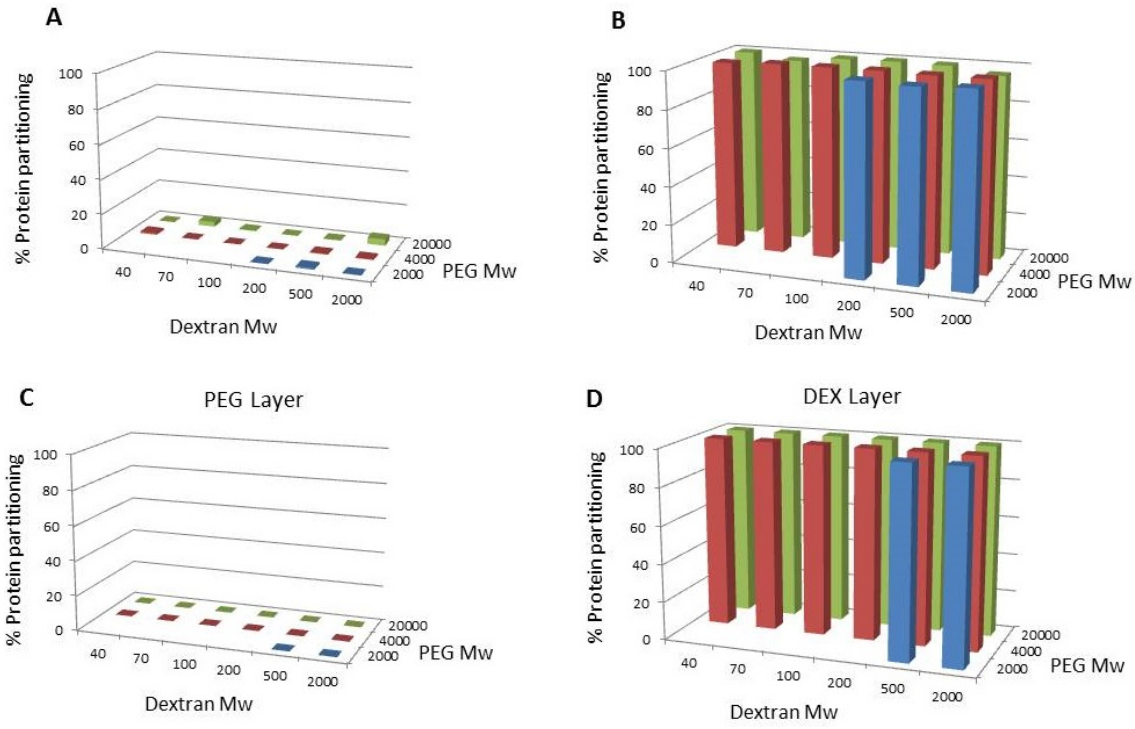



**Figure 3.3: Phase partitioning of cellular protein to upper PEG (A) and lower DEX (B) phases when utilising the settling methodology.** The phase partitioning to upper PEG (C) and lower DEX (D) phases of centrifuged systems are also shown (2000g for 2 minutes).

In Figure 3.3 above it can be seen that there is near complete protein partitioning to the lower DEX phase with only two systems (Figure 3.3 A), namely PEG 20000 / DEX 70 and PEG 20000 / DEX 2000 having proteins present in the upper PEG phase when formed using the settling methodology. This indicates that there is little difference between the two methodologies of settling and centrifugation when partitioning proteins within a cell lysate.

To further analyse the molecular partitioning characteristics of these PEG-DEX systems, DNA partitioning analysis was performed using agarose gel electrophoresis (2.6.1). ATPS that were formed using cell lysate (Figure 3.4) were used in these experiments.

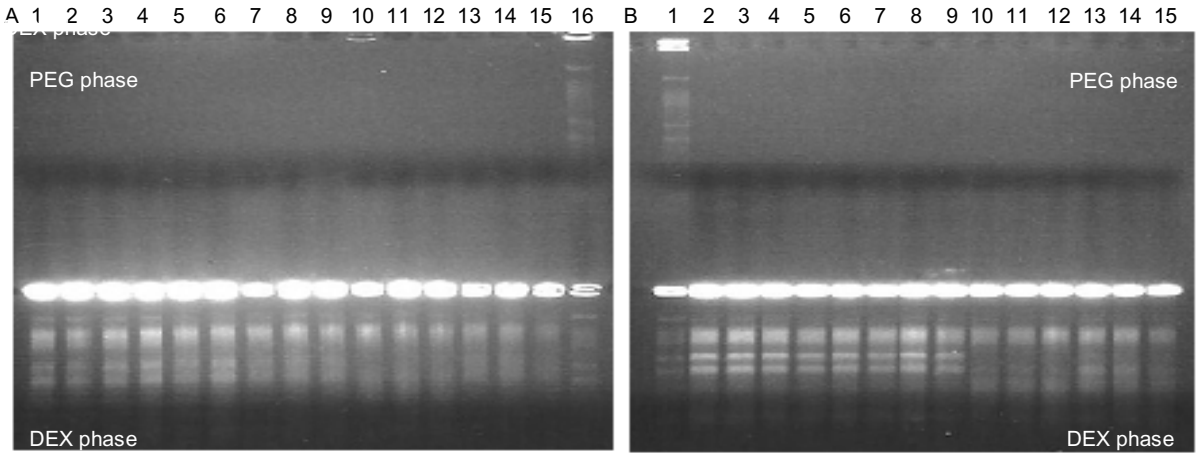



**Figure 3.4: Agarose electrophoresis gels showing the partitioning of genomic DNA in ATPS formed by settling (A) and centrifugation (B).** 10µl of each ATPS layer was electrophoresed on a 1% agarose gel that was pre-stained with ethidium bromide as described in section (2.6.1) of "Materials and Methods". Lanes A1: PEG 2000 / DEX 200 ATPS, A2: PEG 2000 / DEX 500 ATPS, A3: PEG 2000 / DEX 2000 ATPS, A4: PEG 4000 / DEX 40 ATPS, A5: PEG 4000 / DEX 70 ATPS, A6: PEG 4000 / DEX 100 ATPS, A7: PEG 4000 / DEX 200 ATPS, A8: PEG 4000 / DEX 500 ATPS, A9: PEG 4000 / DEX 2000 ATPS, A10: PEG 20000 / DEX 40 ATPS, A11: PEG 20000 / DEX 70 ATPS, A12: PEG 20000 / DEX 100 ATPS, A13: PEG 20000 / DEX 200 ATPS, A14: PEG 20000 / DEX 500 ATPS, A15: PEG 20000 / DEX 2000 ATPS and A16: original cell lysate (no ATPS). B1: original cell lysate (no ATPS); B2: PEG 2000 / DEX 500 ATPS; B3: PEG 2000 / DEX 2000 ATPS; B4: PEG 4000 / DEX 40 ATPS; B5: PEG 4000 / DEX 70 ATPS; B6: PEG 4000 / DEX 100 ATPS; B7: PEG 4000 / DEX 200 ATPS; B8: PEG 4000 / DEX 500 ATPS; B9: PEG 4000 / DEX 2000 ATPS; B10: PEG 20000 / DEX 40 ATPS; B11: PEG 20000 / DEX 70 ATPS; B12: PEG 20000 / DEX 100 ATPS; B13: PEG 20000 / DEX 200 ATPS; B14: PEG 20000 / DEX 500; ATPS and B15: PEG 20000 / DEX 2000 ATPS where in each case the upper half on the lane represents contents of the upper, PEG phase and the lower half of the lane represents the contents of the lower, DEX phase.

There was no genomic material present in the top PEG layers of any of the ATPS (Figure 3.4), indicating complete DNA partitioning to the lower DEX layer when either settling or centrifugation methodologies are employed.

## 3.5 Purified GFPuv Partitioning Analysis in Settled and Centrifuged ATPS

To determine whether other components of the lysate would influence the partitioning behaviour of the GFPuv protein, the experiments in section 3.2 were repeated using GFPuv that had been purified by his-tag purification.

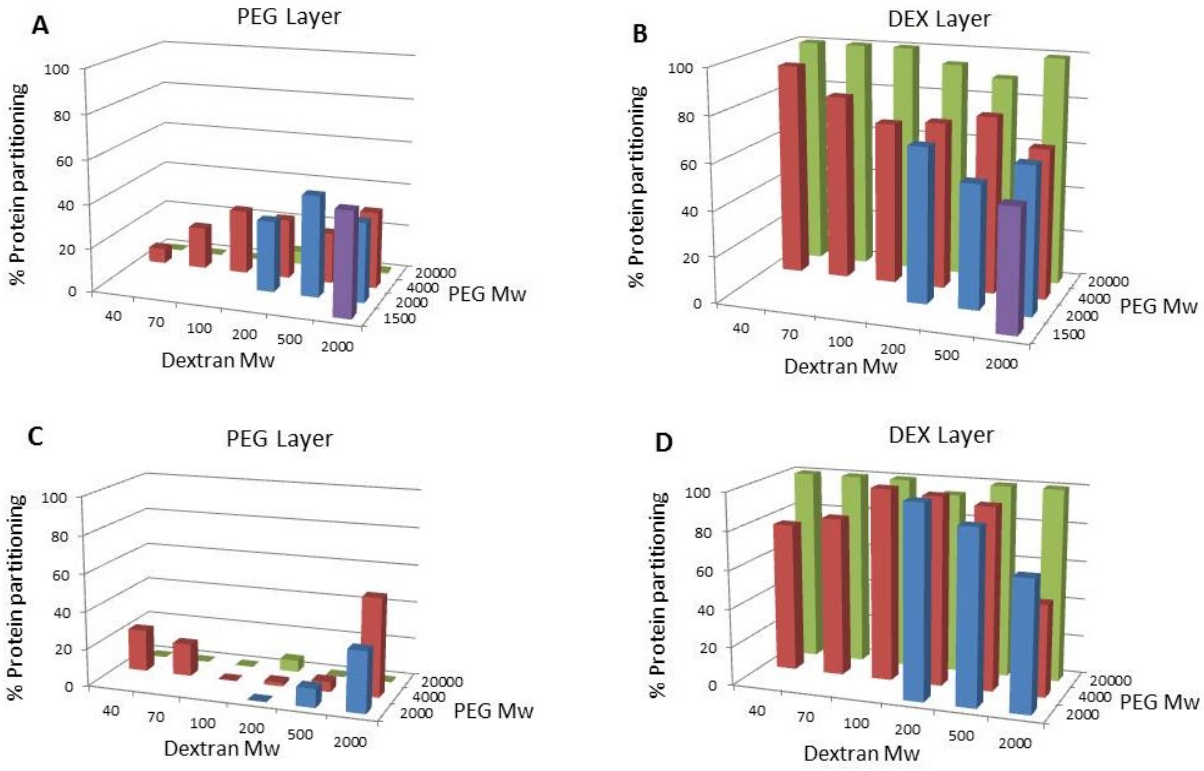



**Figure 3.5: Phase partitioning of 100µl purified GFPuv to upper PEG (A) and lower DEX (B) phases when utilising the settling methodology.** The phase partitioning to upper PEG (C) and lower DEX (D) phases of centrifuged systems are also shown (2000g for 2 minutes).

It can be seen above (Figure 3.5 A and B) that there is a clear trend in settled ATPS in which the majority of protein partitioned to the lower DEX phase when higher molecular weight polymers are employed. Whereas the lower molecular weight polymers, showed more even partitioning of purified GFPuv between the two layers. This trend is exemplified most strongly by the PEG 1500 Da and DEX 2000 kDa system in which partitioning is approaching a 50 / 50 distribution. In contrast, when using the centrifugation method to partition purified

GFPuv, there is no clear trend in partitioning behaviour therefore making the use of centrifuged systems less predictable. As there was no genomic material present in the purified GFPuv, DNA partitioning analysis was not performed.

## 3.6 Conclusion

The work in this chapter has laid the foundations for the proceeding chapters experiments by determining the experimental parameters to be employed. As the physical properties of the polymer solutions with differing molecular weights and concentrations were unknown, these initial experiments were essential to assess their suitability. Due to there not being a general consensus in current literature with regards to a generic ATPS composition, this work is the basis for a step towards a simple, generic system that could be prepared easily and quickly in both a research and industrial setting.

In this case, the results (Table 3.1) obtained from section 3.2 indicated that 25% (w/v) polymers had reduced accuracy due to high viscosity and 10% (w/v) polymers did not produce a wide enough variety of ATPS for analysis. Due to this, 20% (w/v) polymer solutions were determined to be suitable polymer concentrations and therefore chosen for further studies. Prior to the secondary study in this chapter starting, the effect of PEG or dextran on protein assays was investigated to ensure that the Bradford assay accuracy wasn't impeded. The results shown in figure 3.1 demonstrate that the polymers are diluted sufficiently to not interfere with protein detection during the Bradford assay, as supported by Barbosa *et al,* 2009. This finding was further supported in Table 3.2 where all polymers showed very low readings when analysed using Bradford reagent in the absence of protein. These results demonstrate that samples can be analysed for protein concentration directly from the ATPS, without the need for an additional purification step. These results demonstrate that this separation method can be easily analysed and therefore has potential for development into a commercial methodology.

The secondary study in this chapter systematically analysed the effect of cell lysate addition to the previously investigated 20% (w/v) combinations of polymers on the formation of two distinct phases. Upon the addition of cell lysates to ATPS formed using 20% (w/v) polymer solutions of varying molecular weights, it was determined that polymer combinations

of PEG and DEX, both of higher molecular weights, were more effective at partitioning of cellular proteins, GFPuv (Figure 3.3) and genomic DNA (Figure 3.4) to the DEX phase of the ATPS. These results (Table 3.3) also showed that allowing the systems to settle, instead of employing the centrifugation methodology, produced greater ATPS variants with complete lower phase partitioning. This was ideal for following experiments due to an affinity method being used to separate GFPuv into the upper PEG phase from the remaining cellular components in the lower DEX phase.

The behaviour of purified GFPuv in ATPS was also analysed to determine the effect of cellular components, within cell lysates, on GFPuv partitioning. The results shown in figure 3.5 demonstrate that his-tag purified GFPuv does not exhibit purely lower DEX phase partitioning but was also spread into the upper PEG phase. This indicates that other cellular components present in cell lysate do have a partitioning effect on GFPuv. This effect could potentially be overcome using PEGylated affinity ligands, developed in subsequent chapters, to ensure upper PEG phase partitioning of the desired protein. This lower DEX phase partitioning effect of cell lysate debris was desirable for the further development of affinity ATPS into a generic system as all detectable levels of undesired contaminants were partitioned away from the upper PEG phase.

# Chapter 4. Affinity ATPS using reduced glutathione and glutathione *S*-transferase fused proteins.

## 4.1 Introduction

The aim of this work was to use PEGylated, reduced glutathione (GSH) as an affinity ligand in ATPS. Studies of ATPS combinations featuring differing concentrations and molecular weights of PEG and DEX have shown that complete partitioning of all cellular components into the dextran layer preferentially occurs at both higher molecular weights and concentrations of the polymers (Chapter 3). It was also determined that the settling methodology was preferable for ATPS formation due to complete protein, cell debris and genomic material partitioning to the lower DEX phase (Chapter 3.4.2). Thus the hypothesis behind work in this chapter was that the PEGylation of GSH could help to partition glutathione *S*-transferase (GST) fused proteins into the upper PEG layer of the ATPS. (Figure 4.1).

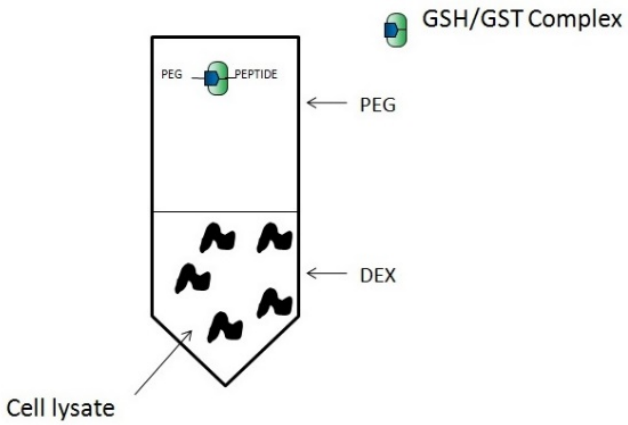



**Figure 4.1: Schematic of an ATPS system comprising a lower DEX phase and upper PEG phase.** Cell lysate containing the GST-GFPuv fusion protein was added to the system along with PEGylated GSH.

The PEGylated ligand should remove the fusion protein of interest into the upper phase, and therefore away from any cellular debris and native proteins. To facilitate visualisation, green fluorescent protein (GFPuv) was expressed as a glutathione *S*-transferase fusion (GST-GFPuv) for use in these studies. The expression of GST-GFPuv (Chapter 4.2), the synthesis of PEGylated reduced glutathione (PEG-GSH) (Chapter 4.4) and their evaluation in ATPS partitioning (Chapter 4.6) are described.

## 4.2 Expression of GST-GFPuv fusion protein

The GFPuv gene was amplified from the pGFPuv vector (Clontech) using the primers presented in methods 2.4.1 and 2.4.2. Due to presence of a *BamHI* site within the required GFPuv gene (shown in appendix 4.1) it was not possible to simply perform a straightforward restriction digest and ligation. To circumnavigate this (cloning schematic demonstrated in Figure 4.2), the GFPuv gene was amplified from the pGFPuv plasmid with modified primers to insert a *SmaI* restriction enzyme site at the 5' end of the gene and a 3' *EcoRI* site. *SmaI* was employed first to cut the GFPuv insert at the 5' end, followed by *EcoRI* to digest the 3' end of the gene (Methods 2.9.1). The plasmid pGEX2TK was prepared by first digesting with *BamHI* (Methods 2.9.2) resulting in a linearised vector with 5' overhangs. Mung bean nuclease was used to remove these 5' overhangs, followed by *EcoRI* digestion (Methods 2.9.3). The resulting pGFPuv insert was ligated into the pGEX2TK vector using T4 DNA ligase (Methods 2.9.4). The resulting construct was sent for sequencing to the Birmingham University Functional Genomics Laboratory which confirmed the sequence structure (Appendix 4.1 plasmid map + sequence of fusion).

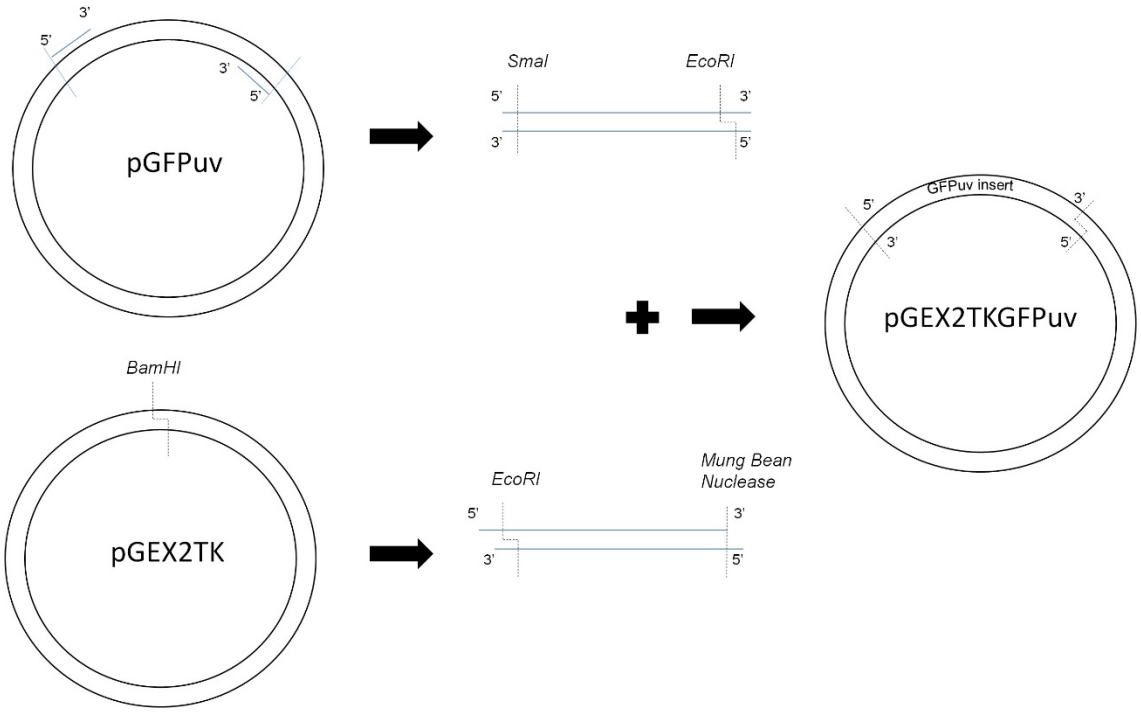



**Figure 4.2: Schematic demonstrating the cloning of the GFPuv gene into the pGEX2TK vector.**

*E. coli* Tuner (DE3) bacterial cells were prepared and transformed with the resulting vector containing the GST-GFPuv fusion gene (Methods 2.8.1). Expression of the GST-GFPuv gene was performed in *E. coli* Tuner DE3 cells induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). Preliminary expression at 37°C gave poor yields of soluble protein, but good yields were obtained successfully at a 30°C growth temperature (Methods 2.8.2). A fraction of the resulting culture (50 ml) was harvested by centrifugation and lysed using the commercial reagent Bugbuster™, with the remainder of the culture (150 ml) frozen at -20°C for future use. The desired GST-GFPuv fusion protein was purified by affinity chromatography (Methods 2.8.3) followed by analysis using polyacrylamide agarose gel electrophoresis (PAGE; Methods 2.6.2; Figure 4.3).

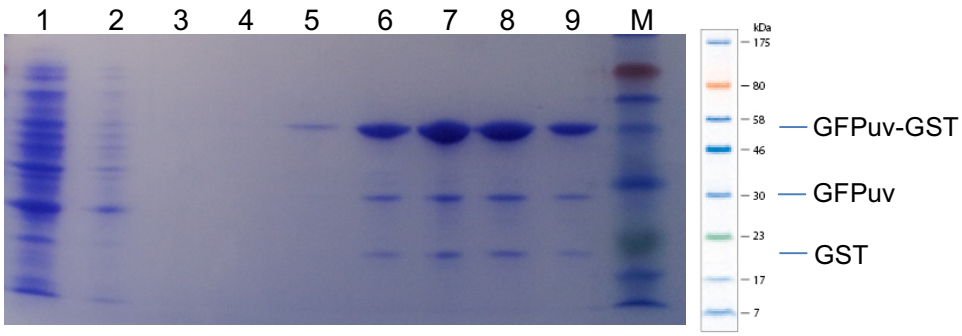



**Figure 4.3: Polyacrylamide gel electrophoresis of GST-GFPuv purification.** 12% PAGE gel, stained with Coomassie Blue. Lanes: 1) Column flow through; 2-4) Washes 1-3; 5-9) Elutions 1-5; M) 15 µl NEB ColorPlus™ Prestained Protein Marker, Broad Range (7-175 kDa). All other samples were 5 µl of original material.

Elutions 2-5 (250 µl each; Figure 4.3) were pooled to generate the purified GST-GFPuv fusion protein for use in future experiments and quantified by the Bradford assay (Method 2.8.14), yielding a total volume of 1 ml GST-GFPuv fusion protein at a final concentration of 4.09 mg/ml.

## 4.3 Partitioning of purified GST-GFPuv

In Chapter 3, it was established that many combinations of higher molecular weight PEG and DEX ATPS, when using the settling methodology, partitioned both total *E. coli* nucleic acid and proteins into the lower DEX layer. To ensure that this finding was consistent with the behaviour of purified GST-GFPuv, the partitioning of this fusion protein was examined with a range of ATPS conditions.

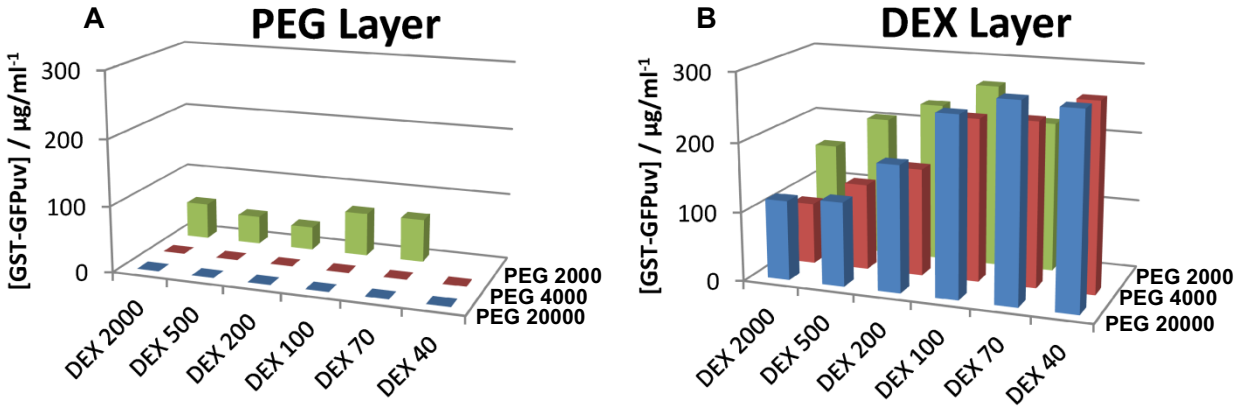
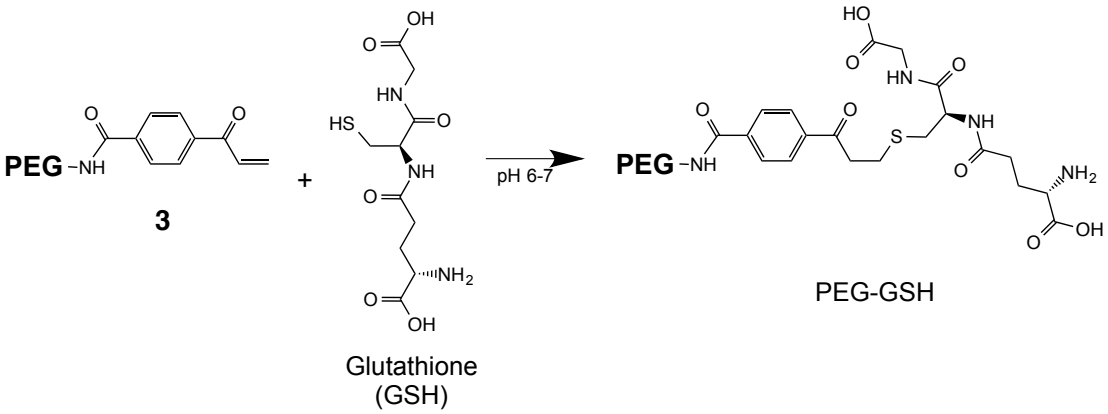



**Figure 4.4: Histograms demonstrating the distribution of purified GST-GFPuv fusion across PEG and DEX layers of ATPS of various compositions.** Solutions of polymers (0.5 ml, 20% (w/v) PEG 20000 Da, 4000 Da, 2000 Da and DEX 2000 kDa, 500 kDa, 200 kDa, 100 kDa, 70 kDa and 40 kDa) were combined as indicated to form ATPS with 100 µl purified GST-GFPuv fusion protein (4.09 mg/ml). Protein concentration was determined by the Bradford assay. Note that PEG 2000 Da and DEX 40 kDa failed to form two distinct phases.

As Figure 4.4 shows, GST-GFPuv partitioned predominantly into the lower DEX phase when the higher molecular weight PEGs (20000 Da and 4000 Da) were used. In contrast, when PEG 2000 Da was used, protein had partitioned into the upper PEG phase indicating that higher molecular weight PEGs are desirable. This is supported in the literature with Johansson *et al.,* (2008) describing a decrease in the partitioning coefficient (K) of GFPuv with an increase of PEG molecular weight. For the purposes of this study, a low K for GFPuv and other cellular components is desirable as the PEGylated affinity ligand should overcome this, allowing for upper PEG phase partitioning of only the required GST-GFPuv fusion protein.

## 4.4 Synthesis of PEGylated GSH Ligands

PEGylation of GSH was performed using the M1a PEGylation reagent (Methods 2.8.5; Scheme 4.1, 1; provided by Polytherics Ltd.) which links one PEG molecule to the thiol of a single GSH molecule (2). The conjugates were made in both 5 kDa and 10 kDa forms of PEG and then the conjugate was purified (Methods 2.8.5). The reaction for the conjugation of M1a PEG-GSH (3) is illustrated in Scheme 4.1.




**Scheme 4.1: Conjugation of PEG M1a to reduced glutathione.** i – PEGylation reaction buffer (2.2.11); 18 hours; room temperature.

The resulting conjugates were examined by $^{1}$H NMR spectroscopy (Figure 4.5 A and B; interpretation in Methods 2.8.5) and then stored at -20°C under argon, until required. $^{1}$H NMR spectroscopy was selected for this analysis as it was possible to confirm the structure of the analyte ensuring its suitability for further experimentation.

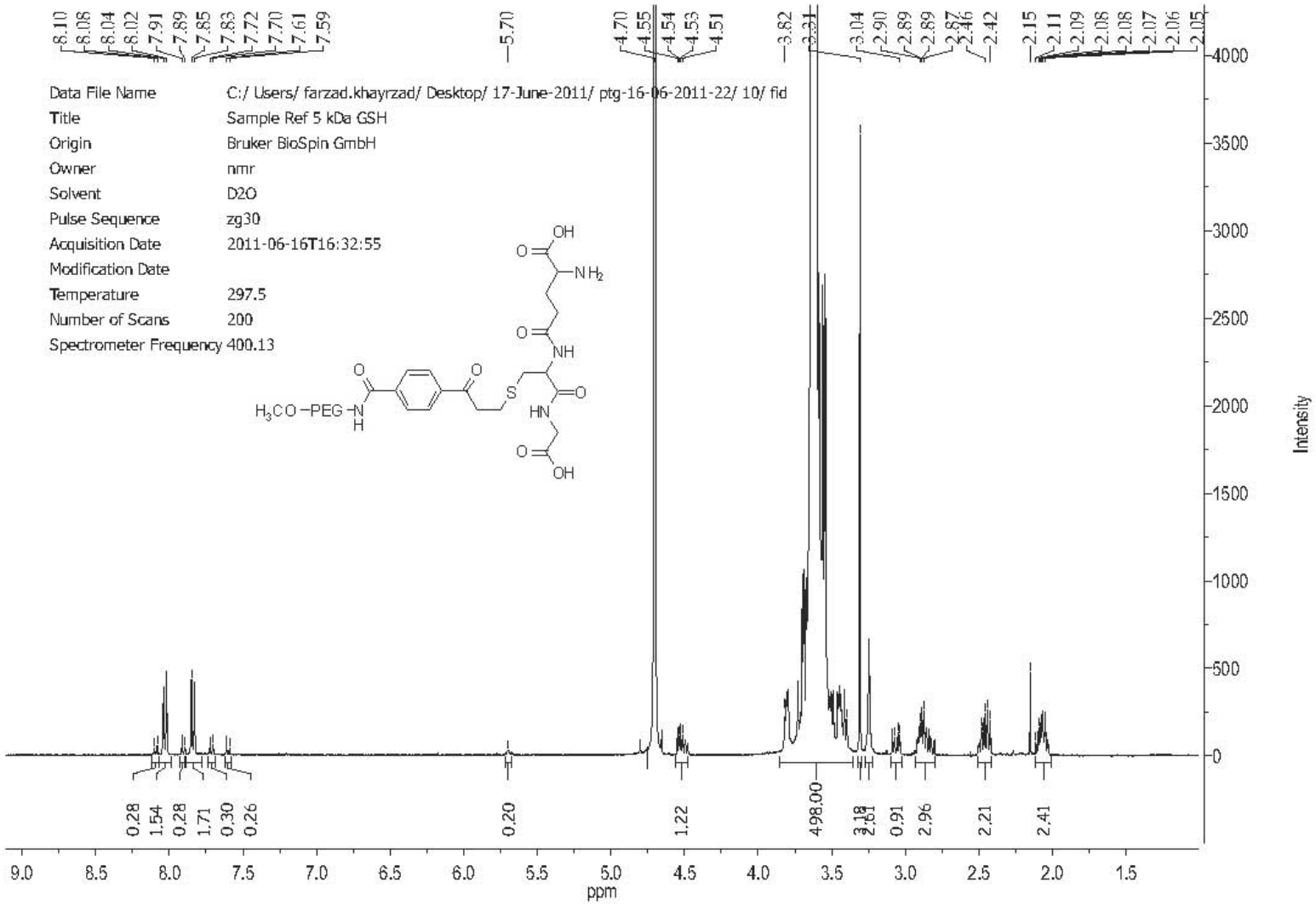



**Figure 4.5: $^{1}$H NMR Spectrum of 5 kDa PEGylated GSH**

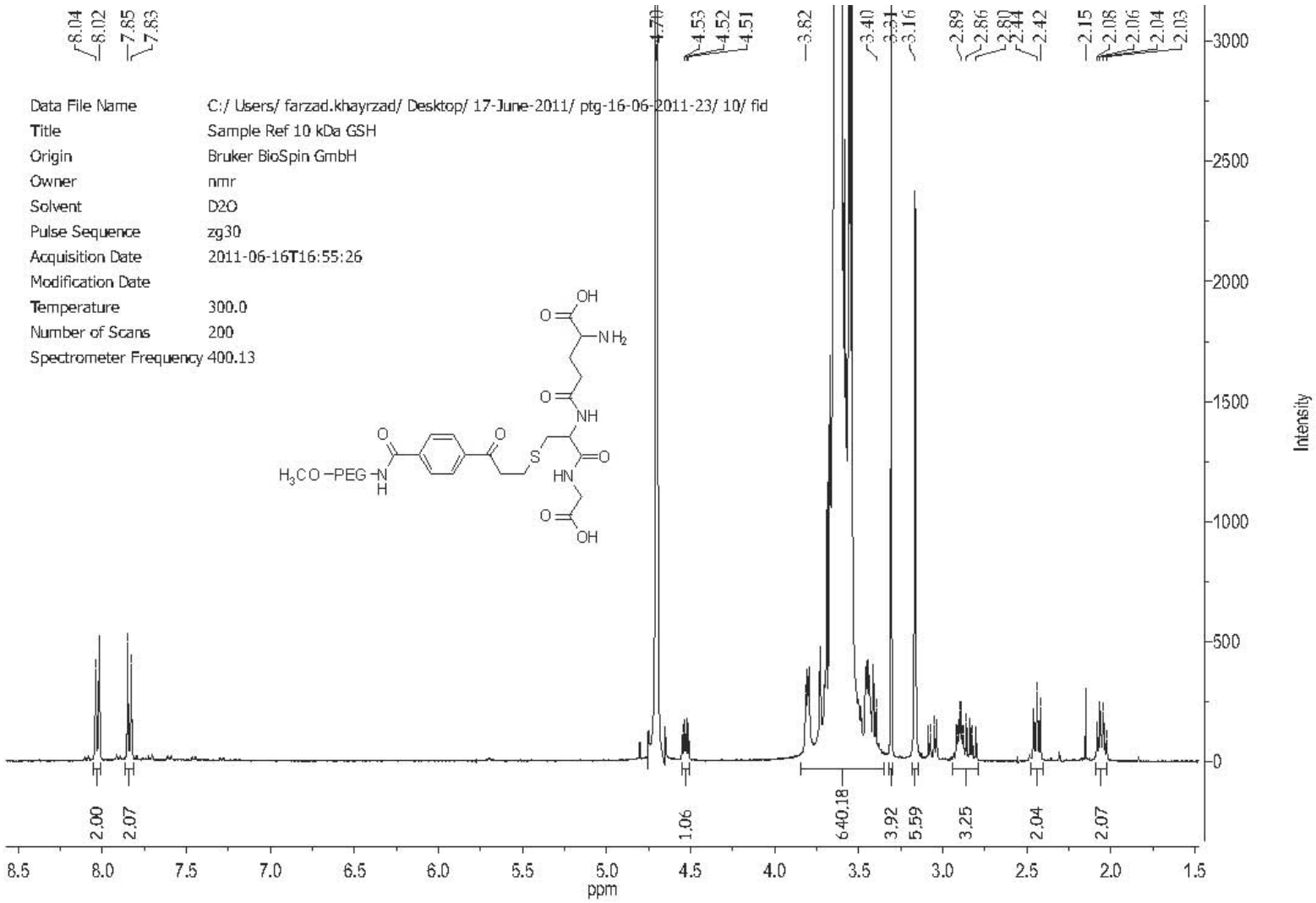



**Figure 4.6: $^{1}$H NMR Spectrum of 10 kDa PEGylated GSH**

## 4.5 Investigation of interactions between PEGylated GSH ligands and GST fused protein

As commercially reduced GSH can be immobilised onto solid support using the thiol group, via a linker (Figure 4.7), it is suggested that the site specific PEGylation of the thiol within GSH should not interfere with its subsequent binding with GST. It is therefore a reasonable assumption, that a free thiol on GSH is not required for binding to GST.


**Figure 4.7: Chemical structure demonstrating the binding of glutathione to an agarose solid support via a 12 atom spacer in the commercially available Thermo Scientific Pierce Glutathione Agarose.** (Image reproduced from http://www.piercenet.com/product/pierce-glutathione-agarose-kits)

However, the steric shielding caused by the 5- or 10- kDa PEG (relative to GSH, which has a MW of 307 Da) might also prevent an effective interaction between PEGylated GSH and GST. To examine this possibility, the ability of both 5- and 10- kDa PEGylated GSH to elute GST-GFPuv fusion protein from glutathione sepharose was investigated. A fresh preparation of cleared, concentrated *E. coli* lysate containing the GST-GFPuv fusion protein was prepared and 0.5 ml concentrated cell lysate (2.8.2) was applied separately to two 0.25 ml bed volume glutathione sepharose columns (GE Healthcare). Initially, to confirm appropriate test conditions, these columns were then eluted with either 10 mM glutathione in 50 mM Tris-HCl pH 8.0 (as per manufacturer's instructions), or by 50 mM Tris-HCl pH 8.0 alone. Interestingly, both conditions caused the fusion protein to elute from the column, with elution being complete

by the end of the third application of 0.25 ml elution buffer. A further control experiment was then performed in which elution with water was attempted. Here, elution was limited to minimal leaching, whilst the vast majority of the fusion protein remained bound to the column. Water was tested for elution since it was planned that subsequent ATPS experiments would be performed in water.

As previously, 0.5 ml cleared, concentrated *E. coli* lysate containing GST-GFPuv fusion protein was applied to two 0.25 ml bed volume glutathione sepharose columns. Each column was eluted with either 4 × 0.25 ml 1 mM 5 kDa PEG-GSH or 4 × 0.25 ml 1 mM 10 kDa PEG-GSH, both in water. Eluted fractions were then examined for GFPUV fluorescence. As previously, elution of the fusion protein was complete by the end of the third application of 0.25 ml elution buffer. Since water alone failed to elute the fusion protein, it was concluded that both 5 kDa PEG-GSH and 10 kDa PEG-GSH were capable of binding to the GST element of the GST-GFPuv fusion protein. Both PEG-GSH conjugations were therefore deemed suitable for further experimentation within ATPS to determine their efficacy at partitioning the GST-GFPuv fusion protein into the upper PEG phase.

## 4.6 Cell Lysate Partitioning Analysis

Use of purified GST-GFPuv had established that this protein is detectable by the Bradford assay within ATPS polymer solutions. Whilst fluorescence is useful for rapid visual observation, the Bradford assay is also required, since contaminating cellular proteins clearly cannot be detected by GFPuv fluorescence. However, purified GST-GFPuv that is already complexed with GSH as a result of its purification would not, therefore, be expected to interact strongly with added PEGylated GSH. Moreover, the point of the affinity ATPS system is to purify proteins directly from crude lysates. Therefore, in order to establish whether the PEGylated GSH conjugates could partition the GST-GFPuv fusion protein into the upper PEG phase of an ATPS, clarified cell lysates containing the target protein were used for evaluation of partitioning in ATPS.

Specifically, a 1 ml ATPS comprising 0.5 ml each of 20% (w/v) 20000 Da PEG and 20% (w/v) 100 kDa DEX was chosen for initial investigation, since this combination is one of the most consistent for lower DEX phase partitioning in the initial screening of ATPS combinations (Chapter 3), as well as having a lower viscosity than its higher molecular weight counterparts. To this was added clarified concentrated *E. coli* lysate (100 µl) and a 100 fold excess of 10 kDa PEG-GSH (section 4.2). The ATPS was then mixed by vortexing and then left to settle for 20 minutes and examined using UV light. Unfortunately, all detectable GST-GFPUV fusion protein had partitioned to the lower DEX layer, rather than being “pulled” to the upper PEG layer by the PEGylated GSH ligand, as had been hypothesised. Further experiments involving a greater, 200 fold excess of 10 kDa PEG-GSH produced a similar result, as did repeat experiments with 5 kDa PEG-GSH. Clearly, this experimental set-up would not form an effective ATPS for affinity partitioning of GST fusion proteins (data not shown).

It was hypothesised that perhaps the interactions between the high molecular weight DEX and the fusion protein was too great to be overcome via a PEGylated ligand and that therefore lower molecular weight polymer should be re-investigated. To this end, it was decided to test the ATPS conditions of a proven protocol described by Barbosa et al (2010), which had used PEGylated IDA with a coordinated cation to effect affinity purification of plasmid DNA, via a LacI-$His_6$-GFPuv fusion protein. Specifically, their ATPS which produced the best partitioning of plasmid and protein to the upper PEG phase (which comprised 16% overall PEG, of which 15.85% was PEG 600 and 0.15% PEG-IDA, 14.3% DEX 40) was adapted for use in the current study. In this experiment, 50 µl clarified cell lysate (diluted in 50 mM Tris pH 7.4 to 1 mg / ml total protein content) was added to PEG 600 (15.85% w/v), PEG-GSH (0.15% w/v) and DEX 40 (14% w/v) and ultrapure water was added to a final system weight of 1.5 g. After mixing by vortexing and settling for 2 hrs at 30°C, fluorescence was examined by UV light. Unfortunately, the fluorescence of the GST-GFPuv fusion protein was again found in the lower DEX phase, albeit just underneath the interface of the two phases in both 10 kDa and 5 kDa PEG-GSH systems (data not shown).

To determine whether the simple addition of a higher concentration of PEG-GSH would result in upper PEG phase partitioning of the GST-GFPuv fusion protein, a series of ATPS were conducted in similar conditions to those described above. Increased concentrations of the PEG-GSH ligand were used whilst maintaining an overall PEG concentration of 16% (Table 4.1).

| DEX 40 (%) | PEG 600 (%) | PEG 10000-GSH | PEG 5000-GSH |
|---|---|---|---|
| 14.0 | 15.5 | 0.5 | |
| 14.0 | 15.0 | 1.0 | |
| 14.0 | 14.5 | 1.5 | |
| 14.0 | 14.0 | | 2.0 |
| 14.0 | 12.5 | | 3.5 |

**Table 4.1: ATPS formed using a variety of PEG and PEG-GSH concentrations.** Partitioning of GST-GFPuv is demonstrated by green shading.

Analysis showed that all ATPS had high levels of fluorescence in the lower DEX layer and at the interface, but not within the PEG layer.

## 4.7 Discussion

The GST-GSH fusion protein purification method has been established in the research community for many years, having first been described by Smith, D.B. and Johnson, K.S. in 1988. It has been a favoured protein purification method due to the high level of protein expression as well as the N-terminal GST moiety acting as a chaperone and therefore assisting the expression of the target proteins as soluble protein instead of inclusion bodies. In addition to these benefits, the GST fusion protein can be efficiently eluted from the GSH-agarose using a mild, detergent free wash buffer containing reduced glutathione which does not denature the desired fusion. In addition to the ease of purification, there is a site-specific protease recognition sequence located next to the protein of interest allowing for complete cleavage resulting in a target protein with no additional tags (Harper, S. & Speicher, D.W., 2011). This is particularly advantageous as tag free proteins are required for biomedical research and therapeutic applications. These benefits put GST-GSH affinity purification at a distinct advantage over other purification methodologies and therefore it was chosen as the initial fusion system to be studied in this project.

SDS-PAGE was performed on the purification solution to confirm the presence of the GST-GFPuv fusion protein required for affinity purification within the two phase system. Figure 4.3 shows a high concentration of fusion protein production, approximately in line with the 56 kDa marker, allowing for a final concentration of 4.09 mg/ml. A small amount of protein degradation can be seen with additional bands present which are the two constituent proteins GFPuv and GST separately. In this methodology protease inhibitors were not utilised due to their toxic properties. Their use to prevent the small amount of protein degradation present could render the affinity ATPS method unsuitable for commercial therapeutic protein purification. The fusion protein obtained from this purification was used in various combinations of different molecular weight PEG and dextran ATPS and figure 4.4 demonstrates the distribution of purified GST-GFPuv fusion protein within the systems.

Partitioning of the fusion protein into the upper PEG phase when using PEG 2000 Da indicated that higher molecular weight PEGs were needed to ensure that non-specific partitioning of proteins was not occurring. Functionalised PEG could then be used to overcome the interfacial tension and only partition the targeted fusion protein into the upper PEG phase. In this instance, PEG would be functionalised with reduced glutathione (Scheme 4.1) allowing for specific affinity purification of the protein fused to GST. Two functionalised PEG species (5 kDa PEG-GSH and 10 kDa PEG-GSH) were produced and verified using $^{1}H$ NMR Spectroscopy (Figures 4.5 and 4.6 respectively).

The percentages of functionalised PEG described in Table 4.1 represent an approximate 150 to 1000 fold molar excess of PEGylated GSH over the estimated concentration of GST-GFPuv fusion protein. Although this high molar excess is used in column protein purification methods to elute the required fusion protein from the column, the competing GSH molecules are PEGylated so were located within the upper PEG phase. This should allow for all GST-GFPuv fusion proteins to partition across the interface into the upper PEG phase. In comparison to PEG/Salt and other solvent based ATPS systems, the interfacial tension present in PEG / dextran systems are very low which has traditionally been thought to allow for gentle partitioning of target proteins without disruption or damage occurring to the protein structure. (Ryden, J *et al.,* 1971) When forming an ATPS, the concentration of the constituent polymers determines interfacial tension, for example – the higher the concentration of polymer, the higher the interfacial tension (Mittal, KL, 1981). As can be seen in Table 4.1, a number of different concentrations of PEG 600 were used in order to assess whether a reduction in interfacial tension had an effect on the partitioning of the GST-GFPuv protein. In this instance there was no effect and thus unfortunately, it is clear that PEGylated glutathione (PEG-GSH) is unlikely to form the basis of an effective affinity ATPS for purification of fusion proteins. It is hypothesised that the interaction between GST and GSH is too weak to permit effective partitioning across the interfacial tension present between the PEG and dextran layers and as such, that a protein ligand interaction with a lower $K_d$ will be required for use in an effective affinity ATPS for fusion protein purification (Zachariou, M. & Bailon, P., 2007). In

addition to this, the two proteins fused together, GST and GFPuv, are both large and carry an overall polar charge which may inhibit movement between the layers (Asenjo, JA *et al.,* 2011).

It could also be hypothesised that the act of sonication when releasing the soluble fusion protein from the bacterial cells may have degraded resulting in cleavage of the GST protein from the GFPuv protein. This would automatically prevent a proportion of GFPuv protein from being partitioned to the upper PEG phase. Another inhibiting factor may be steric hindrance caused by the sheer size of the 5 kDa and 10 kDa PEG molecules attached to the 307 Da reduced glutathione molecule. If the PEG molecule is causing GSH to become unable to bind into the active site of GST, partitioning to the upper PEG phase via the affinity method will not occur. To this effect, further investigation will be undertaken into an alternative affinity ATPS model using PEGylated Nitrilotriacetic Acid (NTA) with conjugated copper ions alongside the addition of a small histidine tag to the target protein. This method should allow for the investigation of effect of binding affinities between the ligand and substrate as well as the size of the protein on partitioning behaviours.

# Chapter 5. Affinity ATPS using PEGylated Copper-Nitrilotriacetic acid (PEG-NTA-$Cu^{2+}$) and Poly-Histidine ($His_6$) Tagged Proteins.

## 5.1 Introduction

Studies using a GST and GSH affinity receptor and ligand complex to pull a protein from the DEX to the PEG phase had proved unsuccessful (Chapter 4). The relative weakness of the interaction between GSH and its affinity partner, GST might explain those results, since the $K_d$ for the GST and GSH interaction is in the μM range. In contrast, that of a $His_6$ tag / divalent cation interaction is in the nM range (Knecht *et al.,* 2009). Moreover, PEGylation of the chelating ligand iminodiacetic acid (IDA) had previously proven successful in purifying a plasmid within an affinity ATPS system (Barbosa *et al.,* 2010). It was therefore decided to attempt ATPS purification of protein via a $His_6$-tag and PEGylated chelating agent. Because commercial protein purification resins use nitrilotriacetic acid (NTA), it was decided to PEGylate NTA and to evaluate PEG-NTA-$Cu^{2+}$ for affinity ATPS of His-tagged protein. The synthesis of PEGylated NTA-$Cu^{2+}$ and its evaluation in ATPS's of both purified protein and clarified *E. coli* lysate are described.

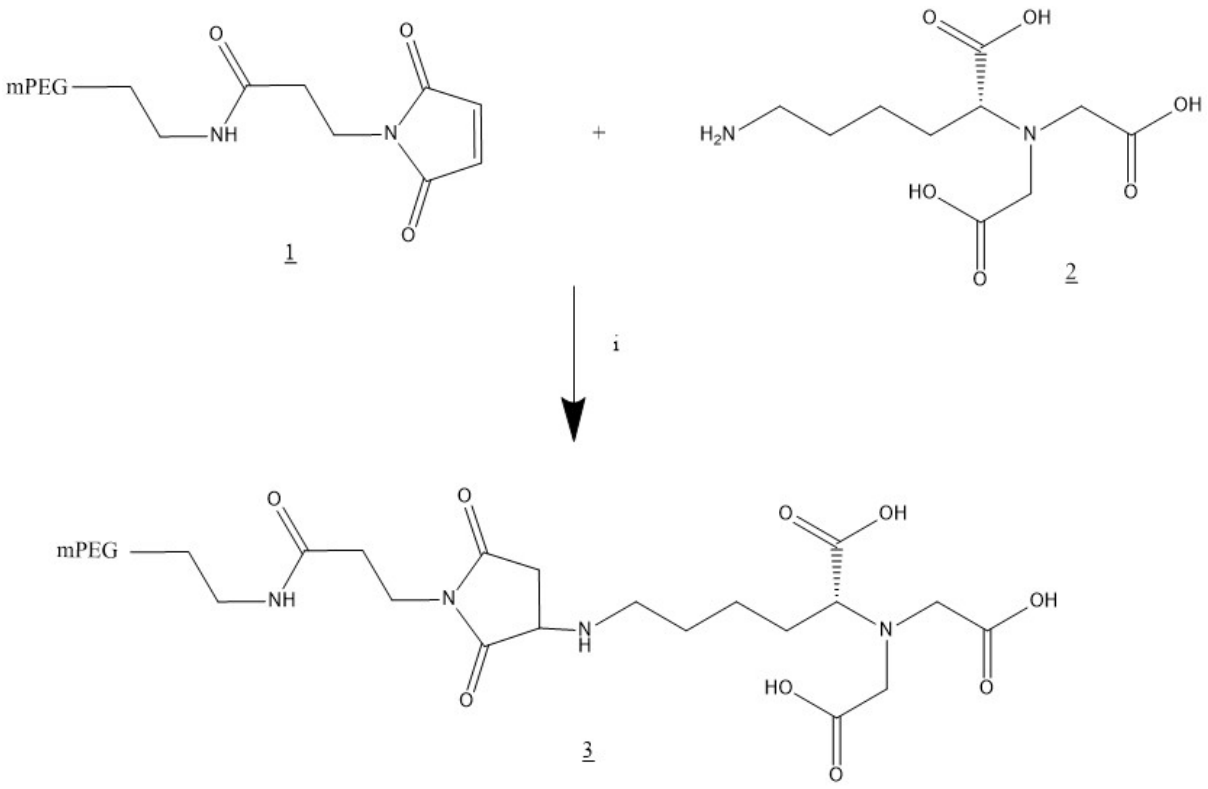
## 5.2 Synthesis of PEGylated NTA-$Cu^{2+}$ ligands

Three alternative classes of PEGylated ligands were conjugated with lysine NTA (2). One (methoxy PEG maleimide) (1) was obtained directly from JenKem Technology USA Ltd. and closely mimics the proprietary M1 PEGylation reagent provided by Polytherics for PEGylation of glutathione (Chapter 4) and will similarly link PEG to one

NTA molecule. Second and third reagents (B1 and *bis*-B1) were prepared according to the methods of Brocchini *et al.,* (2006) and were used to link PEG to two NTA molecules and four NTA molecules respectively. Each reagent was used in both 5 kDa and 10 kDa forms of PEG.

### 5.2.1 Preparation of methoxy PEG maleimides

Both 10 kDa and 5 kDa methoxy PEG maleimide (1) were conjugated to lysine-NTA (2) and the products purified as described (Methods 2.8.6; Scheme 5.1). The starting materials (1 and 2) were examined by $^{1}$H NMR spectroscopy (Figures 5.1 and 5.2).




**Scheme 5.1: 1 Coupling of Methoxy PEG Maleimide to Lysine NTA.** i – 50 mM Sodium Phosphate pH 6.0, 24 hrs, Room Temperature.

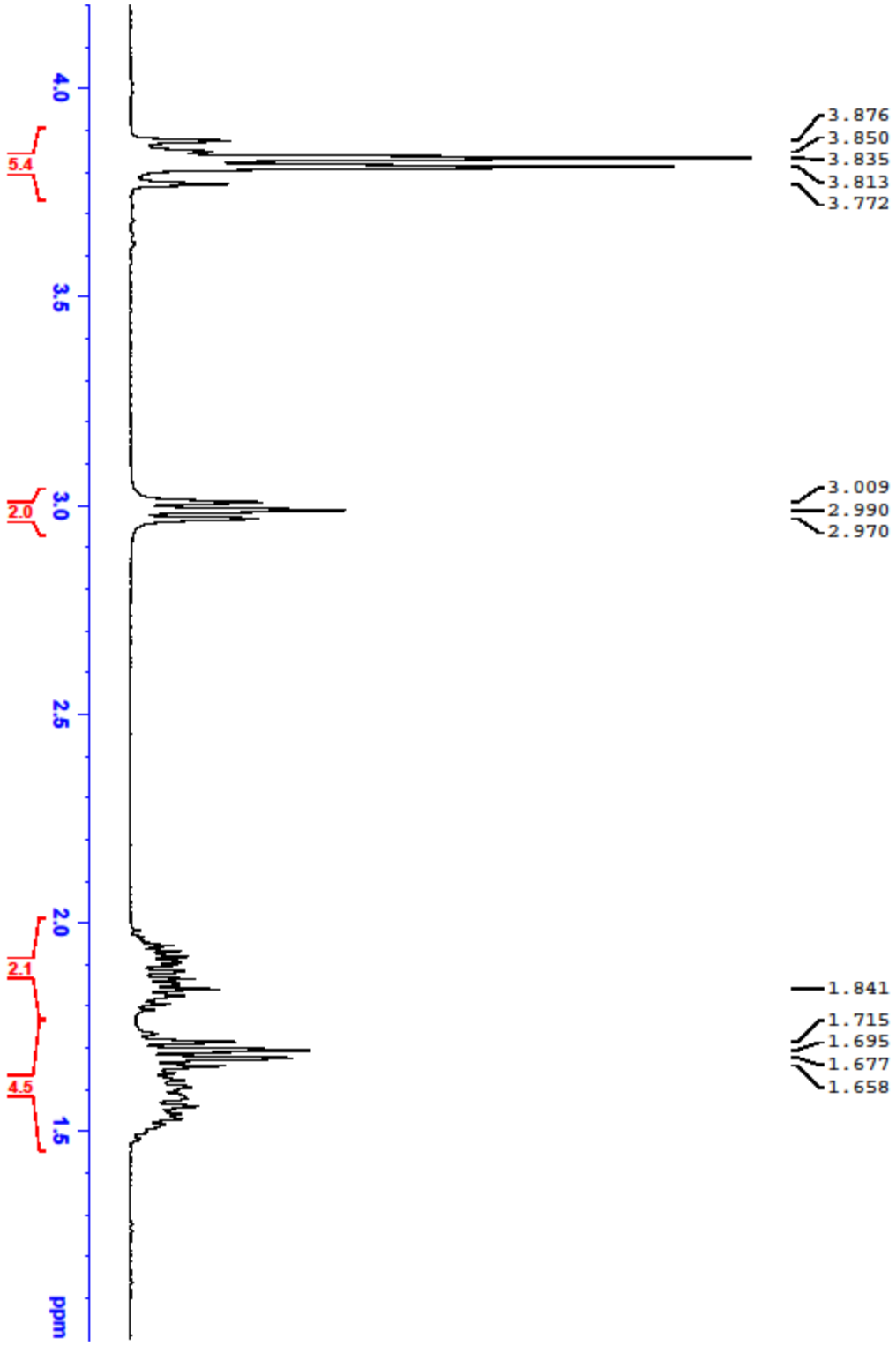



**Figure 5.1: $^{1}$H NMR of Lysine NTA (2)**

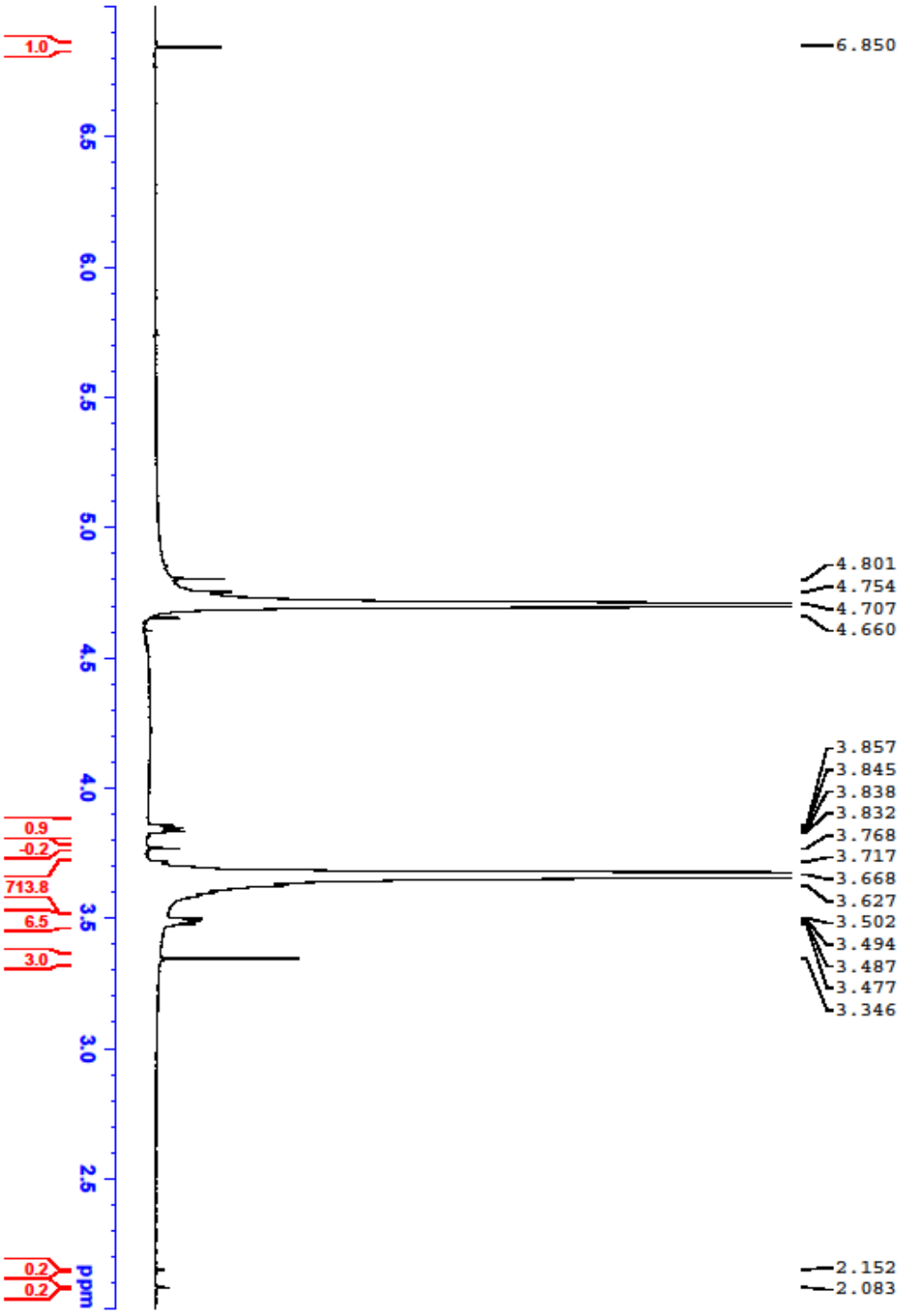
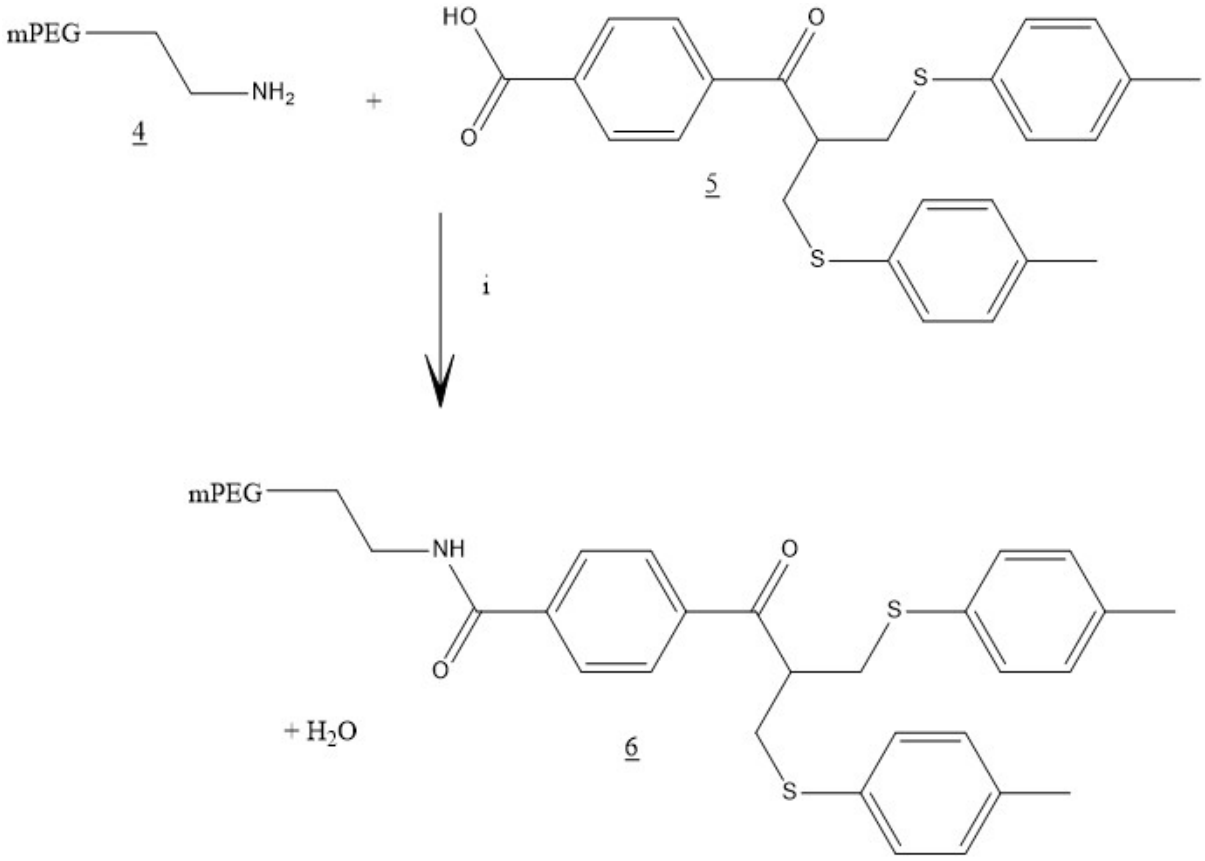



**Figure 5.2: $^{1}$H NMR of Methoxy PEG Maleimide 10 kDa (1)**

### 5.2.2 Preparation of B1-NTA

The carboxylic DB-*bis*-sulfide linker (B1, obtained from Prof. Steve Brocchini, UCL) (5) was conjugated to both 10 kDa and 5 kDa methoxy PEG amine (4) (JenKem Technology USA Ltd.) and purified as described (Methods 2.8.7). The reaction scheme for the conjugation of mPEG B1 (6) is illustrated in Scheme 5.2.




**Scheme 5.2: Coupling of methoxy PEG amine (4) to B1 linker (5). i – Anhydrous dichloromethane under argon, room temperature, 48 hrs.**

The resulting conjugates were examined by $^{1}$H NMR spectroscopy (Figure 5.3) and then stored at -20°C under argon, until required, since the sulfide is more stable than the corresponding sulfone and is thus more suitable for long term storage.

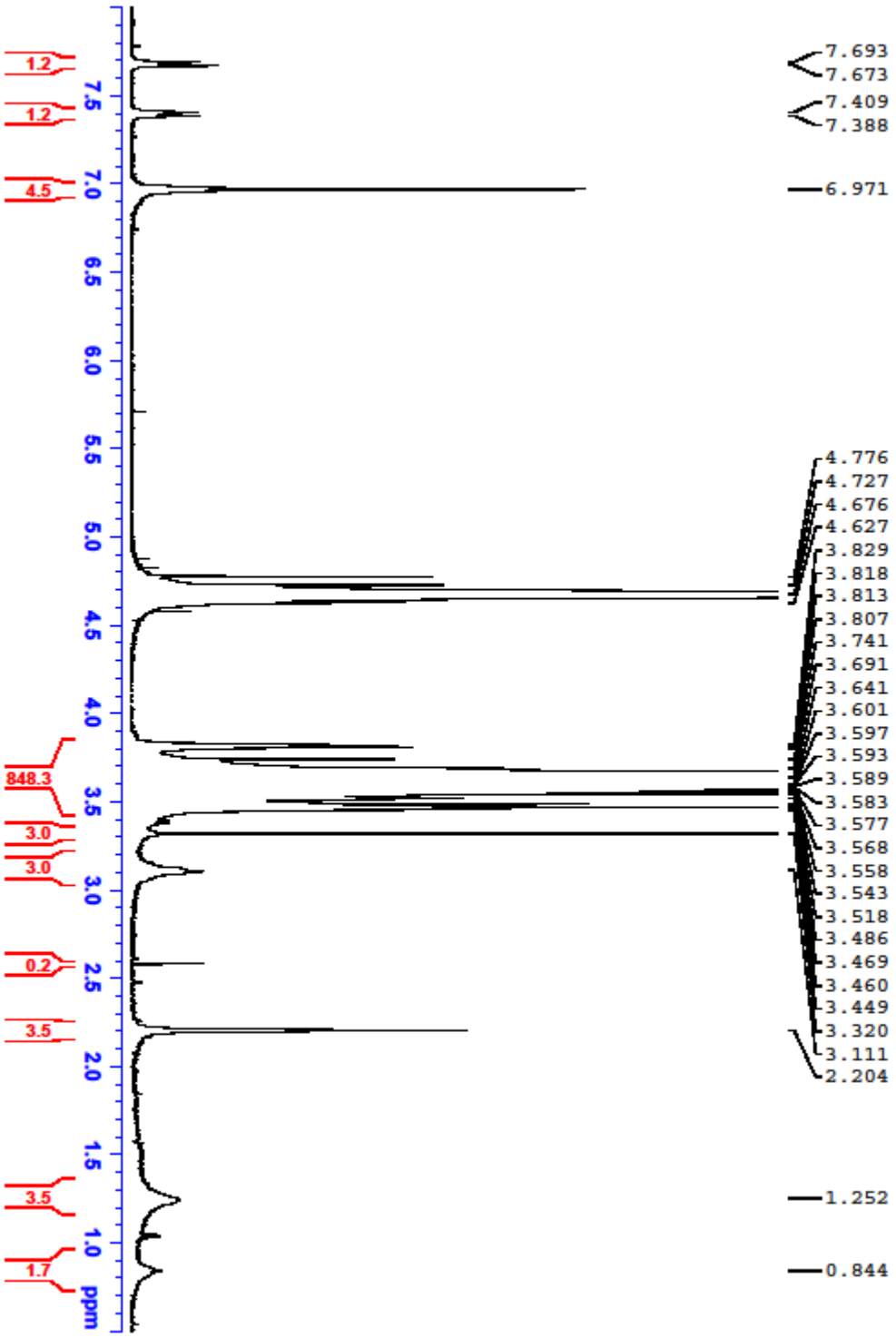
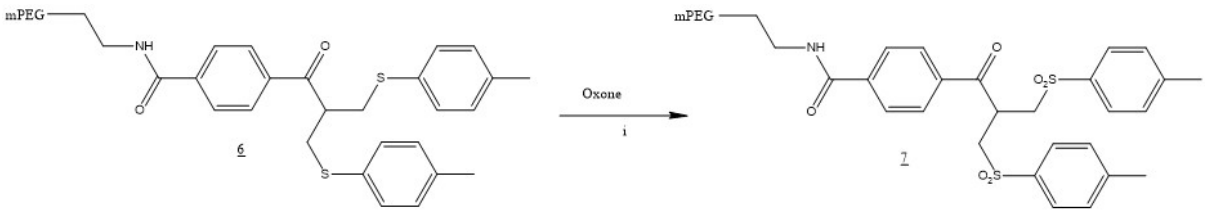



**Figure 5.3: $^{1}$H NMR Spectrum of 10 kDa PEG B1 Sulfide**

Prior to use, the required quantity of sulfide conjugate was oxidised to the corresponding sulfone with Oxone® (Sigma, 228036) (Methods 2.8.8). A representative NMR of a sulfone product is shown in Figure 5.4. The reaction scheme for the production of oxidised B1 (7) is illustrated in Scheme 5.3.




**Scheme 5.3: Oxidation of PEG B1 Sulfide to Sulfone using OXONE®.** i – Methanol, overnight, room temperature.

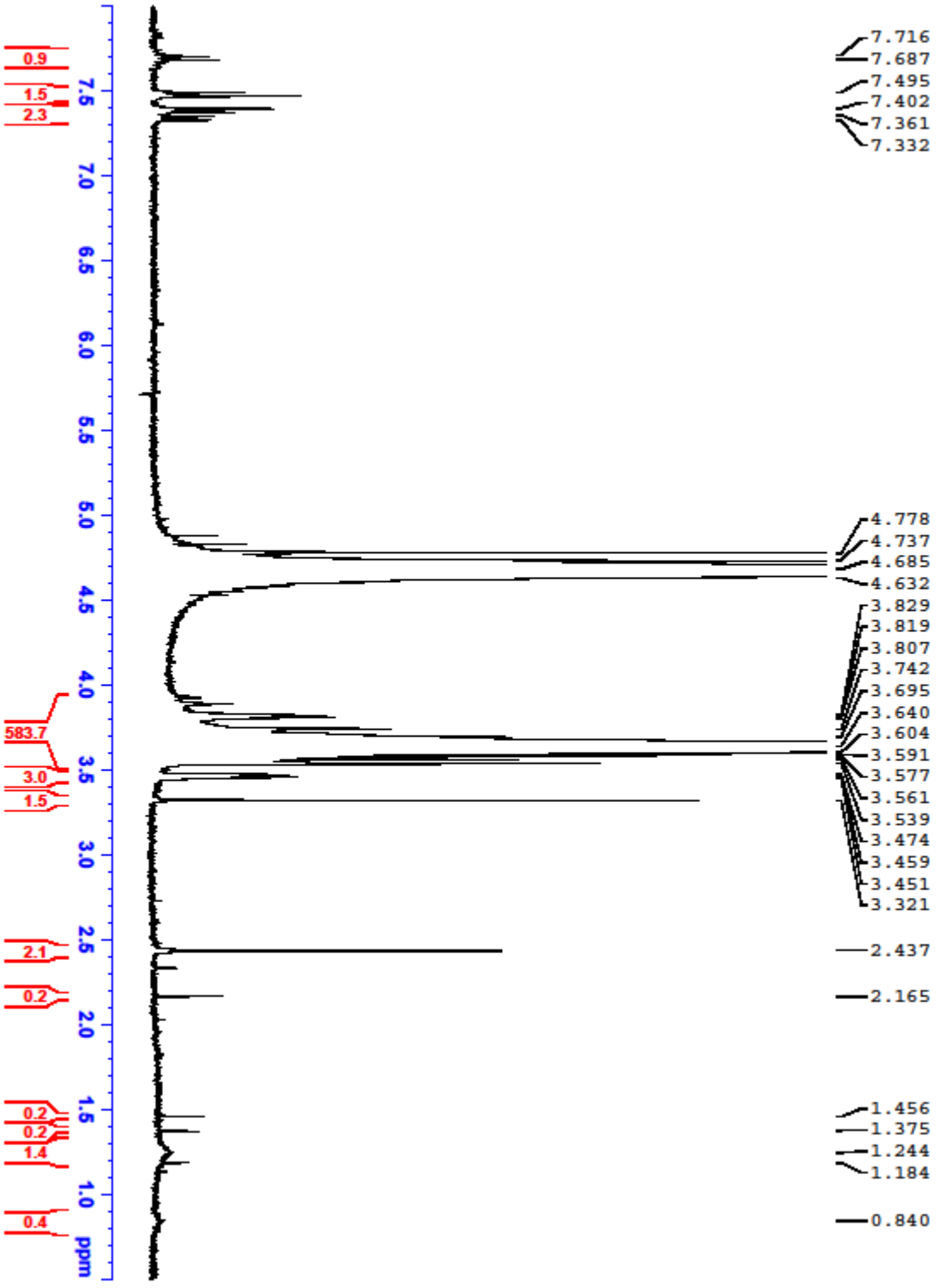
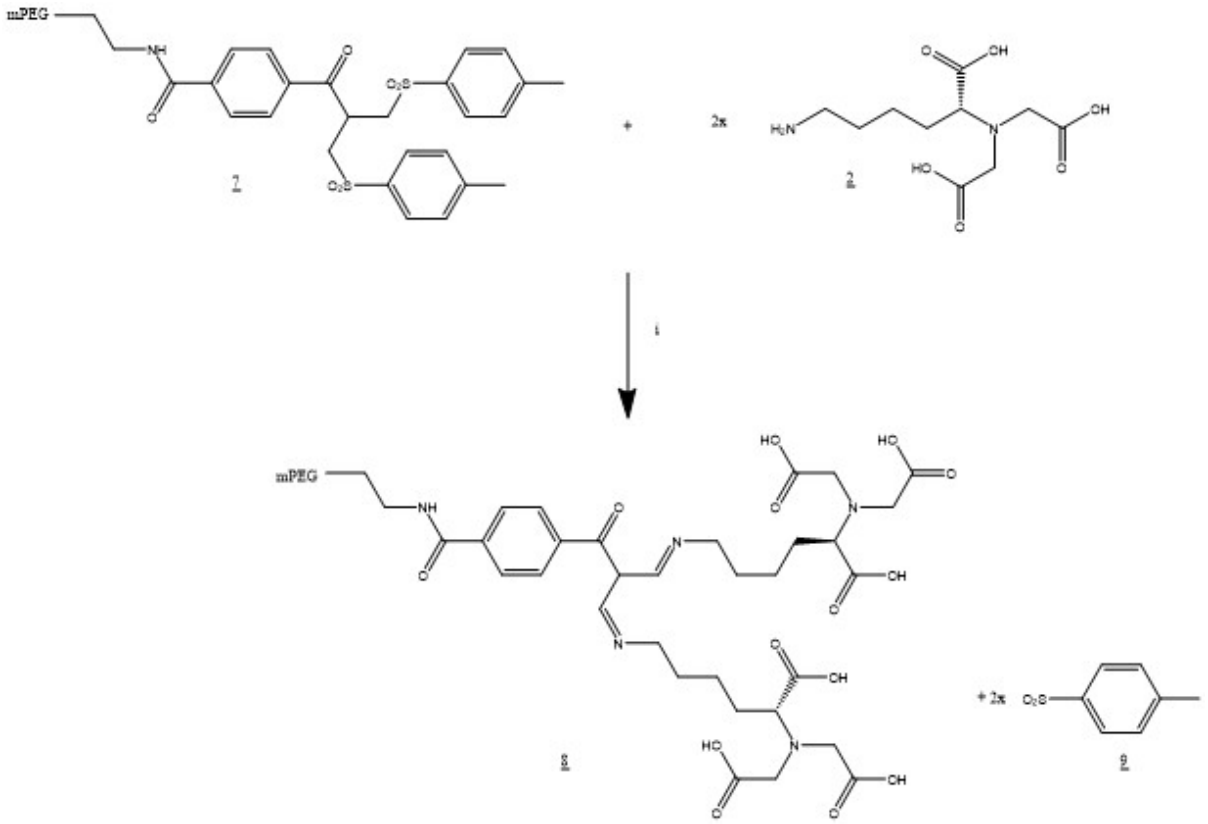



**Figure 5.4: $^{1}$H NMR Spectrum of Oxidised 10 kDa PEG B1 Sulfone**

Both 10 kDa and 5 kDa PEG B1 were conjugated to lysine-NTA and purified as described (Methods 2.8.9). The reaction scheme for the production of mPEG B1 lysine NTA is illustrated in Scheme 5.4 and the resulting conjugates were examined by $^{1}$H NMR spectroscopy (Figure 5.5) and then stored at -20°C under Argon, until required.




**Scheme 5.4: Conjugation of mPEG B1 to Lysine NTA.** i – 20 mM Tris buffer pH 9.0, 24 hrs, 37°C.

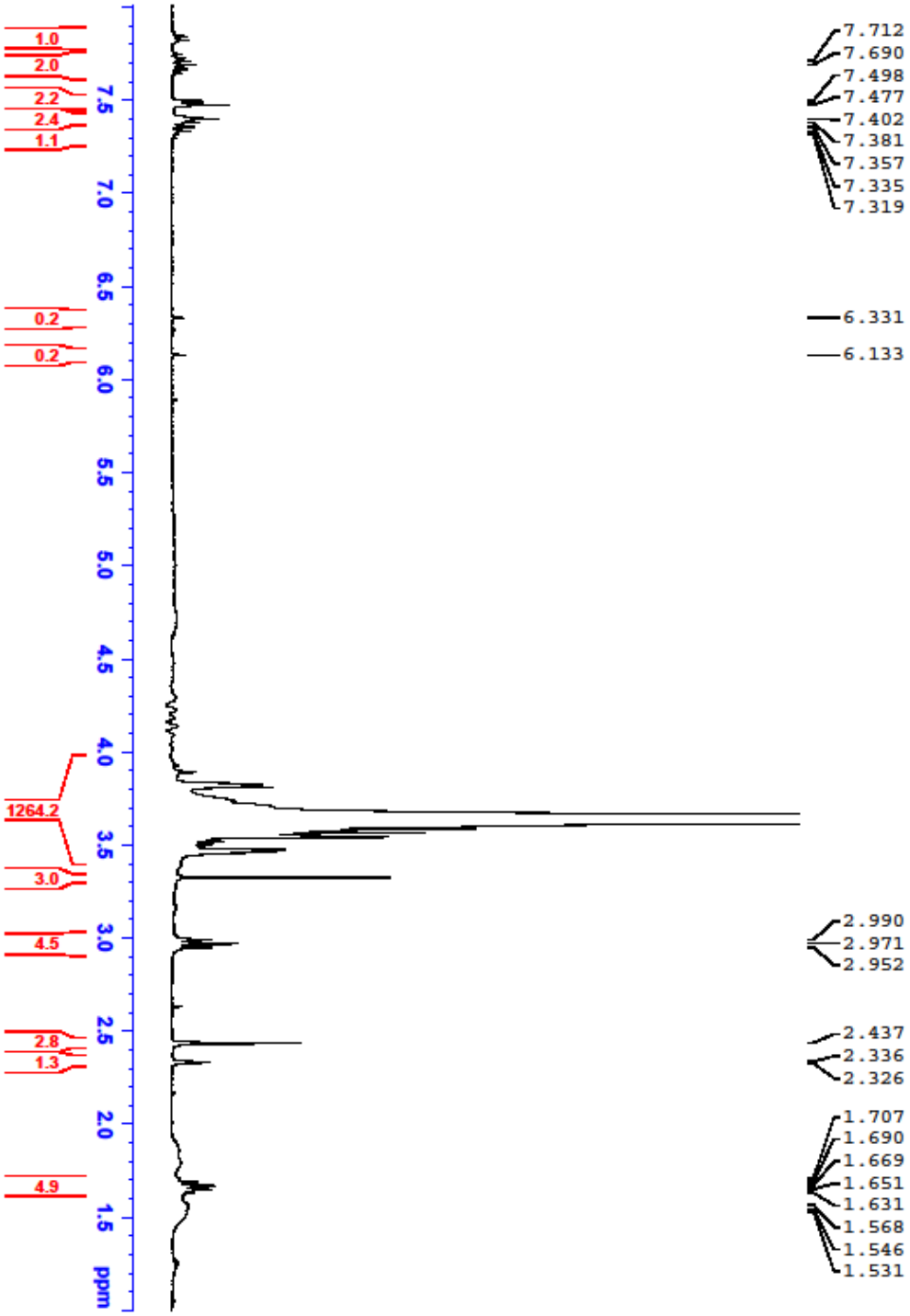
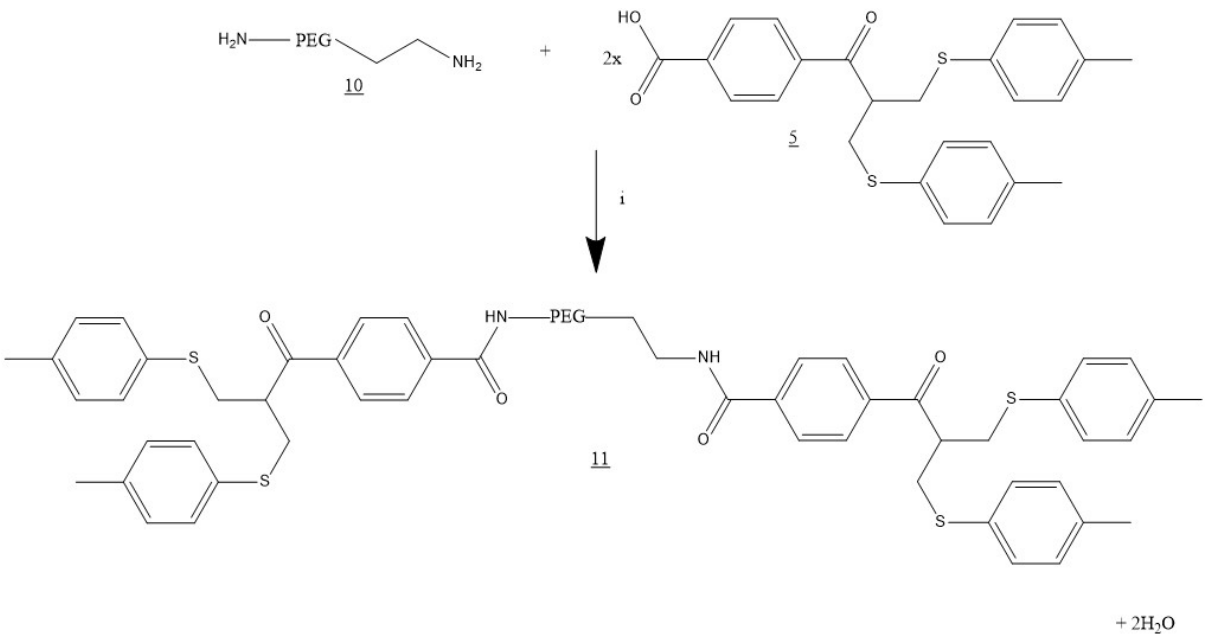



**Figure 5.5: $^{1}$H NMR Spectrum of 10 kDa PEG B1 lysine NTA**

### 5.2.3 Preparation of *bis*-B1-NTA

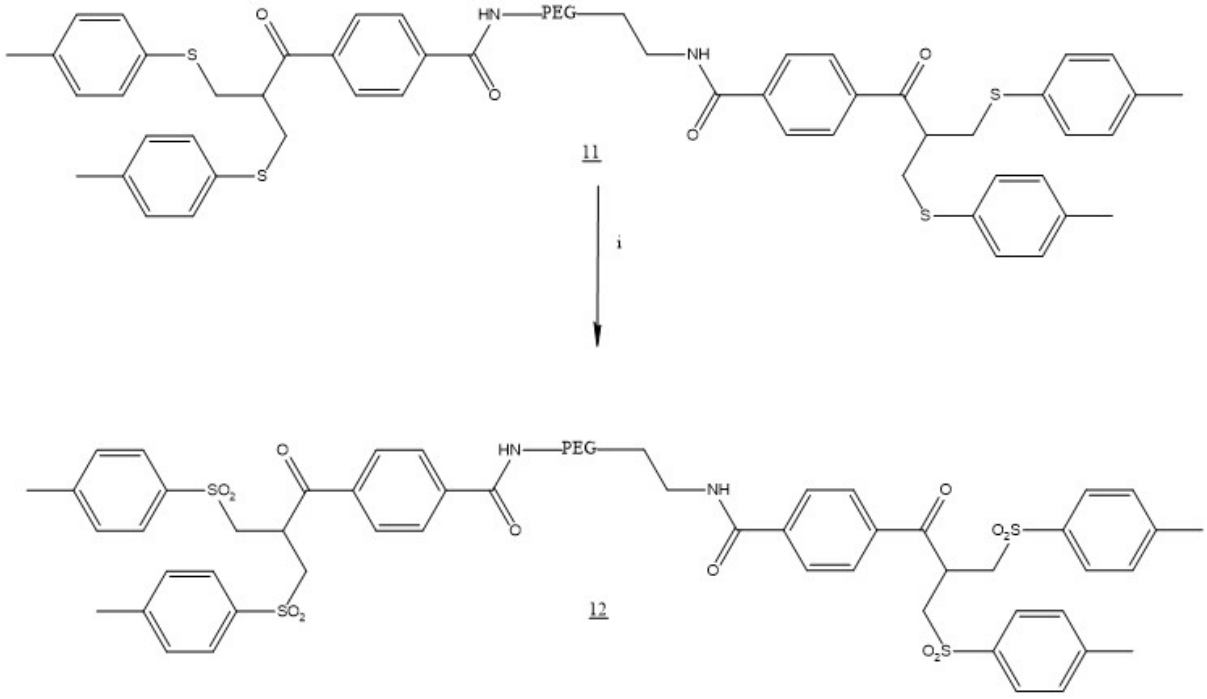
The carboxylic DB-*bis*-sulfide linker (B1, obtained from Prof. Steve Brocchini, UCL) was conjugated to both 10 kDa and 5 kDa amine PEG amine (JenKem Technology USA Ltd.) and purified as described (Methods 2.8.10). The reaction scheme for the conjugation of amine mPEG amine to B1 is illustrated in Scheme 5.5.




**Scheme 5.5: Coupling of Amine PEG Amine to B1 linker.** i – Anhydrous Dichloromethane under Argon, 48 hrs, room temperature.

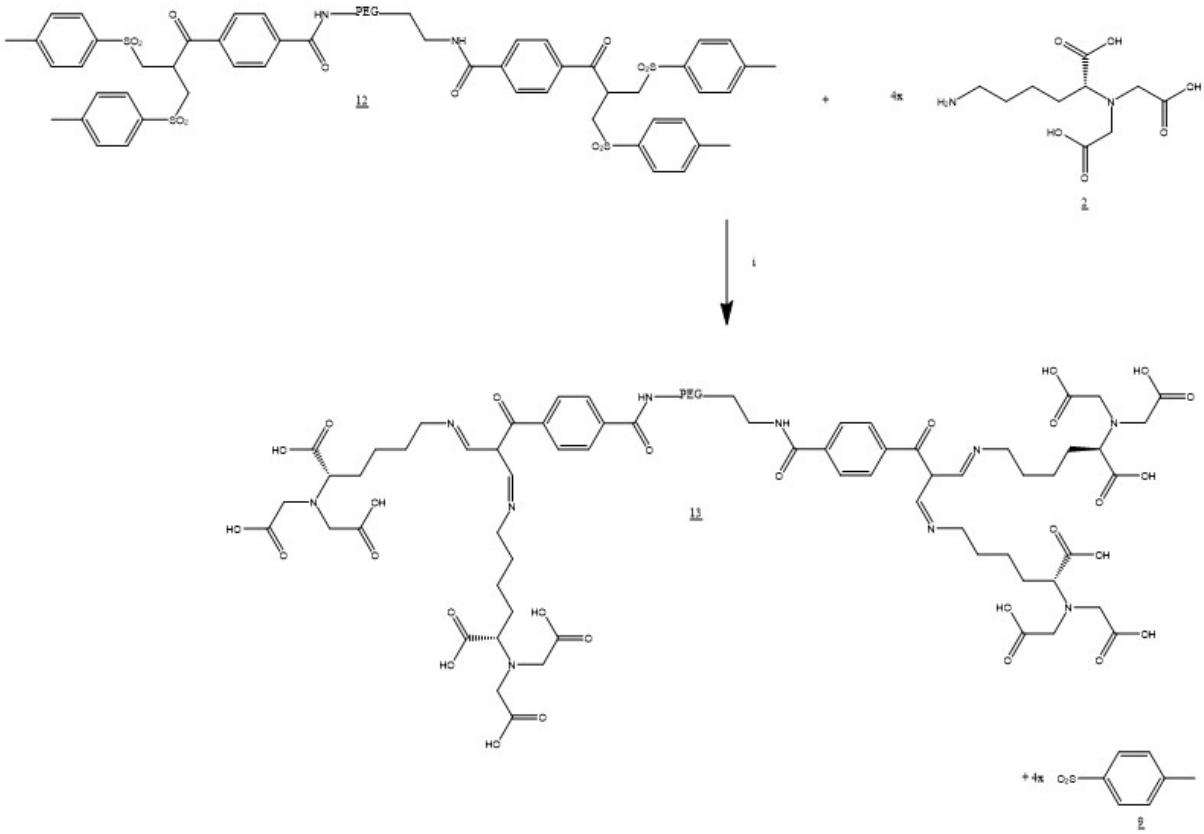
As previously, the resulting conjugates were then stored at -20°C under Argon, until required, since the sulfide is more stable than the corresponding sulfone and is thus more suitable for long term storage.

Prior to use, the required quantity of sulfide conjugate was oxidised to the corresponding sulfone with Oxone (Sigma) (Methods 2.8.11). A reaction scheme for the production of oxidised B1 is illustrated in Scheme 5.6.




**Scheme 5.6: Oxidation of PEG *Bis*-B1 Sulfide to Sulfone.** i – 12 Molar equivalents OXONE, methanol, overnight, room temperature.

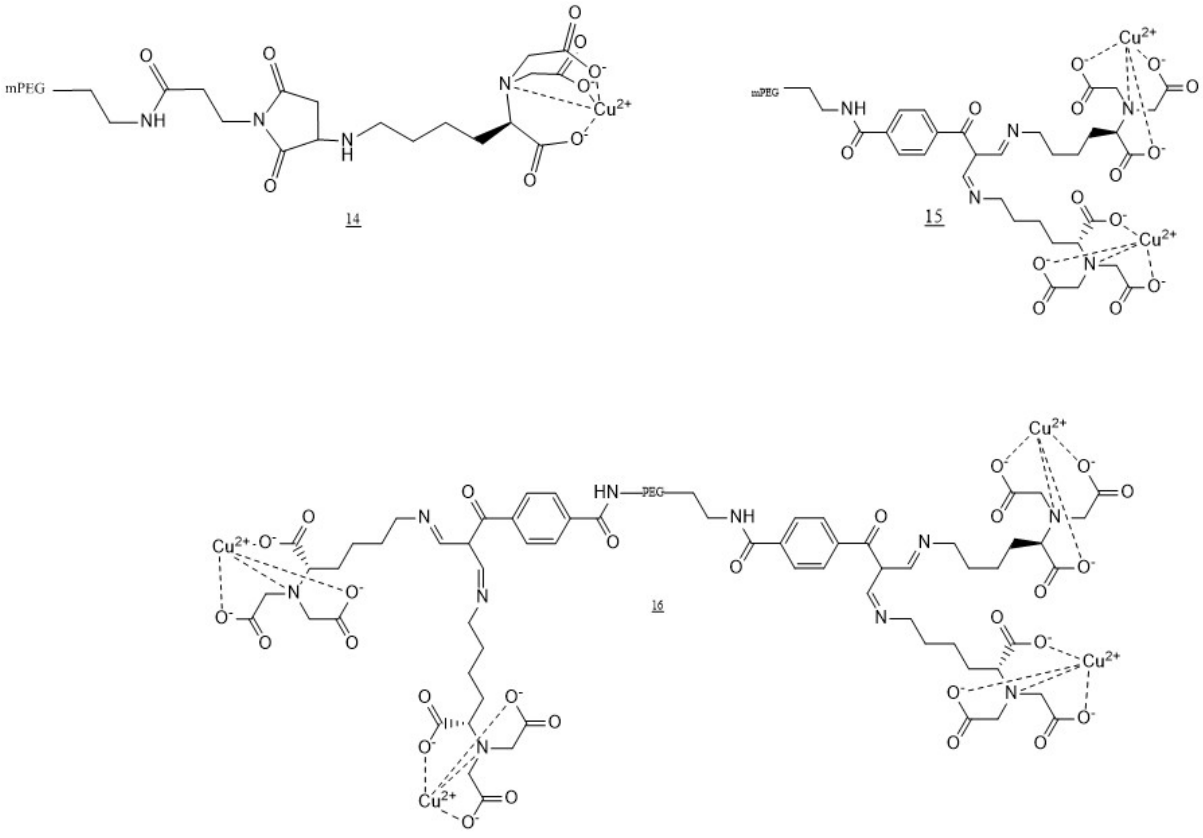
Both 10 kDa and 5 kDa PEG B1 were conjugated to lysine-NTA and purified as previously described (Methods 2.8.12). The reaction scheme for the production of PEG *bis* B1 lysine NTA is illustrated in Scheme 5.7. The resulting conjugates were then stored at -20°C under Argon.



**Scheme 5.7: Conjugation of mPEG bis B1 to Lysine NTA.** i – 20 mM Tris Buffer pH 9.0, 24 hrs, 37°C.

### 5.2.4 Coordination of the Chelating Metal Ion

In order for the PEGylated ligand to bind to the $His_6$ tag it was necessary for a divalent cation to be coordinated using the carboxyl groups of the NTA molecule. This coordination was performed using copper sulphate and was achieved as described in method (2.8.13). A buffer with pH 4.0 was chosen as this is the pKa of the carboxyl groups of NTA producing the necessary reactive $COO^-$ groups via deprotonation. The coordinated copper ion and PEGylated lysine NTA compounds are shown in Scheme 5.8.




**Scheme 5.8: Structures showing Methoxy PEG Maleimide lysine NTA (14), mPEG B1 lysine NTA (15) and PEG *bis* B1 lysine NTA (16) with coordinated copper ions.** i – 0.1 M Sodium Acetate buffer pH 4.0, Room Temperature, 2 hrs.

The resulting copper-coordinated PEG lysine NTA conjugates (5 kDa & 10 kDa mPEG mal lysine NTA $Cu^{2+}$; 5 kDa & 10 kDa PEG B1 lysine NTA $Cu^{2+}$ and 5 kDa & 10 kDa mPEG *bis* B1 lysine NTA $Cu^{2+}$) were each purified away from excess uncoordinated $CuSO_4$ using a PD10 desalting column into 10 mM Sodium Phosphate buffer pH 8.0 according to manufacturer's instructions.

## 5.3 Preliminary ATPS formation

A representative conjugate, 10 kDa mPEG B1 NTA $Cu^{2+}$ (Section 5.2.4) was then tested in model ATPS using purified $His_8$-tagged Interferon alpha 2a (2.5.1; $His_8$-IFN α-2a, Polytherics) as a target protein. This represents a clean, model system in which the affinity tag is simply required to pull the target protein from the DEX phase to the PEG phase without potential interference from cell debris, native proteins or other biomolecules in the sample to be partitioned.

In order to optimise ATPS composition, ATPS were formed according to Barbosa et al (2010). Using their most successful composition, a 1.5 g ATPS combining of 18% overall PEG 600, including 12% functionalised PEG (10 kDa PEG B1 NTA $Cu^{2+}$), and 14% DEX 40 was set up containing 50 µg purified, $His_8$-IFN α-2a. Surprisingly, this combination did not form two phases. Moreover, the large quantity of functionalised PEG (1000 molar equivalents) required to achieve 12% modified PEG, rendered this composition impractical for scale-up. A second combination described by Barbosa et al (2010) comprising 16% PEG 600, including 0.15% functionalised PEG, 14.3% DEX 40, and 100 µg IFN α-2a was also tested, as this composition had previously shown the best phase partitioning for pLacI-$His_6$-GFPuv protein. This combination also failed to form two distinct phases. Further examination of the protocol published by Barbosa *et al.,* revealed that whilst their ATPS contained 15 mM Tris HCl pH 7.4, the ATPS described herein actually contained 50 mM Tris HCl pH 7.4. Although this difference in buffer concentration is considerable, it is perhaps surprising that it is sufficient to disrupt ATPS formation.

To further investigate the influence of 50 mM Tris HCl pH 7.4, simple 20% and 10% PEG / DEX ATPS combinations were formed as described in Chapter 3. Two separate ATPS, each containing 50 µg $His_8$-IFN α-2a were constructed with compositions of 20% PEG 20000 / 20% DEX 100 and 10% PEG 20000 / 10% DEX 100. In the absence of both 50 mM Tris HCl pH 7.4 and functionalised PEG, ATPS formed successfully in both cases. Protein

concentration was measured in both layers (Table 5.1, 1; 2.8.14) Fifty molar equivalents of functionalised PEG (10 kDa PEG B1 NTA $Cu^{2+}$) were then added, the ATPS remixed and allowed to settle. Protein concentration measurements were then repeated (Table 5.1, 2; 2.8.14). Finally, Tris HCl buffer pH 7.4 was added to a concentration of 50 mM, the ATPS re-mixed, allowed to settle and protein concentration measurements were once again taken (Table 5.1, 3; 2.8.14).

| | | 10% Polymer Concentration | | 20% Polymer Concentration | |
|---|---|---|---|---|---|
| | | PEG 600 | DEX 40 | PEG 600 | DEX 40 |
| 1 | PEG / DEX ATPS with 50 µg $His_8$-IFN α-2a | 0.036 | 0.125 | 0.045 | 0.087 |
| 2 | As above, + 50 molar equiv. PEG B1 NTA $Cu^{2+}$ | 0.054 | 0.212 | 0.044 | 0.110 |
| 3 | As above + Tris HCl buffer pH 7.4 to 50 mM | 0.042 | 0.158 | 0.048 | 0.096 |

**Table 5.1: Absorbance Readings ($A^{595}$) of ATPS Layers using Bradford Reagent (2.8.14) to Analyse Protein Concentration**

The absorbance readings shown above do not indicate any partitioning to the upper PEG phase in either 10% or 20% ATPS combination. Upon addition of 50 molar equivalents of 10 kDa B1 PEG-NTA-$Cu^{2+}$ an increase can be seen in the absorbance of the lower DEX layer indicating an increase in protein concentration. However, it was noted visually that the volume of the lower DEX layer had decreased due to the increased quantity of PEG in the system causing water movement into the upper PEG phase and therefore produced an increased protein concentration. Most interestingly however, the addition of 50 mM Tris HCl pH 7.4 to the systems caused the dissociation of $Cu^{2+}$ from the NTA molecule. This was observed immediately as a bright blue colour appearing just above the interface of the ATPS (the PEG NTA complex is normally a pale blue colour). To confirm that the copper conjugate is incompatible with this buffer, 50 mM Tris HCl pH 7.4 was added to the copper conjugate in isolation. Again, an immediate, intense bright blue colour was observed, confirming

dissociation of the copper from the functionalised PEG (data not shown). This is perhaps not surprising in hindsight, given the potential for Tris and NTA to hydrogen bond to each other, owing to their similar structures. The differing structure of IDA may explain why Tris HCl was an acceptable buffer in the experiments conducted by Barbosa *et al.,* (2010). However, regardless of the differing concentrations of Tris buffer between the current study and that of Barbosa *et al.* (2010), it was clear that Tris would not be an acceptable buffer in an affinity ATPS where NTA (rather than IDA, as used by Barbosa et al) is the metal-co-ordinator.

To determine a suitable alternative buffer, 10 mM sodium phosphate buffer pH 8.0 (2.2.15) was added to functionalised PEG (10 kDa PEG B1 NTA $Cu^{2+}$). No colour change resulted (data not shown) and since pH 8.0 is optimal for poly-histidine tag binding (Schmitt *et al.,* 1993), 10 mM sodium phosphate buffer pH 8.0 was selected to buffer all future His-tag ATPS experiments.

Interestingly, close visual investigation of the PEG 600 / DEX 40 ATPS in the absence of Tris buffer (Table 5.1, 1) also showed that the pale blue 10 kDa PEG B1 NTA $Cu^{2+}$ complex partitions underneath the 600 Da PEG, just above the interface and so demonstrated a likely polymer gradient and suggested the need for samples to be taken from multiple positions within an ATPS.

## 5.4 Cell Lysate Partitioning Analysis

Having developed a suitable ATPS methodology for use with NTA $Cu^{2+}$ conjugates, attention now turned to determining whether a PEG-DEX ATPS containing functionalised, $Cu^{2+}$ conjugated PEG could partition $His_6$-tagged protein into the upper PEG phase. Accordingly, a large batch of *E. coli* expressing His-tagged GFPuv was prepared (2.8.2) and aliquoted into 50 ml samples. Cells were then harvested by centrifugation and stored frozen and then resuspended in 5 ml PBS containing 0.1% Triton X100 when required, to enable all subsequent experiments to be performed with the same batch of protein. To estimate GFPuv content, a cell aliquot was lysed and its total protein content measured. His-tagged GFPuv was then purified from 1 ml of this aliquot and the concentration of the GFPuv measured. As demonstrated in Table 5.2, GFPuv represented approximately 20% of total soluble protein.

| Total soluble protein (mg/ml) | $His_6$-tagged GFPuv (mg/ml) | % GFPuv |
|---|---|---|
| 4.2 | 0.9 | 21.4 |

**Table 5.2. Protein Concentrations of Total Soluble Protein and $His_6$-tagged GFPuv as determined using the Bradford Method (2.8.14).** Overall percentage of $His_6$-tagged GFPuv is also shown.

It was initially anticipated that fluorescence could be used to follow GFPuv partitioning. Unfortunately, early observations demonstrated that $Cu^{2+}$ ions quench GFPuv fluorescence (Lakowicz, 2006). UV-based detection of protein partitioning was therefore impossible and thus standard protein concentration analysis was required to detect total protein partitioning between the layers.

### 5.4.1 Affinity Conjugate Concentration

B1 PEG-NTA-$Cu^{2+}$ (10 kDa) was again selected as a representative conjugate to determine the effect of conjugate concentration on protein partitioning within a phosphate-buffered ATPS, as defined in section 5.3. Both 1% and 2% concentrations of this conjugate were examined for their effects on protein partitioning. Specifically, a phosphate-buffered ATPS with composition described by Barbosa *et al.* (2010) (18% (w/v) PEG 600 Da / 14% (w/v) DEX 40 kDa) was compared with 20% (w/v) combinations of these polymers, as described in Chapter 3. PEG 20000 Da was also analysed for protein partitioning as higher molecular weight polymers were favoured when partitioning all non-required proteins into the lower DEX phase (again, as determined in Chapter 3). In order to account for a possible PEG gradient, samples were analysed from both the top and bottom of each phase in addition to the interface. Owing to the large sample numbers, these experiments were analysed for protein concentration using the microplate method (Methods 2.8.15; Figure 5.6).

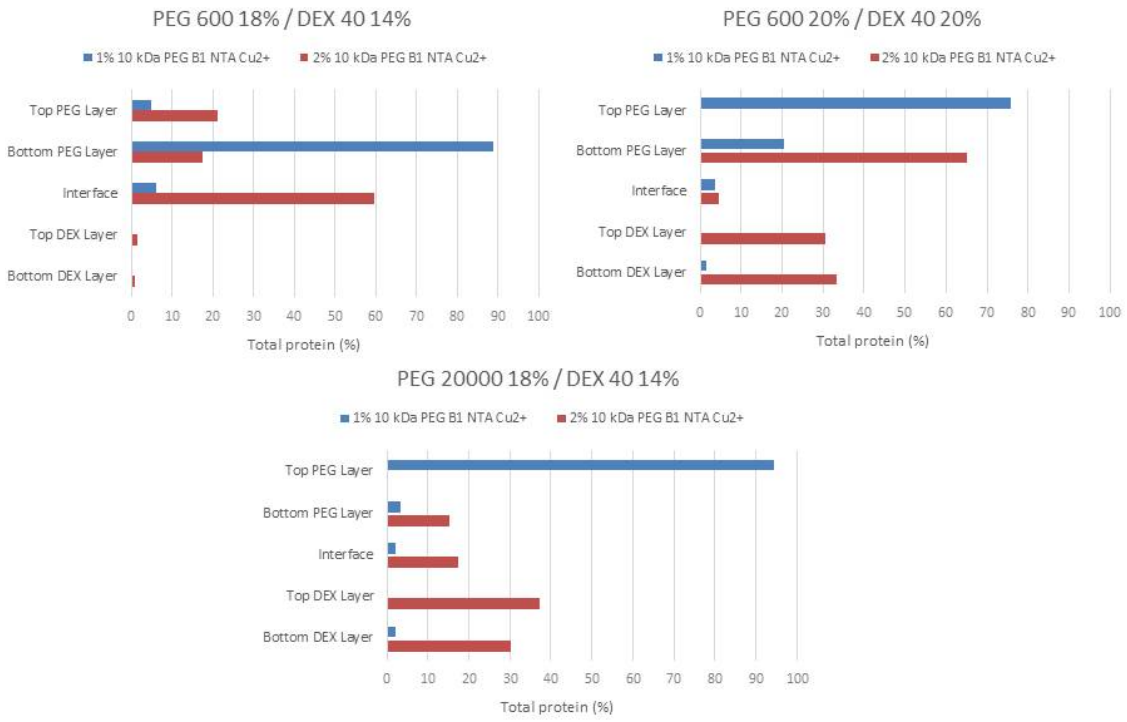



**Figure 5.6: Total protein partitioning in ATPS, composed as labelled supplemented with 1% 10 kDa PEG B1 NTA $Cu^{2+}$ or 2% 10 kDa PEG B1 NTA $Cu^{2+}$.** ATPS were set up as described in methods (2.8.16) using 100 µl of crude cell lysate expressing his-tagged GFPuv as a test sample.

Figure 5.6 demonstrates that 2% of the 10 kDa B1 PEG-NTA-$Cu^{2+}$ conjugate was preferable to 1%, since the 2% concentration resulted in a wider distribution of the proteins between the layers of the ATPS. In contrast, 1% conjugate tended to lead to the majority of protein being partitioned to the PEG layer, which would preclude separation of the required, His-tagged protein away from other cellular proteins, since only ~20% of total; protein can be His-tagged GFPuv (Table 5.2). The 2% concentration of conjugate was therefore selected for future study.

### 5.4.2 Analysis of 2% PEG NTA $Cu^{2+}$ conjugates

All six of the PEG NTA $Cu^{2+}$ conjugates made in section 5.2 were tested in phosphate-buffered ATPS to examine the protein partitioning of clarified cell lysates containing $His_6$-tagged green fluorescent protein ($His_6$-GFPuv). Owing to the previous distribution of total proteins (Figure 5.6), both 18% PEG 600 / 14% DEX 40 and 18% PEG 20000 / 14% DEX 40 ATPS were selected for further examination. The 20% PEG 600 / 20% DEX 40 identified from Chapter 3 was not selected for further study owing to the high percentage of total protein in the upper PEG phase of this system (Figure 5.6) which considerably exceeds the estimated content of GFPuv within the total protein of the cell lysate. Instead, Barbosa et al's 16% PEG 600 / 14.3% DEX 40 was selected as the third alternative ATPS.

Initially, to again investigate the potential for polymer gradient formation, ATPS were sampled at multiple levels (Figures 5.7; 5.8; 5.9).

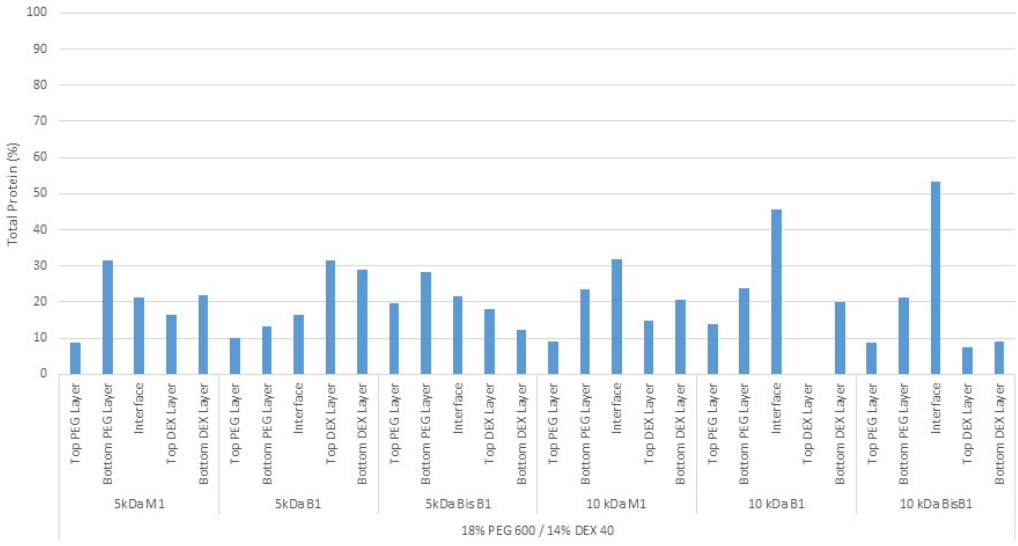



**Figure 5.7. Protein partitioning of Total Protein in ATPS composed of 18% PEG 600 / 14% DEX 40 with 2% of overall PEG being functionalised (PEG-NTA-$Cu^{2+}$) as described.** 5 µl samples were taken from both the top and bottom of the PEG and DEX phases along with a sample from the interface. These were analysed using the microplate Bradford assay as described in method 2.8.15.

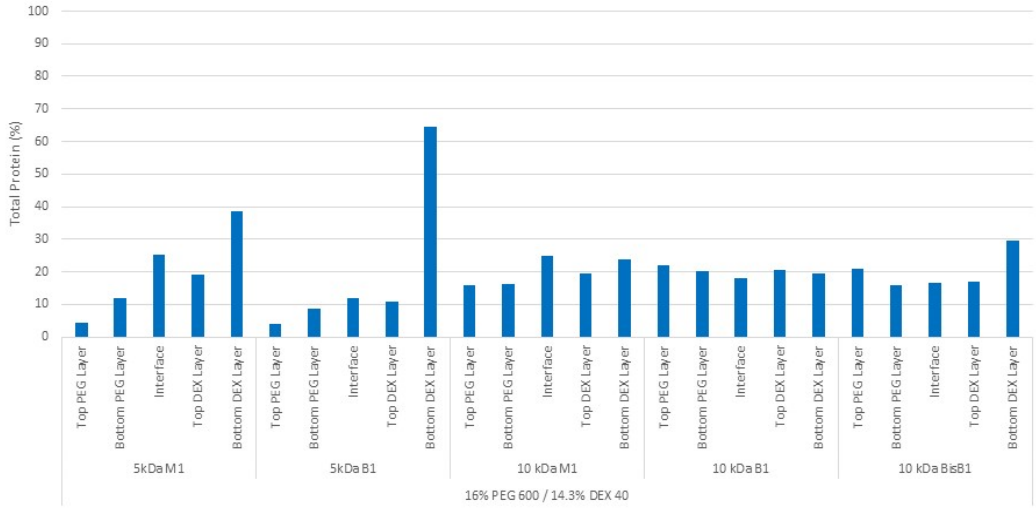



**Figure 5.8. Protein partitioning of Total Protein in ATPS composed of 16% PEG 600 / 14.3% DEX 40 with 2% of overall PEG being functionalised (PEG-NTA-$Cu^{2+}$) as described.** 5 µl samples were taken from both the top and bottom of the PEG and DEX phases along with a sample from the interface. These were analysed using the microplate Bradford assay as described in method 2.8.15.

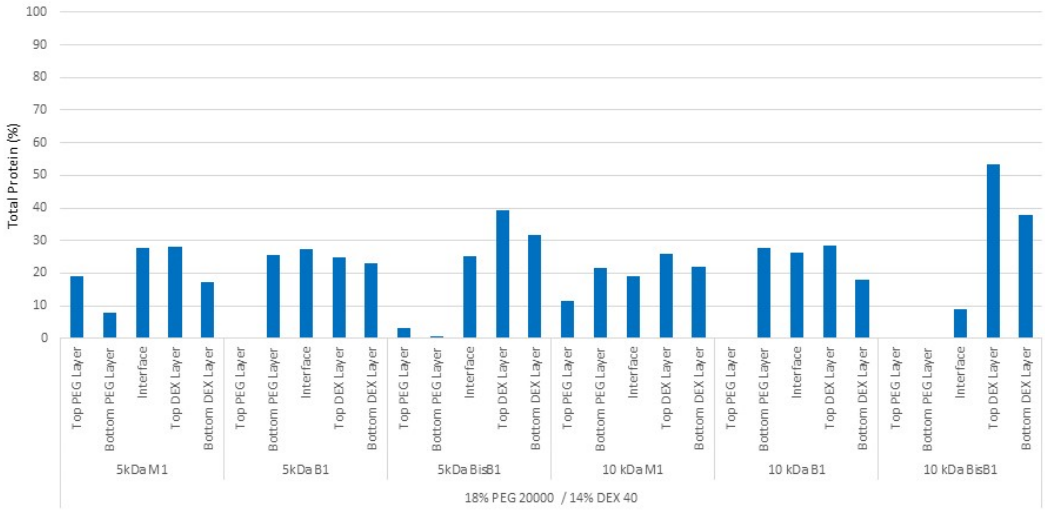



**Figure 5.9. Protein partitioning of Total Protein in ATPS composed of 18% PEG 20000 / 14% DEX 40 with 2% of overall PEG being functionalised (PEG-NTA-$Cu^{2+}$) as described.** 5 µl samples were taken from both the top and bottom of the PEG and DEX phases along with a sample from the interface. These were analysed using the microplate Bradford assay as described in method 2.8.15.

Examination of Figures 5.7, 5.8 and 5.9 demonstrate no clear trends. Therefore total protein concentrations in the PEG, DEX and interface layers were summed and examined graphically (Figure 5.10; 5.11; 5.12).

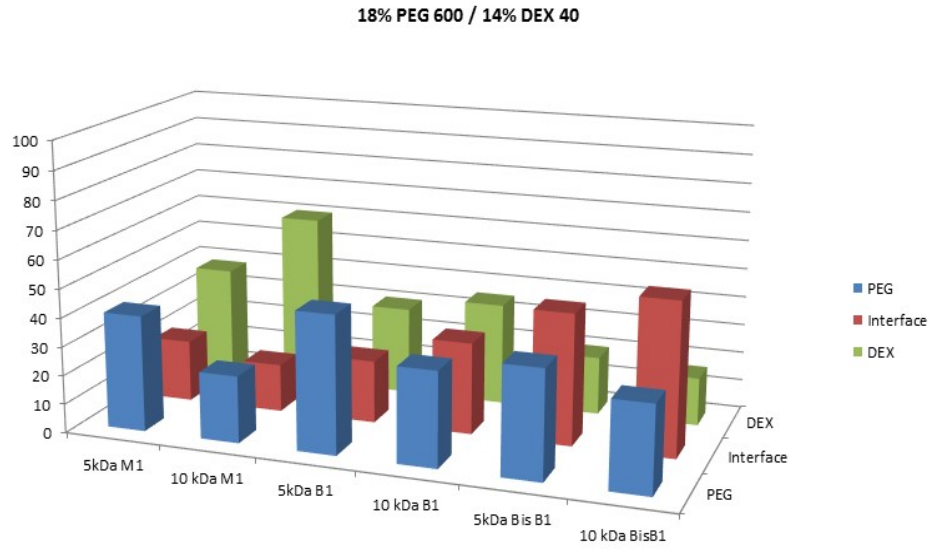



**Figure 5.10. Protein partitioning of Total Protein in ATPS composed of 18% PEG 600 / 14% DEX 40 with 2% of overall PEG being functionalised (PEG-NTA-$Cu^{2+}$) as described.** 5 µl samples were taken from both the top and bottom of the PEG and DEX phases along with a sample from the interface. These were analysed using the microplate Bradford assay as described in method 2.8.15. Percentage values of total protein in the upper and lower of each phase were combined to show an overview of protein distribution.

Examination of Figure 5.10 shows a trend towards a spread of protein over both PEG and DEX layers as well as at the interface. As approximately 20% of overall protein is $His_6$-tagged GFPuv all 18% PEG 600 / 14% DEX 40 ATPS, with the exception of the system containing 10 kDa M1 PEG-NTA-$Cu^{2+}$, demonstrated non-specific partitioning of protein into the PEG phase. The system containing 10 kDa M1 PEG-NTA-$Cu^{2+}$ demonstrated a total protein partitioning in the PEG phase of 23%, and whilst this is still a couple of percent higher than the approximate GFPuv content, this system was closest to the required partitioning profile suggesting that further analysis of the proteins in each layer was required (Figure 5.14).

A secondary trend noticeable within this figure is an increase in protein partitioning at the interface and a reduction of protein within the DEX phase when the functionalised PEG used is 10 kDa B1 PEG-NTA-$Cu^{2+}$, 5 kDa *bis*-B1 PEG-NTA-$Cu^{2+}$ and 10 kDa *bis*-B1 PEG-NTA-$Cu^{2+}$.

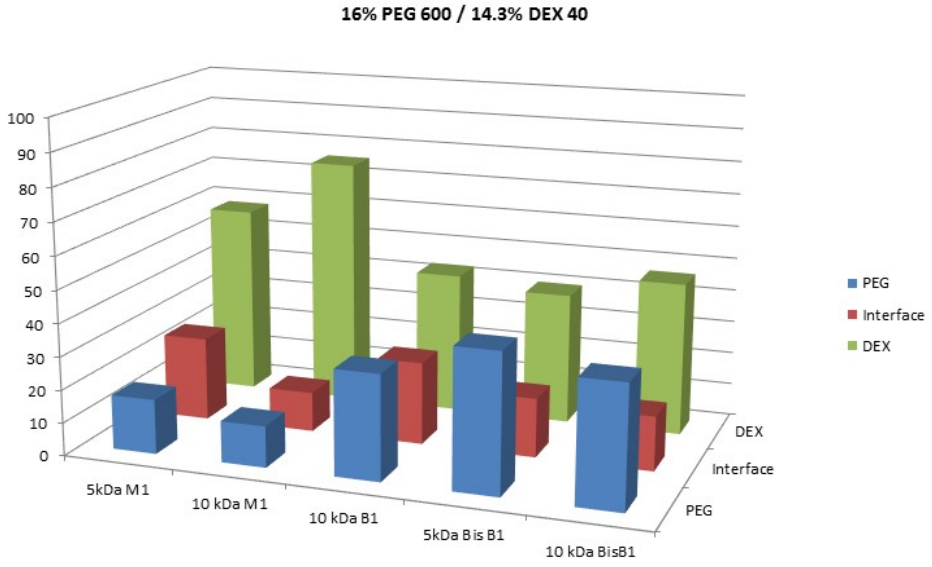



**Figure 5.11. Protein partitioning of Total Protein in ATPS composed of 16% PEG 600 / 14.3% DEX 40 with 2% of overall PEG being functionalised (PEG-NTA-$Cu^{2+}$) as described.** 5 µl samples were taken from both the top and bottom of the PEG and DEX phases along with a sample from the interface. These were analysed using the microplate Bradford assay as described in method 2.8.15. Percentage values of total protein in the upper and lower of each phase were combined to show an overview of protein distribution.

Examination of Figure 5.11 shows a high level of protein partitioning to the lower DEX phase in all ATPS formed, especially within those ATPS formed with 5 kDa and 10 kDa M1 PEG-NTA-$Cu^{2+}$. These two ATPS show low levels of protein partitioning to the upper PEG phase, 17% and 13% respectively, with the remaining systems demonstrating protein partitioning higher than the available ~20% $His_6$-tagged GFPuv. These values suggest that there is non-specific binding and partitioning occurring to the PEG phase. The system formed using 5 kDa M1 PEG-NTA-$Cu^{2+}$ was selected for further protein analysis by SDS-PAGE (Figure 5.15) as the percentage protein partitioned to the PEG phase was close to the $His_6$-tagged GFPuv cleared cell lysate content.

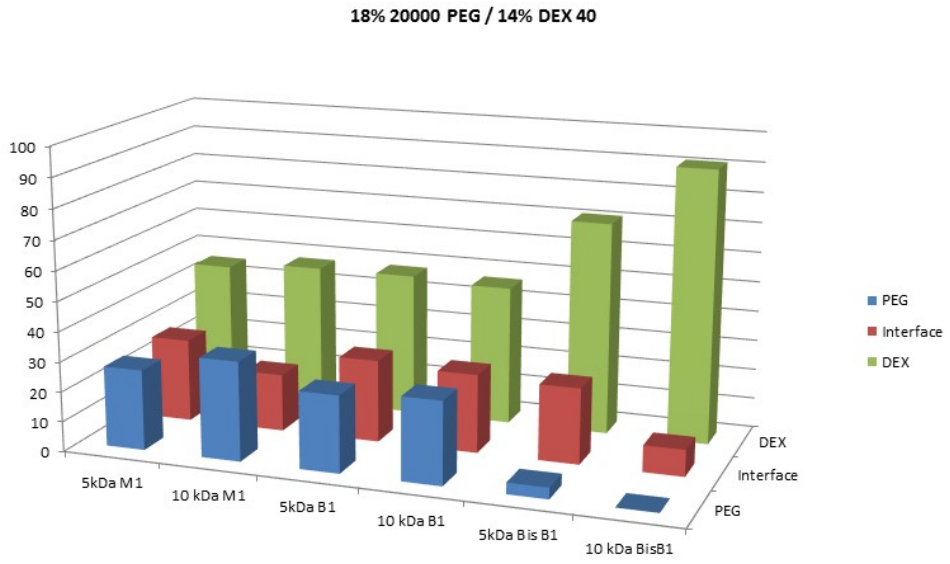



**Figure 5.12. Protein partitioning of Total Protein in ATPS composed of 18% PEG 20000 / 14% DEX 40 with 2% of overall PEG being functionalised (PEG-NTA-$Cu^{2+}$) as described.** 5 µl samples were taken from both the top and bottom of the PEG and DEX phases along with a sample from the interface. These were analysed using the microplate Bradford assay as described in method 2.8.15. Percentage values of total protein in the upper and lower of each phase were combined to show an overview of protein distribution.

As in Figure 5.10, Figure 5.12 also shows a trend towards a spread of protein partitioning across the layers. However, in contrast to Figure 5.10, these ATPS combinations of 18% PEG 20000 / 14% DEX 40 with 2% functionalised PEG demonstrate protein partitioning predominantly to the DEX phase with greater than 50% of total protein in the DEX layer of all ATPS rising to almost 80% when using 5 kDa *bis*-B1 PEG-NTA-$Cu^{2+}$ and almost 90% of total protein when using 10 kDa *bis*-B1 PEG-NTA-$Cu^{2+}$. Correspondingly, there is also a large decrease in protein partitioning to the PEG phase when this rise in partitioning to the DEX phase is seen.

When considering the His6-tagged GFPuv content of overall protein is approximately 20%, it is interesting to note the protein contents within PEG layers of ATPS formed with 5 kDa M1 PEG-NTA-$Cu^{2+}$, 5 kDa B1 PEG-NTA-$Cu^{2+}$ and 10 kDa B1 PEG-NTA-$Cu^{2+}$ are 27%, 26% and 28% respectively. These ATPS were also selected for further protein analysis by SDS-PAGE (Figures 5.13 and 5.15).

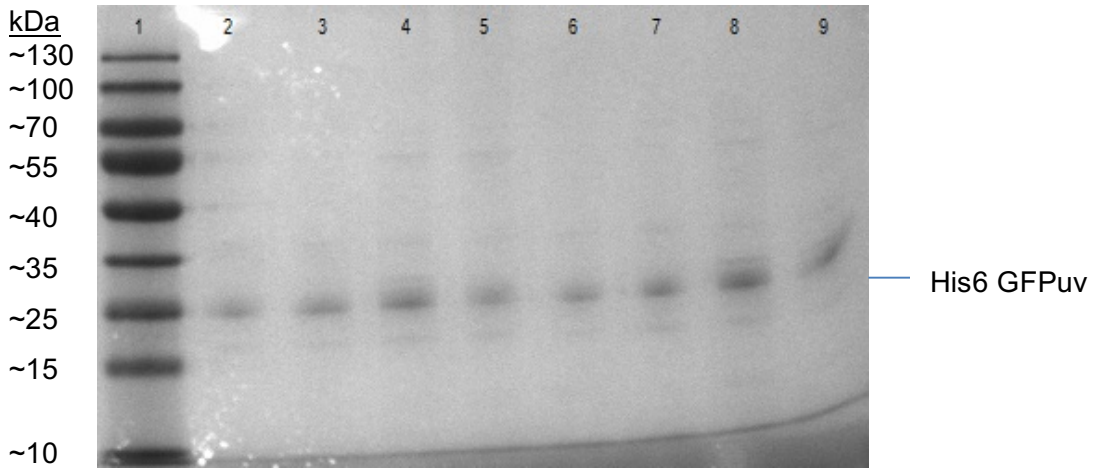



**Figure 5.13. SDS-PAGE of Protein Content within ATPS Layers from 18% PEG 600 / 14% DEX 40 with 2% 10 kDa B1 PEG-NTA-$Cu^{2+}$ and 18% PEG 600 / 14% DEX 40 with 2% 5 kDa B1 PEG-NTA-$Cu^{2+}$ Systems.** The gel was prepared and run as described in methods 2.6.2 with 20 µl of sample loaded into each well. The gel was stained using Coomassie Brilliant Blue.His6-tagged GFPuv (~ 27 kDa) is shown as a dense band. Lane 1 – 5 µl PageRuler™ Prestained Protein Ladder (Thermo Scientific); Lanes 2 – 5 – ATPS 18% PEG 600 / 14% DEX 40 with 2% 10 kDa B1 PEG-NTA-$Cu^{2+}$; 2 – Upper PEG Phase, 3 – Lower PEG Phase, 4 – Upper DEX Phase, 5 – Lower DEX Phase; Lanes 6 – 9 – ATPS 18% PEG 600 / 14% DEX 40 with 2% 5 kDa B1 PEG-NTA-$Cu^{2+}$; 6 – Upper PEG Phase, 7 – Lower PEG Phase, 8 – Upper DEX Phase, 9 – Lower DEX Phase.

Figure 5.13 above shows His6-tagged GFPuv (indicated by the dark bands at approximately 25 kDa) spread across both layers with soluble native proteins also present in both layers. The bands representing soluble native proteins are not as distinctive and can be described as a smear which appears darker when more protein is present in the sample. In both ATPS, a greater concentration of protein can be seen in the DEX samples than PEG samples but there is still contamination of native cellular proteins in the phase.

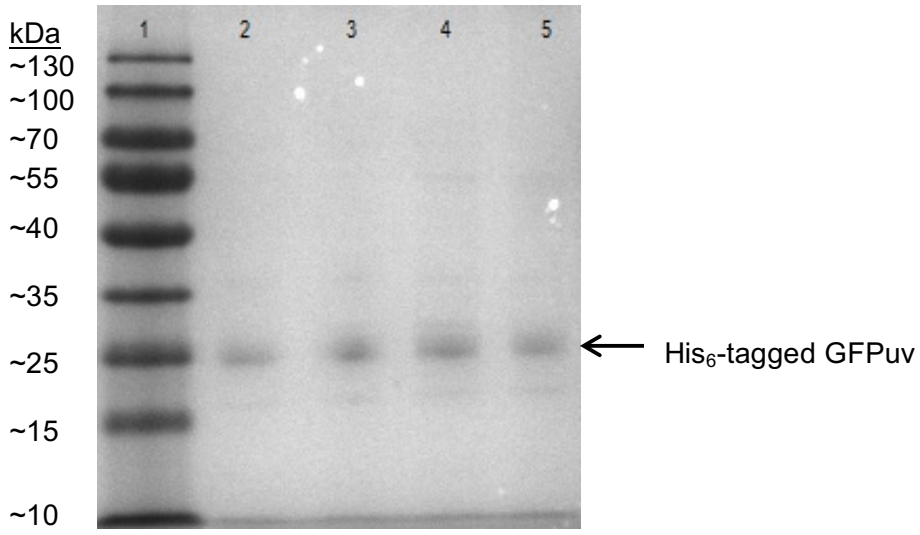



**Figure 5.14. SDS-PAGE of Protein Content within ATPS Layers from ATPS 18% PEG 20000 / 14% DEX 40 with 2% 10 kDa M1 PEG-NTA-$Cu^{2+}$Systems.** The gel was prepared and run as described in methods 2.6.2 with 20 µl of sample loaded into each well. The gel was stained using Coomassie Brilliant Blue. Lane 1 – 5 µl PageRuler™ Prestained Protein Ladder (Thermo Scientific); Lanes 2 – 5 – ATPS 18% PEG 20000 / 14% DEX 40 with 2% 10 kDa M1 PEG-NTA-$Cu^{2+}$; 2 – Upper PEG Phase, 3 – Lower PEG Phase, 4 – Upper DEX Phase, 5 – Lower DEX Phase.

As in Figure 5.13, $His_6$-tagged GFPuv can be seen in all four samples from the ATPS indicating a spread across both the PEG and DEX layers. However, in contrast with Figure 5.13, there are only two contaminating bands present in lane 2 from the native soluble proteins of *E.coli* expression cells in the upper part of the PEG phase. Lane 3 shows slightly more contamination of native proteins in the lower part of the PEG layer of this ATPS. Thus, although the yield of $His_6$-tagged GFPuv in the upper PEG phase is comparatively poor, relatively good purification has been obtained in this sample.

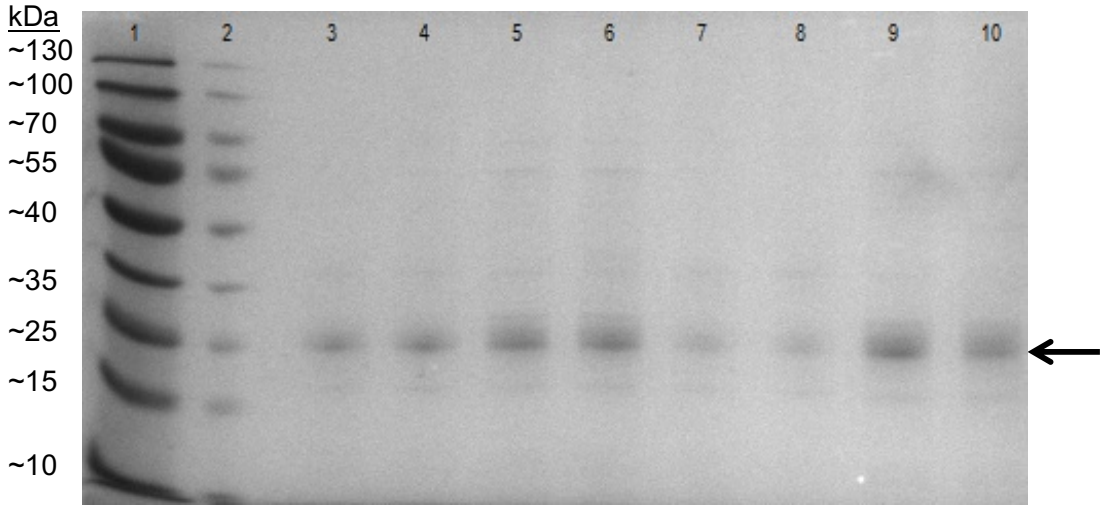



**Figure 5.15. SDS-PAGE of Protein Content within ATPS Layers from ATPS 18% PEG 20000 / 14% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$ and ATPS 16% PEG 600 / 14.3% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$Systems.** The gel was prepared and run as described in methods 2.6.2 with 20 µl of sample loaded into each well. The gel was stained using Coomassie Brilliant Blue. Arrow indicates $His_6$-tagged GFPuv. Lanes 1 & 2 – 5 µl PageRuler™ Prestained Protein Ladder (Thermo Scientific) & overflow; Lanes 3 – 6 – ATPS 18% PEG 20000 / 14% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$; 3 – Upper PEG Phase, 4 – Lower PEG Phase, 5 – Upper DEX Phase, 6 – Lower DEX Phase; Lanes 7 – 10 – ATPS 16% PEG 600 / 14.3% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$; 7 – Upper PEG Phase, 8 – Lower PEG Phase, 9 – Upper DEX Phase, 10 – Lower DEX Phase.

Examination of Figure 5.15 also demonstrates a spread of $His_6$-tagged GFPuv across all layers in both ATPS described. It is noticable that the system comprising 16% PEG 600 / 14.3% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$ (lanes 7 – 10) partitions the majority of total protein content to the DEX layer. In contrast, there is a greater degree of $His_6$-tagged GFPuv partitioning into the PEG phase of the ATPS shown in lanes 3 – 6 (18% PEG 20000 / 14% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$).

Similarly to Figure 5.14, the presence of native soluble proteins is far reduced in the PEG phase but not totally eliminated. Only two contaminant protein bands are present in the PEG phase of these two ATPS (lanes 3, 4, 7 and 8; Figure 5.15), and the relative intensity of these contaminant bands in comparison to $His_6$-tagged GFPuv is low.

## 5.5 Discussion

In section 5.3 it was determined that Tris buffer was an unsuitable buffer to use for affinity ATPS with PEGylated NTA compounds with coordinated copper ions as demonstrated in table 5.1 where the addition of Tris buffer caused the dissociation of coordinated copper ions. This is due to Tris being similar in structure to NTA but as Tris has a pKa of 8.07, close to the pH of the buffer, the carboxyl groups will be deprotonated and therefore compete with the NTA to coordinate the copper ions. To avoid dissociation of the copper ion, a phosphate buffer (pH 8.0) was selected for use in further experiments within this chapter.

The results in section 5.4.1 indicate that a 2% concentration of functionalised PEG was preferable over 1% functionalised PEG as the ATPS containing 2% concentration of functionalised PEG demonstrated a spread of proteins over the two layers (Figure 5.6). This was beneficial in comparison to 1% functionalised PEG concentration as these systems partitioned the majority of proteins to the PEG phase. These results were subsequently carried through into the experiments performed in section 5.4.2.

Successful development of an affinity ATPS would be indicated by partitioning of the protein of interest to the PEG phase using the PEGylated NTA-$Cu^{2+}$ affinity conjugates. This goal has been partially achieved. Purification is always a balance of purity of individual fractions against yield of the required protein. Whilst yield is poor (the results in this section showed that the $His_6$-tagged GFPuv was spread across both of the layers within all of the ATPS tested- see figures 5.7, 5.8 and 5.9), the purity is relatively good in all or part of the PEG phase of three ATPS tested (16% PEG 600 / 14.3% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$ shown in figure 5.15, 18% PEG 20000 / 14% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$ shown in figure 5.15 and 18% PEG 20000 / 14% DEX 40 with 2% 10 kDa M1 PEG-NTA-$Cu^{2+}$ shown in figure 5.14) with only two minor contaminants detectable by SDS-PAGE (Section 5.4).

Whilst the ATPS comprising 16% PEG 600 / 14.3% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$ (Figure 5.15) produced a relatively clean separation of $His_6$-tagged GFPuv into the PEG phase, the yield of protein was very low in this system. In contrast, the ATPS comprised of 18% PEG 20000 / 14% DEX 40 with 2% 10 kDa M1 PEG-NTA-$Cu^{2+}$ demonstrated a higher yield of required protein, again with only two contaminating proteins in the upper part of the PEG phase. A similar yield of protein was obtained with the ATPS comprising 18% PEG 20000 / 14% DEX 40 with 2% 5 kDa M1 PEG-NTA-$Cu^{2+}$, in this case, with purity being maintained throughout the whole PEG phase.

When investigating which metal ion to use for the purification of a therapeutic protein, it was prudent to identify which metal ions would pose a risk to the patient if contamination occurred during the production process. The European Medicines Agency (EMA) have three classifications of metal ranging from class one – a significant safety concern to class three – minimal safety concern with class two being a low safety concern. Copper was chosen for this methodology as it is in class two and therefore has a lower safety concern when administered to the body in comparison to the class one significant safety concern associated with nickel, the other commonly used ion in metal ion affinity chromatography (European Medicines Agency, 2007). Copper has been assigned class two status as it is a functional component of cuproenzymes for example cytochrome c oxidase and ascorbic acid oxidase as well as playing a key role in the scavenging of radicals in the body reducing oxidative stress (Stanstead, HH., 1995). According the guidelines released by the EMA in 2007, oral exposure to copper is limited to 2500 µg / day and a concentration of 250 ppm. Parenteral exposure to copper has a reduced limit 250 µg / day and a concentration of 25 ppm. When following good manufacturing practice, levels of excess copper ions should be kept to a minimum via the use of filtration methods isolating the free copper ions away from the functionalised PEG with conjugated metal ions. It is however noted that a concentration of less than 5 ppm in an administered medicine is unlikely to cause an allergic response even in highly sensitive patients (Uter *et al.,* 1995)

Due to the high costs of therapeutic protein production it is advantageous to keep costs to a minimum especially when considering the scaling up of a method to industrial levels. An advantage of using copper ions conjugated to functionalised PEG when compared to the previous protein fusion methods is the ability to reuse the copper bound functionalised PEG as well as recycling any filtered, unbound copper to freshly prepared PEG-NTA conjugates. This ability to reuse and recycle any unbound copper would allow for an initially expensive component to develop a lower cost per use.

# Chapter 6. Conclusion

The initial study was performed with a systematic approach in order to determine which w/v concentrations of polymer molecular weights formed two distinct phases, and the optimal conditions for this to happen. This was due to the mechanisms for partitioning not currently being understood in the literature. The present systematic review of the literature identified that the following gaps in knowledge required further attention: the need for additional research into the exact mechanisms and influences on protein partitioning as well as the development of a large scale purification process from lab bench scale. The selection of PEG and Dextran have therefore provided the most comprehensive information about separation to date and provide the most suitable materials to effect the most efficient separation when considering yield, cost and purity.

When the present study examined specifically PEG & Dextran combinations in chapter 3, it was evident that when the 10% w/v polymer concentrations were combined only 10 ATPS were formed compared to 25 ATPS when formed using 25% w/v concentrations (Table 3.1). Due to the high viscosity of the 25% w/v polymer, particularly the dextran solutions, it was difficult to accurately pipette the required volume. It was therefore decided that a lower concentration, 20% w/v, would be used as this would maintain the high occurrence of ATPS formation. An advantage to using a simple 20% w/v polymer solution would have been easily and quickly produced in a laboratory or industrial scale with minimal measuring although having an increased cost due to quantities of polymer being used. Should this method have worked, this would have been a distinct advantage of using affinity ATPS as a generic affinity protein purification system through not only ease of use, but the potential for the production of a high purity protein.

As previously discussed in section 5.5 it is essential to have a high purity within medicines to avoid any unwanted side effects or reactions. Whilst some medicines need to be highly pure, such as those delivered intravenously (for example - intravenous nutrition) with

subcutaneous injectable medicines having a slightly higher level of impurity tolerance (e.g. – insulin), whilst some medicines can have a higher tolerance of impurities such as those delivered orally. Whilst orally administered drugs can have a higher impurity tolerance, the dosages have to be higher due to the process of digestion occurring in the stomach and intestines before a proportion of the drug can be absorbed. An advantage to administering therapeutic proteins parenterally is an increased efficacy of the protein as it will not be digested in the digestive system but the purity of the end product will have to be much higher to gain EMA/FDA approval. There does, however, need to be additional consideration for any impurities that may cause an anaphylactic response, such as nut allergies that can be triggered from contamination on a production line. In this instance, it was determined that the level of contamination from copper ions would be unlikely to trigger an anaphylactic response as long as the levels were lower than 5 ppm (Uter *et al.,* 1995).

The use of PEG and dextran in this study have the obvious advantage that any trivial degree of contamination by these two polymers can be dealt with by the body. This is demonstrated by the regular use of PEG as a bio conjugate to improve the in vivo circulation half-life of the drug through reducing urinary excretion of the molecule (Yang *et al.,* 2004) as well as enabling reduced dosages to be administered (Alconcel, S., *et al*, 2011). In order to remove PEG from the body it is metabolised via oxidation of the alcohol groups to form a carboxylic acid, usually performed by alcohol dehydrogenase enzymes (Herold *et al.,* 1989). PEG metabolism also appears to be driven by molecular weight with a higher proportion of small molecular weight polymers (less than 5000 Da) being metabolised (Schaffer *et al.,* 1950). In contrast, larger molecular weight PEG polymers (above 5000 Da) tend to be removed by biliary elimination (Friman *et al.,* 1993, Webster *et al.,* 2007). Dextran is also regularly used within medicine as a volume expander when hypovolemic shock is occurring. As dextran is simply a polymer constructed of glucose monomer units, the body can easily handle and digest the polymer, via glucosidases, and utilise it for normal bodily processes such as respiration. The biocompatibility of dextran is of particular importance for future development of this method as a second ATPS could be used alongside a PEGylated protease to cleave the target

therapeutic protein from the histidine tag and metal ion complex. This cleaved target protein would then settle into the lower dextran layer and would not require further purification for use.

To ensure that the analysis of protein concentration in the different ATPS layers was reliable, it was important to determine whether the presence of PEG or dextran would affect a Bradford assay (section 3.3). Figure 3.1 and table 3.2 both demonstrated that the presence of polymer molecules produced a negligible effect as supported by Barbosa *et al.,* 2009. The Bradford assay was therefore considered suitable for use in future ATPS protein concentration analysis. This is advantageous for future development of the method as the Bradford assay is a well-established and industrially recognised method of protein quantification.

Before the addition of cell lysate or purified protein to the 20% w/v polymer solutions 20 ATPS were formed when left to settle. There was shown to be a material effect on ATPS formation by the addition of cell lysate or purified protein as the number of ATPS formed was reduced to just 15 systems. This could be due to the addition of 100 µl of $H_2O$ therefore adjusting the overall polymer concentrations and therefore water densities of the two layers. The change in ATPS formation could also be due to the surface properties of any proteins or cellular material being added interacting with the polymers and therefore reducing the interactions between polymer and $H_2O$ (Israelachvili, J., 1997).

There was also a small difference noted in the number of ATPS that were formed when either settled or centrifuged with cell lysate or purified protein. When left to settle 15 ATPS formed compared to 14 when formed by centrifugation (Table 3.3). Although this difference is not large it does indicate that there is more efficient polymer separation when left to settle for 30 minutes. When considering the scale up of the system to potentially hundreds of litres on an industrial scale, allowing the systems to settle would be advantageous compared to centrifugation as larger batches could be processed when settling allowing for faster processing and a potential reduction in costs due to not needing expensive large scale centrifugation equipment.

As the proteins are being produced by bacteria and purified away from the complete bacterial cell contents there will be other biopolymers present such as DNA as well as proteins.

It was appropriate to also test the partitioning of genomic material to ensure that it wouldn't contaminate the target therapeutic protein. The 20% w/v ATPS were then further tested to determine which systems demonstrated complete protein and genomic DNA partitioning to the lower dextran phase as this is a necessity for the further development of the ATPS as a generic affinity peptide purification system. It was found that the combination of higher molecular weight PEG and Dextran were the most successful when formed both by being left to settle and being centrifuged (Figure 3.3). This is thought to be due to there being a greater difference in the water density produced in each layer as the increased size of the polymer exhibits greater solvent ordering properties than their lower molecular weight counterparts (Zaslavsky, B. 1995). It has also been shown that both PEG and DEX are monopolar Lewis bases which means, when hydrogen bonding and Lifshitz-van der Waals interactions are considered, they will both repel each other strongly when immersed in water. This repulsion energy can be further increased with an increase in the molecular weight of the polymer explaining why the higher concentration polymer solutions as well as the higher molecular weight polymer solutions preferentially formed two distinct phases (Van Oss, C.J. *et al.,* 2011).

The partitioning of the genomic DNA was also analysed from the systems that were prepared using GFPuv in cell lysates and was seen to separate to the lower dextran layer in all formed ATPS when both formed by the centrifugation method and the settling method (Figure 3.4). This is due to the dextran layer having a more water rich and therefore polar environment for the DNA. It was also observed that there was a greater tendency for protein partitioning to the lower dextran phase in the presence of the cell lysate, whether the ATPS was formed by settling or centrifugation.

When examining the protein partitioning behaviours it was established that leaving the systems to settle was much more effective than centrifugation when full protein partitioning into the lower dextran layer was required. This could be because the act of leaving the layers to form naturally allowed the proteins to interact with the environment for a longer period of time and therefore partitioning into their optimal environment; compared to the centrifugation method which forced the separation of the polymers but tended to result in a spread of the

protein across the two phases. Such mixed partitioning is thought to be due to the surface charge of the proteins interacting with both of the polymers but possibly not having time to settle into the desired phase. This is especially thought to be the case as this was predominantly seen in systems that were used to separate purified GFPuv which has a polyhistidine tag (Figure 3.5). The presence of this tag causes the protein to be approximately 1 kDa larger than the wild type GFPuv and also contributes to the overall charge of the protein due to its imidazole groups and therefore may cause there to be a difference in the partitioning behaviours (Fan, W. 1997).

The ATPS that look most promising for future studies are the settled ATPS that partitioned both cell lysates and purified GFPuv to the lower dextran layer. There are four systems that will be used consisting of PEG 4000 / Dextran 40, PEG 20000 / Dextran 40, PEG 20000 / Dextran 100 and PEG 20000 / Dextran 200. These selected systems look the most promising, as it is necessary for all contents of the cell lysate to remain in the dextran layer (Figure 4.4), as well as utilising the higher molecular weight polymers as previously suggested by Van Oss, C.J. *et al.,* (1987). This will then allow the PEGylated GSH (of which PEGylation has previously been shown to influence the partitioning of a molecule to the PEG layer (Barbosa *et al.,* 2010)) to drag the GST tagged peptide into the upper PEG layer. PEG-GSH did however prove unsuccessful as an affinity partner within ATPS (Table 4.1) as it is thought the binding affinity between the receptor and ligand was not strong enough to overcome the interfacial tension present.

To overcome the interfacial tension, PEGylated NTA with a coordinated metal ion, in this case copper, was selected as research suggested a higher binding affinity between a coordinated metal ion and polyhistidine tagged proteins. Copper was selected after careful consideration as it has a lower risk level when consumed or injected compared to nickel which is commonly used in metal ion affinity protein purification. The European Medicines Agency (EMA) in 2007 recommended that copper, as a class two metal, demonstrates a low safety concern due to its requirement in dietary forms for a number of biological processes (Stanstead, HH., 1995). If the therapeutic protein is to be administered orally then the

tolerance for contaminating copper is 2500 µg/day compared to a tolerance level of 250 µg/day however the dosage required for parenterally administered therapeutic protein would be far lower. These experiments were performed initially with a Tris buffer, as previously performed by Barbosa, H. *et al.,* 2009, that proved unsuitable due to its ability to disassociate the metal ion from NTA at pH 7.4 (Table 5.1). A more suitable phosphate buffer was selected at pH 8.0, a pH where the interaction between metal ions and polyhistidine tags is reported to be strongest. Barbosa would not have encountered the same unsuitability due to using IDA which has a structure distinct to that of Tris and NTA.

Three Af-ATPS have been identified that successfully partition $His_6$-tagged GFPuv into the PEG phase via the affinity ligands 5 kDa M1 PEG-NTA-$Cu^{2+}$ and 10 kDa M1 PEG-NTA-$Cu^{2+}$, albeit with a relatively low yield (Figures 5.13 and 5.14). Of these, 18% PEG 20000 / 14% DEX 40 with 10 kDa M1 PEG-NTA-$Cu^{2+}$ (Figure 5.14) has the best potential for future use since purity is maintained throughout the entire PEG phase with a relatively high yield of protein. The protein yield and purity is of great importance when producing a therapeutic protein. As previously discussed, a protein administered as a drug must be free from impurities or within prescribed limits as described with free copper ions. If a high protein purity can be obtained, then little to no further processing would be required once collected in the dextran layer of the second shake. This is highly advantageous in a commercial environment as downstream processing costs would be vastly reduced with a reduction in processing steps.

Future work could investigate the effect of pH on this optimal system. As *E.*coli contains a number of native proteins that can bind to NTA resin (binding buffers are pH 8.0) it seems reasonable to assume that these proteins may also bind to the PEGylated NTA ligands used in the current study. It is thought that using a binding buffer at pH 8.0 deprotonates the external histidine residues present on native cellular proteins, which in turn leads to their co-purification via NTA ligands. Therefore it may be possible to reduce or eliminate the binding of these native proteins by slightly reducing the pH of the buffer systems used, such that $His_6$-tags remain deprotonated, whilst the individual residues on the surface of native proteins become protonated and so lose their affinity for the PEGylated ligands described herein. The reduction

of non-specific binding to the PEGylated NTA ligands would elevate commercial appeal of this purification method, particularly at laboratory bench scale. After consideration of the techniques used, a combination of the ability to filter and reuse any free copper ions as well as settling the systems instead of using centrifugation lend this method to being scaled up to commercial levels. Costs should also be considered, so fine-tuning the methodology to use lower concentrations of PEG and dextran would make the purification technique industrially appealing. The proposed mechanism of using a PEGylated protease to cleave the therapeutic protein directly from the histidine tag in order to be removed in the lower dextran layer of a secondary ATPS would add additional costs but the final product would be simply the target therapeutic protein, without any additional amino acids which can prove essential for FDA / EMA approval, located within a biocompatible polymer solution. A further consideration would be when working with temperature sensitive proteins. In this study, ATPS were difficult to repeatedly produce at 4 °C due to the increased viscosity of the polymers and therefore currently recommended for therapeutic proteins that are stable at room temperature.

Overall this thesis has provided evidence that ATPS purification of therapeutic proteins is possible when using an 18% PEG 20000 / 14% DEX 40 system containing 10 kDa M1 PEG-NTA-$Cu^{2+}$. This system allowed for a relatively high yield and purity of the desired therapeutic protein into the upper PEG phase when allowed to settle at room temperature. The practicalities of this methodology have been explored and discussed, particularly with regard to technical achievability, yield, cost effectiveness and purity. It is suggested that for therapeutic proteins to be delivered parenterally, the yield can be sacrificed slightly in order to obtain a higher purity. Conversely, for orally administered drugs a medium purity, larger scale process could be employed.

# Chapter 7. References


Aatsinki, J.T. & Rajaniemi, H.J., 2005. An alternative use of basic pGEX vectors for producing both N- and C-terminal fusion proteins for production and affinity purification of antibodies. *Protein Expression and Purification*, 40(2), pp.287-291.

Aguilar, M.-I., 2004. *HPLC of peptides and proteins: methods and protocols*, Totowa, N.J.: Humana Press.

Albertsson, P.A., Johansson, G. & Tjerneld, F., 1990. Separation processes in biotechnology. Aqueous two-phase separations. *Bioprocess Technol*, 9, pp.287-327.

Albertsson, P.A.A., 1971. *Partition of cell particles and macromolecules: distribution and fractionation of cells, mitochondria, chloroplasts, viruses, proteins, nucleic acids, and antigen-antibody complexes in aqueous polymer two-phase systems*, New York ; [Chichester]: Wiley-Interscience.

Alconel, S *et al*., 2011. FDA-approved poly(ethylene glycol) - protein conjugate drugs. *Poly Chem*. 2, pp. 1442 - 1448

Alpar, Emin Kaya *et al.,* 2004. Effects of hypertonic dextran in hypovolaemic shock: a prospective clinical trial. *Injury*, 35(5), pp.500-506

Arnau, J. *et al.,* 2006. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification*, 48(1), pp.1-13.

Asenjo, J.A. & Andrews, B.A., 2011. Aqueous two-phase systems for protein separation: a perspective. *J Chromatogr A*. 1218(49), pp. 8826-35.

Azevedo, A.M. *et al.,* 2009. Affinity-enhanced purification of human antibodies by aqueous two-phase extraction. *Separation and Purification Technology*, 65(1), pp.31-39.

Barbosa, H. *et al.,* 2008. Affinity partitioning of plasmid DNA with a zinc finger protein. *Journal of Chromatography A*, 1206(2), pp.105-112.
Barbosa, H. *et al.,* 2009. Protein quantification in the presence of poly(ethylene glycol) and dextran using the Bradford method. *Analytical Biochemistry*, 395, pp.108-110.

Barbosa, H.S.C. *et al.,* 2010. Dual affinity method for plasmid DNA purification in aqueous two-phase systems. *Journal of Chromatography A*, 1217(9), pp.1429-1436.

Bradford M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry,* 72. pp.248–254.

Brocchini, S. *et al.,* 2006. PEGylation of native disulfide bonds in proteins. *Nature Protocols*, 1(5), pp.2241-2252.

Brocchini, S. *et al.,* 2007. Disulfide bridge based PEGylation of proteins. *Advanced Drug Delivery Reviews*, 60, pp.3-12.

Bulger EM. 7.5% Saline and 7.5% Saline/6% Dextran for Hypovolemic Shock. *The Journal of trauma*. 2011;70(5):S27-S29.

Casey, J.L. *et al.,* 1995. Purification of bacterially expressed single chain Fv antibodies for clinical applications using metal chelate chromatography. *Journal of Immunological Methods*, 179(1), pp.105-116.

Chaga, G.S., 2001. Twenty-five years of immobilized metal ion affinity chromatography: past, present and future. *Journal of Biochemical and Biophysical Methods*, 49(1-3), pp.313-334.

Chakravarthi, S., Jessop, C.E. & Bulleid, N.J., 2006. The role of glutathione in disulphide bond formation and endoplasmic-reticulum-generated oxidative stress. *EMBO Rep*, 7(3), pp.271-275.

Chen, H.-M. *et al.,* 1999. Affinity purification of Schistosoma japonicum glutathione-S-transferase and its site-directed mutants with glutathione affinity chromatography and immobilized metal affinity chromatography. *Journal of Chromatography A*, 852(1), pp.151-159.

Edahiro, J. *et al.,* 2005. Separation of cultured strawberry cells producing anthocyanins in aqueous two-phase system. *J Biosci Bioeng*, 100(4), pp.449-54.

Effio, C. *et al.*, 2015. Downstream processing of virus-like particles: Single-stage and multi-stage aqueous two-phase extraction. *JOURNAL OF CHROMATOGRAPHY A,* 1383(1), pp.35-46.

Fan, Weiyu, "Charge effect on protein partitioning in aqueous two-phase systems " (1997). Retrospective eses and Dissertations. Paper 11980.

Fong, R.B. *et al.,* 2002. Affinity separation using an Fv antibody fragment-“smart” polymer conjugate. *Biotechnology and Bioengineering*, 79(3), pp.271-276.

Friman S, Egestad B, Sjovall J, and Svanvik J (1993) Hepatic excretion and metabolism of polyethylene glycols and mannitol in the cat. *J Hepatol* **17:** 48–55.

Gaberc-Porekar, V. & Menart, V., 2001. Perspectives of immobilized-metal affinity chromatography. *Journal of Biochemical and Biophysical Methods*, 49(1-3), pp.335-360.

Garza-Madrid, M. *et al.,* 2010. Potential of Aqueous Two-Phase Systems constructed on flexible devices: Human serum albumin as proof of concept. *Process Biochemistry*, 45(7), pp.1082-1087.

Gottschalk, U., 2008. Bioseparation in antibody manufacturing: The good, the bad and the ugly. *Biotechnology Progress*, 24(3), pp.496-503.

Goubran Botros, H. *et al.,* 1991. Immobilized metal ion affinity partitioning of cells in aqueous two-phase systems: erythrocytes as a model. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1074(1), pp.69-73.

Gupta, R., Bradoo, S. & Saxena, R. K., 1999. Aqueous two phase systems: An attractive technology for downstream processing of biomolecules. *Current Science*, 77(4), pp.520-523.

Hagel, L, 1998. Gel-Filtration Chromatography. *Current Protocols in Molecular Biology,* J.E., *et al.*(eds). New York: John Wiley and Sons, pp. *8.3.1–8.3.30*.

Harper, S. & Speicher, D.W., 2011. Purification of proteins fused to glutathione S-tranferase. *Methods Mol Biol,* 681, pp.259-280.

Herold DA, Keil K, and Bruns DE (1989) Oxidation of polyethylene glycols by alcohol dehydrogenase. *Biochem Pharmacol* **38:** 73–76

Israelachvili, J., 1997. The different faces of poly(ethylene glycol). *Proc. Natl. Acad. Sci. USA*, 94, pp. 8378-8379.

Kapust, R.B. & Waugh, D.S., 1999. Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Science*, 8(8), pp.1668-1674.

Kasprzak, K.S., Sunderman, F.W. & Salnikow, K., 2003. Nickel carcinogenesis. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, 533(1-2), pp.67-97.

Kaul, R.H. 2011. Aqueous two-phase systems: Methods and Protocols. *Methods in Biotechnology*, 11, pp. 1-10.

Kepka, C. *et al.,* 2003. Pilot-scale extraction of an intracellular recombinant cutinase from E. coli cell homogenate using a thermoseparating aqueous two-phase system. *Journal of Biotechnology*, 103(2), pp.165-181.

Korf, U., Kohl, T., van der Zandt, H., Zahn, R., Schleeger, S., Ueberle, B., Wandschneider, S., Bechtel, S., Schnölzer, M., Ottleben, H., Wiemann, S. and Poustka, A. (2005), Large-scale protein expression for proteome research. Proteomics, 5: 3571–3580.

Kumar, A. *et al.*, 2001. Type-specific separation of animal cells in aqueous two-phase systems using antibody conjugates with temperature-sensitive polymers. *Biotechnology and Bioengineering*, 75(5), pp.570-580.

Lakowicz, J.R., 2006. *Principles of Fluorescence Spectroscopy 3rd Edition*. New York ; Springer US.

Li, Y & Cao, X., 2010. Prediction of phase diagrams for new pH-thermo sensitive recycling aqueous two-phase systems. *Fluid phase equilibria* 298, pp.206-211

Liu, H.F. *et al.,* 2010. Recovery and purification process development for monoclonal antibody production. *MAbs*, 2(5), pp.480-499.

Lowe, C.R., 2001. Combinatorial approaches to affinity chromatography. *Current Opinion in Chemical Biology,* 5(3), pp.248-256.

Mazzola, P.G. *et al.*, 2006. Affinity-tagged green fluorescent protein (GFPUV) extraction from a clarified E. coli cell lysate using a two-phase aqueous micellar system. *Biotechnol Bioeng*, 93(5), pp.998-1004.

Mittal, K.L, 1983. Physicochemical Aspects of Polymer Surfaces, Volume 1. Plenum Press, NewYork.

Münchow, G., Schönfeld, F., Hardt, S. and Graf, K., 2008. Protein diffusion across the interface in aqueous two-phase systems. *Langmuir*, *24*(16), pp.8547-8553.

Nguyen, N.Y. *et al.,* 1996. Isolation of a biologically active soluble human interferon-alpha receptor-GST fusion protein expressed in Escherichia coli. *Journal of Interferon & Cytokine Research: The Official Journal of the International Society for Interferon and Cytokine Research*, 16(10), pp.835-844.

Nilsson, J. *et al.,* 1997. Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expression and Purification*, 11(1), pp.1-16.

Porath, J. 1992. Immobilized metal ion affinity chromatography. *Protein Expression and Purification*, 3(4), pp. 263-281.

Rogl, H. *et al.,* 1998. Refolding of Escherichia coli produced membrane protein inclusion bodies immobilised by nickel chelating chromatography. *FEBS Letters*, 432(1-2), pp.21-26.

Rosa, P.A. *et al.,* 2007. Affinity partitioning of human antibodies in aqueous two-phase systems. *Journal of Chromatography A*, 1162(1), pp.103-13.

Rosa, P.A.J. *et al.,* 2010. Aqueous two-phase systems: A viable platform in the manufacturing of biopharmaceuticals. *Journal of Chromatography A*, 1217(16), pp.2296-2305.

Rosa, P.A.J. *et al.,* 2011. Aqueous two-phase extraction as a platform in the biomanufacturing industry: Economical and environmental sustainability. *Biotechnology Advances*, 29(6), pp.559-567

Ruiz-Ruiz, F. *et al.,* 2013. Scaling-up of a B-phycoerythrin production and purification bioprocess involving aqueous two-phase systems: Practical experiences. *Process Biochemistry*, 48(4), pp.738-745

Ryden, J. & Albertsson, P.A., 1971. Interfacial tension of dextran - polyethylene glycol - water two-phase systems. *Journal of Colloid and Interface Science*, 37(1), pp.219-222.

Saraswat, M. *et al.*, 2013. Preparative purification of recombinant proteins: Current status and future trends. *BioMed Research International*, 2013(2013), Article ID 312709, 18 pages.

Schaffer, C. B., Critchfield, F. H. & Nair, J. H. (1950) J. Amer.Pharm. Assoc., Sci. Ed., 39, 340-344

Selber, K. *et al.,* 2004. Large-scale separation and production of engineered proteins, designed for facilitated recovery in detergent-based aqueous two-phase extraction systems. *Process Biochemistry*, 39(7), pp.889-896.

Shuman, H.A., 1982. Active transport of maltose in Escherichia coli K12. Role of the periplasmic maltose-binding protein and evidence for a substrate recognition site in the cytoplasmic membrane. *Journal of Biological Chemistry*, 257(10), pp.5455 -5461.

Simpson, D.M. & Beynon, R.J., 2010. Acetone Precipitation of Proteins and the Modification of Peptides. *Journal of Proteome Research*, 9(1), pp.444-450.

Smith, C., 2005. Striving for purity: advances in protein purification. *Nature Methods*, 2(1), pp.71-77.

Smith D.B. & Johnson K.S., 1988. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione *S*-transferase. *Gene,* 67(1), pp.31-40.

Srinivasan, U. & Bell, J.A., 1998. A convenient method for affinity purification of maltose binding protein fusions. *Journal of Biotechnology*, 62(3), pp.163-167.

Stanstead HH (1995). "Requirements and toxicity of essential trace elements, illustrated by zinc and copper". *Am. J. Clin. Nutr.* **61** (3 Suppl): 621S–624S.

Tjerneld, F. *et al.,* 1986. New Aqueous 2-Phase System Based on Hydroxypropyl Starch Useful in Enzyme-Purification. *Enzyme and Microbial Technology*, 8(7), pp.417-423.

Trexler-Schmidt, M. *et al.,* 2009. Purification Strategies to Process 5 g/L Titers of Monoclonal Antibodies. *Biopharm International*, pp.8-15.

Uter, W., Fuchs, Th., Häusser, M. and Ippen, H. (1995), Patch test results with serial dilutions of nickel sulfate (with and without detergent), palladium chloride, and nickel and palladium metal plates. Contact Dermatitis, 32: 135–142.

Van Oss, C.J. *et al.,* 1987. The mechanism of Partition in Aqeuous Media. *Separation Science and Technology*, 22(6), pp.1515-1526.

Waites, M.J., 2001. *Industrial microbiology*, Oxford: Blackwell Science.

Walker, S.G., Dale, C.J. & Lyddiatt, A., 1996. Aqueous two-phase partition of complex protein feedstocks derived from brain tissue homogenates. *Journal of Chromatography B: Biomedical Sciences and Applications*, 680(1-2), pp.91-96.

Webster R. *et al*., 2007. PEGylated proteins: evaluation of their safety in the absence of definitive metabolism studies. *Drug Metab Dispos*. 35(1):9-16.

Xu, Y. *et al.,* 2003. Purification of glucose-6-phosphate dehydrogenase from baker's yeast in aqueous two-phase systems with free triazine dyes as affinity ligands. *Appl Biochem Biotechnol*, 105 -108, pp.853-65

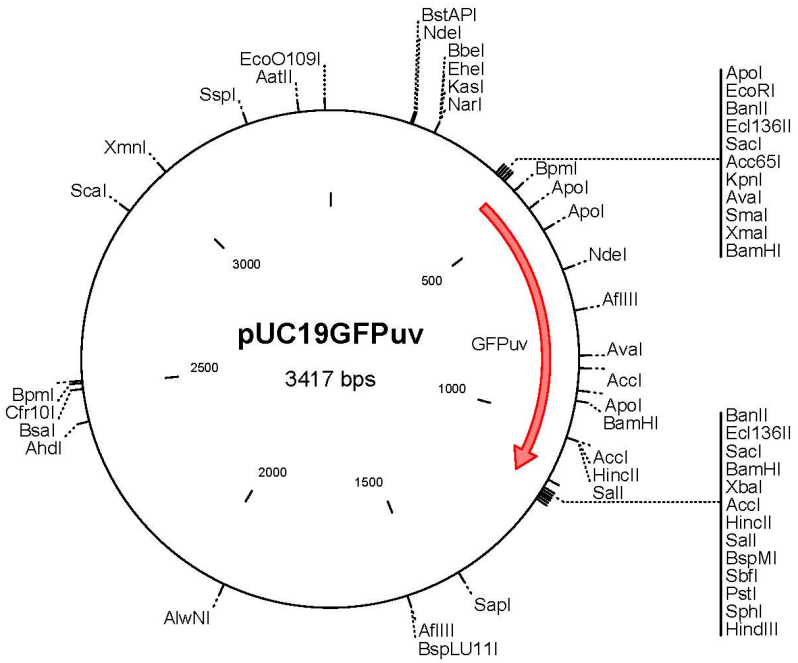
Yang BB, Lum PK, Hayashi MM, and Roskos LK (2004) Polyethylene glycol modification of filgrastim results in decreased renal clearance of the protein in rats. *J Pharm Sci* **93:** 1367–1373.

Zachariou, M. & Bailon, P., 2007. Affinity chromatography : methods and protocols. *Methods in Molecular Biology 2nd Ed*, London: Springer

Zaslavsky, Boris (1995). *Aqueous Two-Phase Partitioning: Physical Chemistry and Bioanalytical Applications*. Marcel Dekker Inc.

# Chapter 8. Appendices

## Appendix 8.1




```
Molecule:       pUC19GFPuv,  3417 bps DNA Circular
File Name:      pUC19GFPuv.cm5,  dated 02 Jul 2012

Description:

Notes:
```

Molecule: pUC19GFPuv, 3417 bps DNA Circular
Description:
File Name: pUC19GFPuv.cm5, dated 02 Jul 2012
Printed: 1-3417 bps (Full), format Annotated: Enzymes, Genes

```
      1  tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca

     61  cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg

    121  ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc

                                                                  NarI
                                                                  KasI
                                                                  EheI
                NdeI                                              BbeI
              BstAPI
    181  accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc

    241  attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat

    301  tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt

                                                  SacI                XmaI
                                          EcoRI   Ecl136II   KpnI      SmaI
                                           ApoI    BanII    Acc65I     AvaI
    361  tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgagctcggt acccccgggg

           BamHI                                  BpmI
    421  ggatccatga gtaaaggaga agaacttttc actggagttg tcccaattct tgttgaatta
         »»»...........................................GFPuv.....................»

                    ApoI
    481  gatggtgatg ttaatgggca caaattttct gtcagtggag agggtgaagg tgatgcaaca
         »...........................................GFPuv.....................»

                ApoI
    541  tacggaaaac ttacccttaa atttatttgc actactggaa aactacctgt tccatggcca
         »...........................................GFPuv.....................»

                                                                  NdeI
    601  acacttgtca ctactttctc ttatggtgtt caatgctttt cccgttatcc ggatcatatg
         »...........................................GFPuv.....................»

    661  aaacggcatg acttttcaa gagtgccatg cccgaaggtt atgtacagga acgcactata
         »...........................................GFPuv.....................»

                                  AflIII
    721  tctttcaaag atgacgggaa ctacaagacg cgtgctgaag tcaagtttga aggtgatacc
         »...........................................GFPuv.....................»

    781  cttgttaatc gtatcgagtt aaaaggtatt gattttaaag aagatggaaa cattctcgga
         »...........................................GFPuv.....................»

           AvaI                                AccI
    841  cacaaactcg agtacaacta taactcacac aatgtataca tcacggcaga caaacaaaag
         »...........................................GFPuv.....................»

                        ApoI                                BamHI
    901  aatggaatca aagctaactt caaaattcgc cacaacattg aagatggatc cgttcaacta
         »...........................................GFPuv.....................»

    961  gcagaccatt atcaacaaaa tactccaatt ggcgatggcc ctgtcctttt accagacaac
         »...........................................GFPuv.....................»
```


SalI
HincII
AccI
1021 cattacctgt cgacacaatc tgccctttcg aaagatccca acgaaaagcg tgaccacatg
GFPuv
SacI
Ecl136II
BanII
1081 gtccttcttg agtttgtaac tgctgctggg attacacatg gcatggatga gctctacaaa
GFPuv
SbfI
BspMI
SalI
HincII
BamHI XbaI AccI PstI SphI HindIII
1141 taatgggga tcctctagag tcgacctgca ggcatgcaag cttggcgtaa tcatggtcat
1201 agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa
1261 gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc
1321 gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc
SapI
1381 aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact
1441 cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac
BspLU11I
AflIII
1501 ggttatccac agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa
1561 aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg
1621 acgagcatca caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa
1681 gataccaggc gtttcccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc
1741 ttaccggata cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac
1801 gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac
1861 ccccgttca gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg
AlwNI
1921 taagacacga cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt
1981 atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagaa
2041 cagtatttgg tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct
2101 cttgatccgg caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga
2161 ttacgcgcag aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg
2221 ctcagtggaa cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct

2281 tcacctagat ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt

2341 aaacttggtc tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc

AhdI

2401 tatttcgttc atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg

BpmI

BsaI Cfr10I

2461 gcttaccatc tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag

2521 atttatcagc aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt

2581 tatccgcctc catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag

2641 ttaatagttt gcgcaacgtt gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt

2701 ttggtatggc ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca

2761 tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg

2821 ccgcagtgtt atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat

ScaI

2881 ccgtaagatg cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta

2941 tgcggcgacc gagttgctct tgcccggcgt caatacggga taataccgcg ccacatagca

XmnI

3001 gaactttaaa agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct

3061 taccgctgtt gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat

3121 cttttacttt caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa

SspI

3181 agggaataag ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt

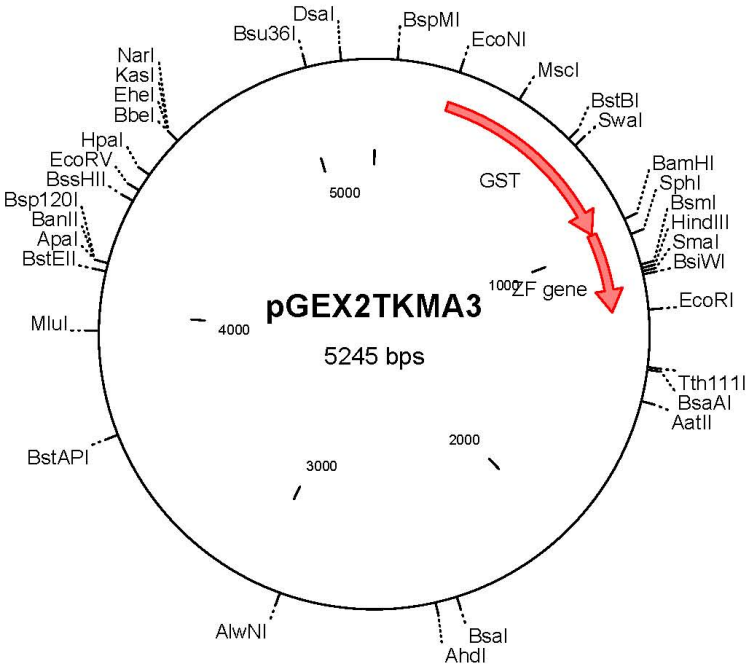
3241 gaagcattta tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa

AatII

3301 ataaacaaat aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa

3361 ccattattat catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtc

## Appendix 8.2




Molecule: pGEX2TKMA3, 5245 bps DNA Circular
File Name: pGEX2TKMA3.cm5, dated 25 Jul 2011

Description:

Notes:

Molecule: pGEX2TKMA3, 5245 bps DNA Circular
Description:
File Name: pGEX2TKMA3.cm5, dated 25 Jul 2011
Printed: 1-5245 bps (Full), format Annotated: Enzymes, Genes

```
      1  acgttatcga ctgcacggtg caccaatgct tctggcgtca ggcagccatc ggaagctgtg

                    BspMI
     61  gtatggctgt gcaggtcgta aatcactgca taattcgtgt cgctcaaggc gcactcccgt

    121  tctggataat gtttttgcg ccgacatcat aacggttctg gcaaatattc tgaaatgagc

    181  tgttgacaat taatcatcgg ctcgtataat gtgtggaatt gtgagcggat aacaatttca

                                  EcoNI
    241  cacaggaaac agtattcatg tcccctatac taggttattg gaaaattaag ggccttgtgc
                          »»».....GST.....»

    301  aacccactcg acttcttttg gaatatcttg aagaaaaata tgaagagcat ttgtatgagc
         »....GST.....»

    361  gcgatgaagg tgataaatgg cgaaacaaaa agtttgaatt gggtttggag tttcccaatc
         »....GST.....»

                                                MscI
    421  ttccttatta tattgatggt gatgttaaat taacacagtc tatggccatc atacgttata
         »....GST.....»

    481  tagctgacaa gcacaacatg ttgggtggtt gtccaaaaga gcgtgcagag atttcaatgc
         »....GST.....»

    541  ttgaaggagc ggttttggat attagatacg gtgtttcgag aattgcatat agtaaagact
         »....GST.....»

                                                                  BstBI
    601  ttgaaactct caaagttgat tttcttagca agctacctga aatgctgaaa atgttcgaag
         »....GST.....»

                          SwaI
    661  atcgtttatg tcataaaaca tatttaaatg gtgatcatgt aacccatcct gacttcatgt
         »....GST.....»

    721  tgtatgacgc tcttgatgtt gttttataca tggacccaat gtgcctggat gcgttcccaa
         »....GST.....»

    781  aattagtttg ttttaaaaaa cgtattgaag ctatcccaca aattgataag tacttgaaat
         »....GST.....»

    841  ccagcaagta tatagcatgg cctttgcagg gctggcaagc cacgtttggt ggtggcgacc
         »....GST.....»

                                                            BamHI
    901  atcctccaaa atcggatctg gttccgcgtg gatctcgtcg tgcatctgtt ggatccgaga
         »....GST.....»»»
                                                                »»»....»

                                                      SphI
    961  aacttcgtaa tggttcgggc gacccaggaa agaagaaaca gcatgcgtgc ccagagtgtg
         »....ZF gene.....»

   1021  gtaagagctt cagtcaatcc tctgatctgc agcgccacca acgtacacat accggggaga
         »....ZF gene.....»
```

BsmI HindIII SmaI BsiWI

1081 aaccgtacaa gtgtccagaa tgcgggaaaa gcttcgttcc cgggatgacg tacgcacacc

»....ZF gene....»

1141 ggggaaaagc catataaatg ccctgaatgt ggcaagtctt tcagccgtag tgatcatctg

»....ZF gene....»

EcoRI

1201 tctcgccatc aacgcacgca tcagaacaag aaatgagaat tcatcgtgac tgactgacga

»....ZF gene....»»

1261 tctgcctcgc gcgtttcggt gatgacggtg aaaacctctg acacatgcag ctcccggaga

1321 cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag ggcgcgtcag

BsaAI

Tth111I

1381 cgggtgttgg cgggtgtcgg ggcgcagcca tgacccagtc acgtagcgat agcggagtgt

1441 ataattcttg aagacgaaag ggcctcgtga tacgcctatt tttataggtt aatgtcatga

AatII

1501 taataatggt ttcttagacg tcaggtggca cttttcgggg aaatgtgcgc ggaaccccta

1561 tttgtttatt tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat

1621 aaatgcttca ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc

1681 ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcacccagaa acgctggtga

1741 aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca

1801 acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg atgagcactt

1861 ttaaagttct gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg

1921 gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc

1981 atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata

2041 acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt

2101 tgcacaacat gggggatcat gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag

2161 ccataccaaa cgacgagcgt gacaccacga tgcctgcagc aatggcaaca acgttgcgca

2221 aactattaac tggcgaacta cttactctag cttcccggca acaattaata gactggatgg

2281 aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg

BsaI

2341 ctgataaatc tggagccggt gagcgtgggt ctcgcggtat cattgcagca ctggggccag

AhdI

2401 atggtaagcc ctcccgtatc gtagttatct acacgacggg gagtcaggca actatggatg

2461 aacgaaatag acagatcgct gagataggtg cctcactgat taagcattgg taactgtcag

2521 accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa tttaaaagga

2581 tctaggtgaa gatccttttt gataatctca tgaccaaaat cccttaacgt gagttttcgt

2641 tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat cctttttttc

2701 tgcgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg gtttgtttgc

2761 cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga gcgcagatac

2821 caaatactgt ccttctagtg tagccgtagt taggccacca cttcaagaac tctgtagcac

AlwNI

2881 cgcctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt

2941 cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgcag cggtcgggct

3001 gaacgggggg ttcgtgcaca gcccagct tggagcgaac gacctacacc gaactgagat

3061 acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag gcggacaggt

3121 atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg

3181 cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt

3241 gatgctcgtc agggggggcgg agcctatgga aaaacgccag caacgcggcc tttttacggt

3301 tcctggcctt ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg

3361 tggataaccg tattaccgcc tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg

3421 agcgcagcga gtcagtgagc gaggaagcgg aagagcgcct gatgcggtat tttctcctta

3481 cgcatctgtg cggtatttca caccgcataa attccgacac catcgaatgg tgcaaaacct

3541 ttcgcggtat ggcatgatag cgcccggaag agagtcaatt cagggtggtg aatgtgaaac

BstAPI

3601 cagtaacgtt atacgatgtc gcagagtatg ccggtgtctc ttatcagacc gtttcccgcg

3661 tggtgaacca ggccagccac gtttctgcga aaacgcggga aaaagtggaa gcggcgatgg

3721 cggagctgaa ttacattccc aaccgcgtgg cacaacaact ggcgggcaaa cagtcgttgc

3781 tgattggcgt tgccacctcc agtctggccc tgcacgcgcc gtcgcaaatt gtcgcggcga

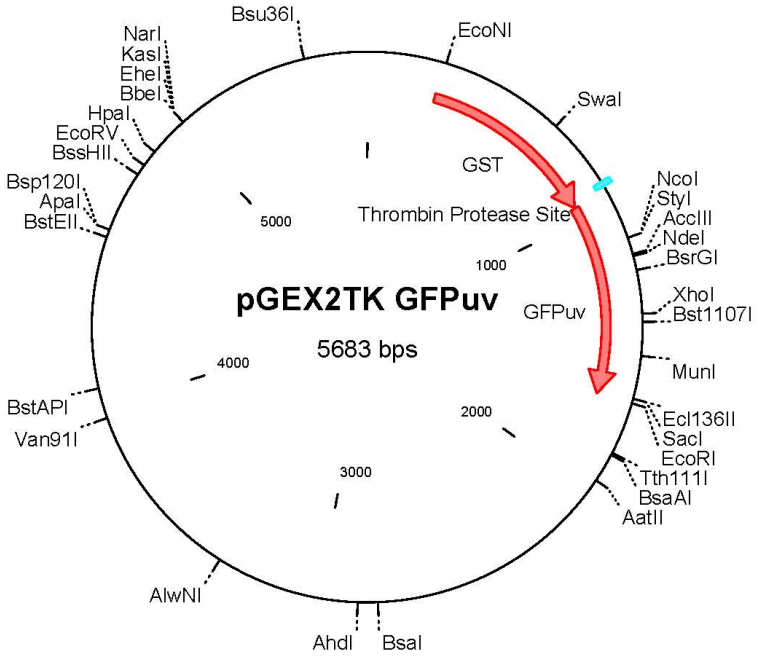
3841 ttaaatctcg cgccgatcaa ctgggtgcca gcgtggtggt gtcgatggta gaacgaagcg

MluI

3901 gcgtcgaagc ctgtaaagcg gcggtgcaca atcttctcgc gcaacgcgtc agtgggctga

```
3961   tcattaacta tccgctggat gaccaggatg ccattgctgt ggaagctgcc tgcactaatg

4021   ttccggcgtt atttcttgat gtctctgacc agacacccat caacagtatt attttctccc

                                                          BstEII
4081   atgaagacgg tacgcgactg ggcgtggagc atctggtcgc attgggtcac cagcaaatcg

                  Bsp120I
                  BanII
                  ApaI
4141   cgctgttagc gggcccatta agttctgtct cggcgcgtct gcgtctggct ggctggcata

4201   aatatctcac tcgcaatcaa attcagccga tagcggaacg ggaaggcgac tggagtgcca

4261   tgtccggttt tcaacaaacc atgcaaatgc tgaatgaggg catcgttccc actgcgatgc

                                                   BssHII
4321   tggttgccaa cgatcagatg gcgctgggcg caatgcgcgc cattaccgag tccgggctgc

                  EcoRV
4381   gcgttggtgc ggatatctcg gtagtgggat acgacgatac cgaagacagc tcatgttata

             HpaI
4441   tcccgccgtt aaccaccatc aaacaggatt ttcgcctgct ggggcaaacc agcgtggacc

4501   gcttgctgca actctctcag ggccaggcgg tgaagggcaa tcagctgttg cccgtctcac

                                 NarI
                                 KasI
                                 EheI
                                 BbeI
4561   tggtgaaaag aaaaaccacc ctggcgccca atacgcaaac cgcctctccc cgcgcgttgg

4621   ccgattcatt aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc

4681   aacgcaatta atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt

4741   ccggctcgta tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat

4801   gaccatgatt acggattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg

4861   cgttacccaa cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga

4921   agaggcccgc accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatggcgctt

                                                          Bsu36I
4981   tgcctggttt ccggcaccag aagcggtgcc ggaaagctgg ctggagtgcg atcttcctga

5041   ggccgatact gtcgtcgtcc cctcaaactg gcagatgcac ggttacgatg cgcccatcta

                                                   DsaI
5101   caccaacgta acctatccca ttacggtcaa tccgccgttt gttcccacgg agaatccgac

5161   gggttgttac tcgctcacat ttaatgttga tgaaagctgg ctacaggaag gccagacgcg

5221   aattattttt gatggcgttg gaatt
```

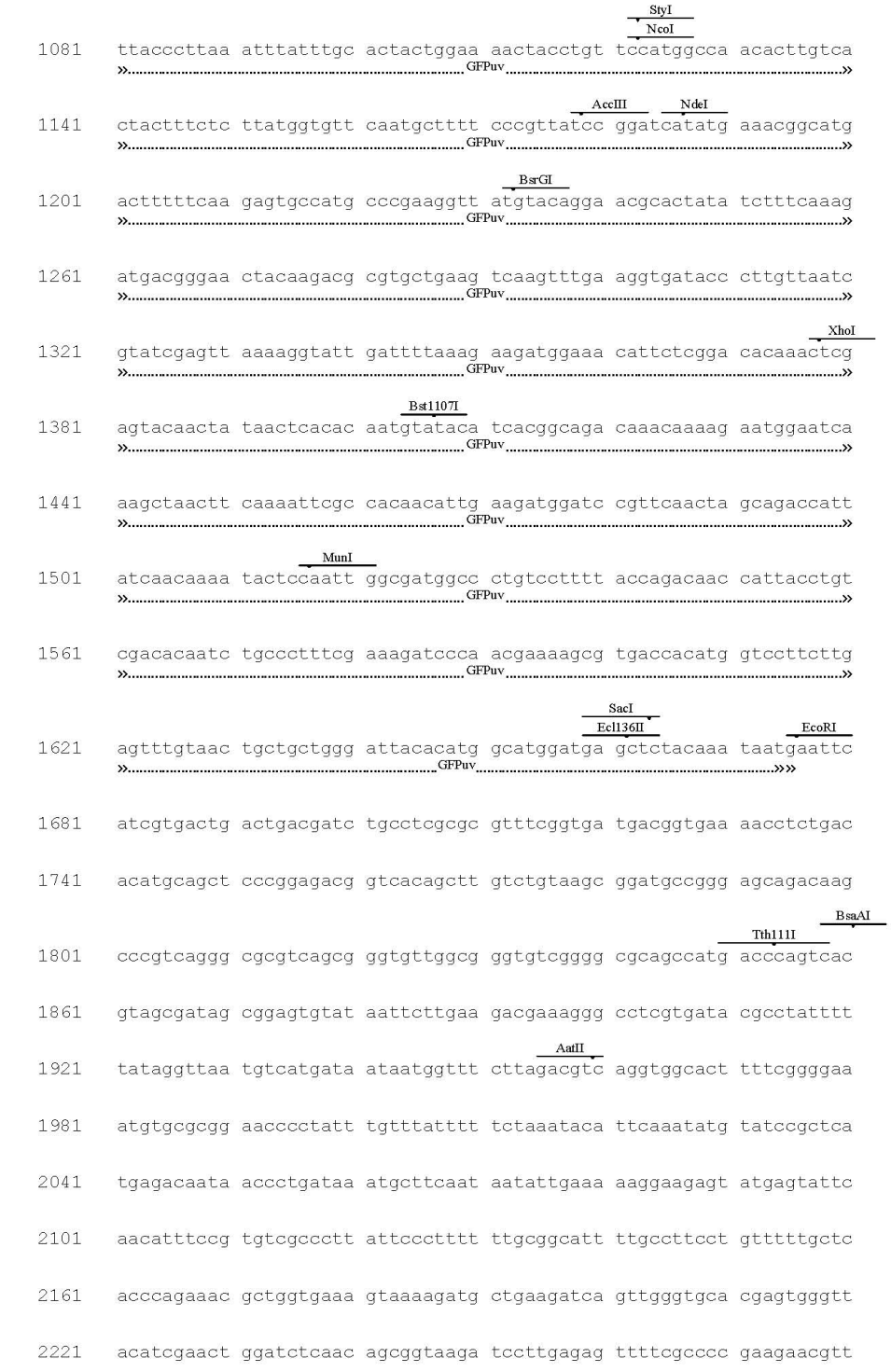
## Appendix 8.3




```
Molecule:       pGEX2TK GFPuv,  5683 bps DNA Circular
File Name:      pGEX2TK GFPuv.cm5,  dated 02 Jul 2012

Description:

Notes:
```

Molecule: pGEX2TK GFPuv, 5683 bps DNA Circular
Description:
File Name: pGEX2TK GFPuv.cm5, dated 02 Jul 2012
Printed: 1-5683 bps (Full), format Annotated: Enzymes, Genes

```
      1   acgttatcga ctgcacggtg caccaatgct tctggcgtca ggcagccatc ggaagctgtg

     61   gtatggctgt gcaggtcgta aatcactgca taattcgtgt cgctcaaggc gcactcccgt

    121   tctggataat gtttttgcg ccgacatcat aacggttctg gcaaatattc tgaaatgagc

    181   tgttgacaat taatcatcgg ctcgtataat gtgtggaatt gtgagcggat aacaatttca

                                          EcoNI
    241   cacaggaaac agtattcatg tcccctatac taggttattg gaaaattaag ggccttgtgc
                           »»».....................GST.....................»

    301   aacccactcg acttcttttg gaatatcttg aagaaaaata tgaagagcat ttgtatgagc
          ».....................GST.....................»

    361   gcgatgaagg tgataaatgg cgaaacaaaa agtttgaatt gggtttggag tttcccaatc
          ».....................GST.....................»

    421   ttccttatta tattgatggt gatgttaaat taacacagtc tatggccatc atacgttata
          ».....................GST.....................»

    481   tagctgacaa gcacaacatg ttgggtggtt gtccaaaaga gcgtgcagag atttcaatgc
          ».....................GST.....................»

    541   ttgaaggagc ggttttggat attagatacg gtgtttcgag aattgcatat agtaaagact
          ».....................GST.....................»

    601   ttgaaactct caaagttgat tttcttagca agctacctga aatgctgaaa atgttcgaag
          ».....................GST.....................»

                                    SwaI
    661   atcgtttatg tcataaaaca tatttaaatg gtgatcatgt aacccatcct gacttcatgt
          ».....................GST.....................»

    721   tgtatgacgc tcttgatgtt gttttataca tggacccaat gtgcctggat gcgttcccaa
          ».....................GST.....................»

    781   aattagtttg ttttaaaaaa cgtattgaag ctatcccaca aattgataag tacttgaaat
          ».....................GST.....................»

    841   ccagcaagta tatagcatgg cctttgcagg gctggcaagc cacgtttggt ggtggcgacc
          ».....................GST.....................»

    901   atcctccaaa atcggatctg gttccgcgtg gatctcgtcg tgcatctgtt ggatccatga
          ».....................GST.....................»»»
                                                          »»»....GFPuv.....»

    961   gtaaaggaga agaacttttc actggagttg tcccaattct tgttgaatta gatggtgatg
          ».....................GFPuv.....................»

   1021   ttaatgggca caaattttct gtcagtggag agggtgaagg tgatgcaaca tacggaaaac
          ».....................GFPuv.....................»
```

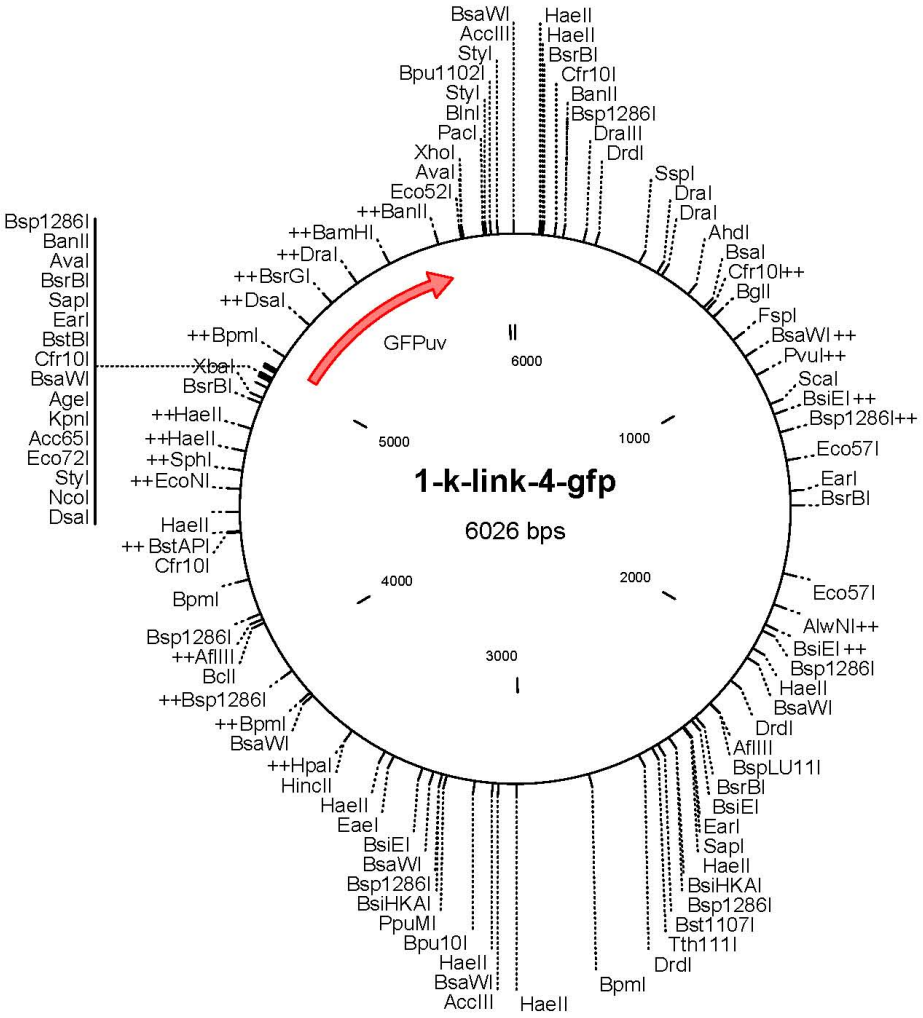

StyI
NcoI
1081 ttacccttaa atttatttgc actactggaa aactacctgt tccatggcca acacttgtca
GFPuv
AccIII
NdeI
1141 ctactttctc ttatggtgtt caatgctttt cccgttatcc ggatcatatg aaacggcatg
GFPuv
BsrGI
1201 actttttcaa gagtgccatg cccgaaggtt atgtacagga acgcactata tctttcaaag
GFPuv
1261 atgacgggaa ctacaagacg cgtgctgaag tcaagtttga aggtgatacc cttgttaatc
GFPuv
XhoI
1321 gtatcgagtt aaaaggtatt gattttaaag aagatggaaa cattctcgga cacaaactcg
GFPuv
Bst1107I
1381 agtacaacta taactcacac aatgtataca tcacggcaga caaacaaaag aatggaatca
GFPuv
1441 aagctaactt caaaattcgc cacaacattg aagatggatc cgttcaacta gcagaccatt
GFPuv
MunI
1501 atcaacaaaa tactccaatt ggcgatggcc ctgtcctttt accagacaac cattacctgt
GFPuv
1561 cgacacaatc tgccctttcg aaagatccca acgaaaagcg tgaccacatg gtccttcttg
GFPuv
SacI
Ecl136II
EcoRI
1621 agtttgtaac tgctgctggg attacacatg gcatggatga gctctacaaa taatgaattc
GFPuv
1681 atcgtgactg actgacgatc tgcctcgcgc gtttcggtga tgacggtgaa aacctctgac
1741 acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag
BsaAI
Tth111I
1801 cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg cgcagccatg acccagtcac
1861 gtagcgatag cggagtgtat aattcttgaa gacgaaaggg cctcgtgata cgcctatttt
AatII
1921 tataggttaa tgtcatgata ataatggttt cttagacgtc aggtggcact tttcggggaa
1981 atgtgcgcgg aacccctatt tgtttatttt tctaaatata ttcaaatatg tatccgctca
2041 tgagacaata accctgataa atgcttcaat aatattgaaa aaggaagagt atgagtattc
2101 aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttcct gtttttgctc
2161 acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt
2221 acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc gaagaacgtt

```
2281   ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc cgtgttgacg

2341   ccgggcaaga gcaactcggt cgccgcatac actattctca gaatgacttg gttgagtact

2401   caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta tgcagtgctg

2461   ccataaccat gagtgataac actgcggcca acttacttct gacaacgatc ggaggaccga

2521   aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt gatcgttggg

2581   aaccggagct gaatgaagcc ataccaaacg acgagcgtga caccacgatg cctgcagcaa

2641   tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct tcccggcaac

2701   aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc

                                          BsaI
2761   cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca

                                                  AhdI
2821   ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga

2881   gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc tcactgatta

2941   agcattggta actgtcagac caagtttact catatatact ttagattgat ttaaaacttc

3001   atttttaatt taaaaggatc taggtgaaga tccttttga taatctcatg accaaaatcc

3061   cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc aaaggatctt

3121   cttgagatcc ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac

3181   cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag gtaactggct

3241   tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta ggccaccact

                                                  AlwNI
3301   tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta ccagtggctg

3361   ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag ttaccggata

3421   aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca gcccagcttg gagcgaacga

3481   cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg cttcccgaag

3541   ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg

3601   agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc cacctctgac

3661   ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca

3721   acgcggcctt tttacggttc ctggcctttt gctggccttt tgctcacatg ttctttcctg
```

```
3781   cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct gataccgctc

3841   gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa gagcgcctga

3901   tgcggtattt tctccttacg catctgtgcg gtatttcaca ccgcataaat tccgacacca

                Van91I
3961   tcgaatggtg caaaaccttt cgcggtatgg catgatagcg cccggaagag agtcaattca

                                                    BstAPI
4021   gggtggtgaa tgtgaaacca gtaacgttat acgatgtcgc agagtatgcc ggtgtctctt

4081   atcagaccgt ttcccgcgtg gtgaaccagg ccagccacgt ttctgcgaaa acgcgggaaa

4141   aagtggaagc ggcgatggcg gagctgaatt acattcccaa ccgcgtggca caacaactgg

4201   cgggcaaaca gtcgttgctg attggcgttg ccacctccag tctggccctg cacgcgccgt

4261   cgcaaattgt cgcggcgatt aaatctcgcg ccgatcaact gggtgccagc gtggtggtgt

4321   cgatggtaga acgaagcggc gtcgaagcct gtaaagcggc ggtgcacaat cttctcgcgc

4381   aacgcgtcag tgggctgatc attaactatc cgctggatga ccaggatgcc attgctgtgg

4441   aagctgcctg cactaatgtt ccggcgttat ttcttgatgt ctctgaccag acacccatca

4501   acagtattat tttctcccat gaagacggta cgcgactggg cgtggagcat ctggtcgcat

                                        Bsp120I
       BstEII                           ApaI
4561   tgggtcacca gcaaatcgcg ctgttagcgg gcccattaag ttctgtctcg gcgcgtctgc

4621   gtctggctgg ctggcataaa tatctcactc gcaatcaaat tcagccgata gcggaacggg

4681   aaggcgactg gagtgccatg tccggttttc aacaaaccat gcaaatgctg aatgagggca

                                                                BssHII
4741   tcgttcccac tgcgatgctg gttgccaacg atcagatggc gctgggcgca atgcgcgcca

                                  EcoRV
4801   ttaccgagtc cgggctgcgc gttggtgcgg atatctcggt agtgggatac gacgataccg

                        HpaI
4861   aagacagctc atgttatatc ccgccgttaa ccaccatcaa acaggatttt cgcctgctgg

4921   ggcaaaccag cgtggaccgc ttgctgcaac tctctcaggg ccaggcggtg aagggcaatc

                                                    NarI
                                                    KasI
                                                    EheI
                                                    BbeI
4981   agctgttgcc cgtctcactg gtgaaaagaa aaaccaccct ggcgcccaat acgcaaaccg

5041   cctctccccg cgcgttggcc gattcattaa tgcagctggc acgacaggtt tcccgactgg

5101   aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcactcatta ggcaccccag
```

```
5161   gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg ataacaattt

5221   cacacaggaa acagctatga ccatgattac ggattcactg gccgtcgttt tacaacgtcg

5281   tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccttcgc

5341   cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct

5401   gaatggcgaa tggcgctttg cctggtttcc ggcaccagaa gcggtgccgg aaagctggct

                    Bsu36I
5461   ggagtgcgat cttcctgagg ccgatactgt cgtcgtcccc tcaaactggc agatgcacgg

5521   ttacgatgcg cccatctaca ccaacgtaac ctatcccatt acggtcaatc cgccgtttgt

5581   tcccacggag aatccgacgg gttgttactc gctcacattt aatgttgatg aaagctggct

5641   acaggaaggc cagacgcgaa ttatttttga tggcgttgga att
```

## Appendix 8.4




Molecule: 1-k-link-4-gfp, 6026 bps DNA Circular
File Name: 1-k-link-4-gfp plasmid.cm5, dated 16 May 2012

Description: Ligation of asc1 cut eco-GFP into 1-k-link-4 cut,pure

Notes:

Molecule: 1-k-link-4-gfp, 6026 bps DNA Circular
Description: Ligation of asc1 cut eco-GFP into 1-k-link-4 cut,pure
File Name: 1-k-link-4-gfp plasmid.cm5, dated 16 May 2012
Printed: 1-6026 bps (Full), format Annotated: Enzymes, Genes

```
      1  tggcgaatgg gacgcgccct gtagcggcgc attaagcgcg gcgggtgtgg tggttacgcg

                                                 BsrBI
                                HaeII      HaeII
     61  cagcgtgacc gctacacttg ccagcgccct agcgcccgct cctttcgctt tcttcccttc

                                                                Bsp1286I
                          Cfr10I                                BanII
    121  ctttctcgcc acgttcgccg gctttccccg tcaagctcta aatcgggggc tccctttagg

    181  gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgattagg gtgatggttc

         DraIII                                                DrdI
    241  acgtagtggg ccatcgccct gatagacggt ttttcgccct ttgacgttgg agtccacgtt

    301  ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct cggtctattc

    361  ttttgattta taagggattt gccgatttcg ggcctattgg ttaaaaaatg agctgattta

                                          SspI
    421  acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaatttctg gcggcacgat

                                                DraI                DraI
    481  ggcatgagat tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt

    541  aaatcaatct aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt

                                                                  AhdI
    601  gaggcaccta tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc

    661  gtgtagataa ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg

                                BpmI
         BsaI          Cfr10I                                   BglI
    721  cgagacccac gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc

    781  gagcgcagaa gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg

                                                FspI
    841  gaagctagag taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca

                                                          BsaWI
    901  ggcatcgtgg tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga

    961  tcaaggcgag ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct

         PvuI
         BsiEI
                                EaeI
   1021  ccgatcgttg tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg

                                                                ScaI
   1081  cataattctc ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca

                                                BsiEI
   1141  accaagtcat tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata

                                                Bsp1286I
                                          DraI  BsiHKAI
   1201  cgggataata ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct
```

```
1261  tcggggcgaa aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact

                Bsp1286I
                BsiHKAI                 Eco57I
1321  cgtgcaccca actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa

1381  acaggaaggc aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc

          EarI                                                    BsrBI
1441  atactcttcc tttttcaatc atgattgaag catttatcag ggttattgtc tcatgagcgg

1501  atacatattt gaatgtattt agaaaaataa acaaataggt catgaccaaa atcccttaac

1561  gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag

1621  atccttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg

                                                              Eco57I
1681  tggtttgttt gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca

1741  gagcgcagat accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga

                                                    AlwNI
1801  actctgtagc accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca

                                                  BsaWI
1861  gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc

          BsiEI                   Bsp1286I
                                  BsiHKAI
1921  agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca

                                        HaeII
1981  ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa

                BsaWI
2041  aggcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc

                                                              DrdI
2101  caggggaaa cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc

2161  gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg

                                          BspLU11I
                                          AflIII
2221  ccttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat

                                                    BsrBI
2281  cccctgattc tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca

                                                HaeII
                                              SapI
          BsiEI                               EarI
2341  gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ctgatgcggt

                                                Bsp1286I
                                                BsiHKAI
2401  attttctcct tacgcatctg tgcggtattt cacaccgcat atatggtgca ctctcagtac

                                  Bst1107I                      Tth111I
2461  aatctgctct gatgccgcat agttaagcca gtatacactc cgctatcgct acgtgactgg

                                                      DrdI
2521  gtcatggctg cgccccgaca cccgccaaca cccgctgacg cgccctgacg ggcttgtctg

2581  ctcccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg
```

2641 ttttcaccgt catcaccgaa acgcgcgagg cagctgcggt aaagctcatc agcgtggtcg

BpmI
2701 tgaagcgatt cacagatgtc tgcctgttca tccgcgtcca gctcgttgag tttctccaga

2761 agcgttaatg tctggcttct gataaagcgg gccatgttaa gggcggtttt ttcctgtttg

2821 gtcactgatg cctccgtgta agggggattt ctgttcatgg gggtaatgat accgatgaaa

2881 cgagagagga tgctcacgat acgggttact gatgatgaac atgcccggtt actggaacgt

2941 tgtgagggta aacaactggc ggtatggatg cggcgggacc agagaaaaat cactcagggt

HaeII
3001 caatgccagc gcttcgttaa tacagatgta ggtgttccac agggtagcca gcagcatcct

BsaWI
AccIII HaeII
3061 gcgatgcaga tccggaacat aatggtgcag ggcgctgact tccgcgtttc cagactttac

Bpu10I
3121 gaaacacgga aaccgaagac cattcatgtt gttgctcagg tcgcagacgt tttgcagcag

3181 cagtcgcttc acgttcgctc gcgtatcggt gattcattct gctaaccagt aaggcaaccc

Bsp1286I
PpuMI BsiHKAI
3241 cgccagccta gccgggtcct caacgacagg agcacgatca tgctagtcat gccccgcgcc

BsaWI BsiEI
3301 caccggaagg agctgactgg gttgaaggct ctcaagggca tcggtcgaga tcccggtgcc

3361 taatgagtga gctaacttac attaattgcg ttgcgctcac tgcccgcttt ccagtcggga

EaeI
3421 aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt

HaeII
3481 attgggcgcc agggtggttt ttcttttcac cagtgagacg ggcaacagct gattgccctt

3541 caccgcctgg ccctgagaga gttgcagcaa gcggtccacg ctggtttgcc ccagcaggcg

HpaI
HincII
3601 aaaatcctgt ttgatggtgg ttaacggcgg gatataacat gagctgtctt cggtatcgtc

3661 gtatcccact accgagatgt ccgcaccaac gcgcagcccg gactcggtaa tggcgcgcat

HaeII
3721 tgcgcccagc gccatctgat cgttggcaac cagcatcgca gtgggaacga tgccctcatt

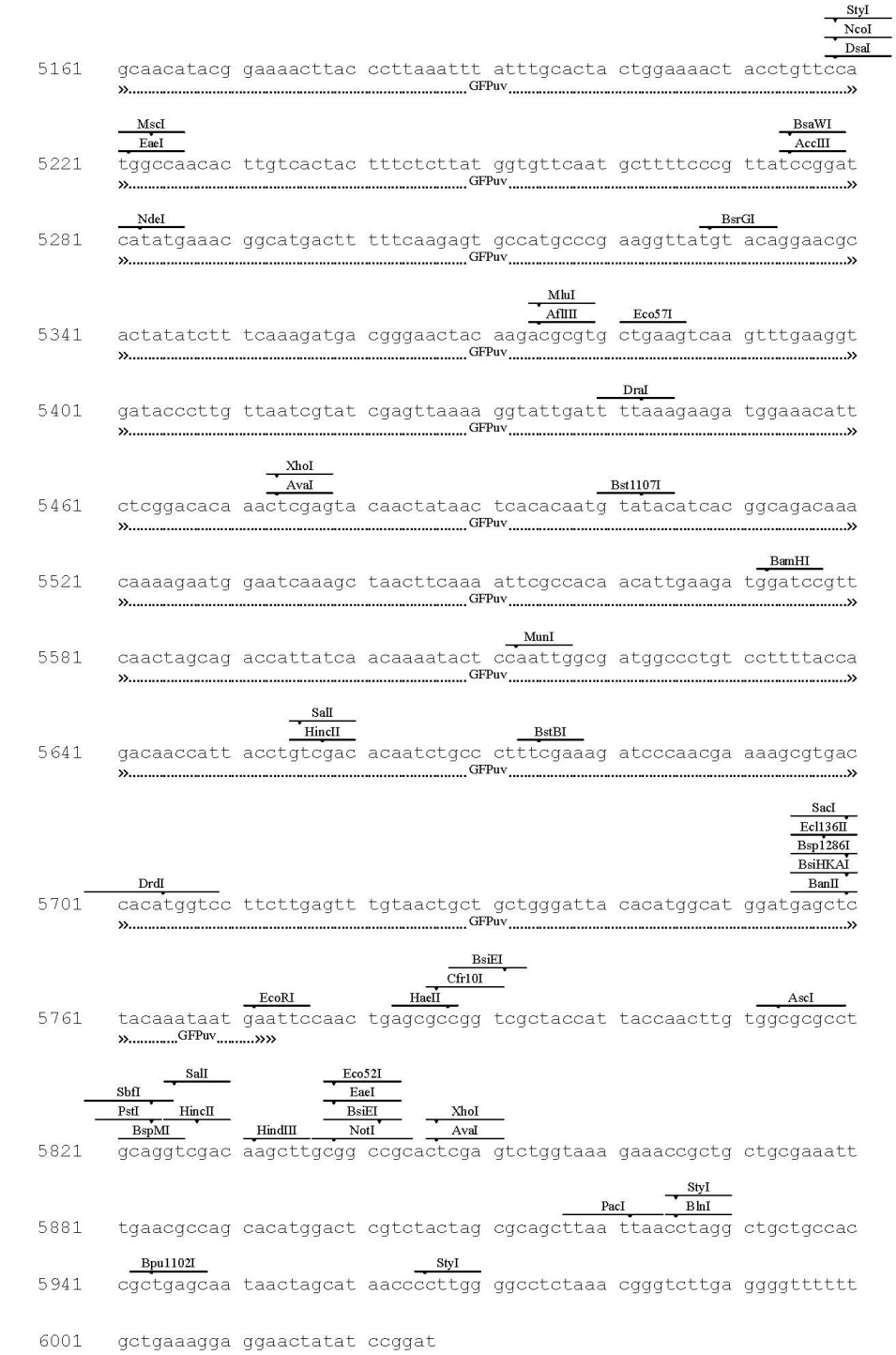
BsaWI BpmI
3781 cagcatttgc atggtttgtt gaaaaccgga catggcactc cagtcgcctt cccgttccgc

3841 tatcggctga atttgattgc gagtgagata tttatgccag ccagccagac gcagacgcgc

Bsp1286I
Bsp120I
BanII
ApaI BstEII
3901 cgagacagaa cttaatgggc ccgctaacag cgcgatttgc tggtgaccca atgcgaccag

3961 atgctccacg cccagtcgcg taccgtcttc atgggagaaa ataatactgt tgatgggtgt


4021 ctggtcagag acatcaagaa ataacgccgg aacattagtg caggcagctt ccacagcaat
MluI
AflIII
BclI
4081 ggcatcctgg tcatccagcg gatagttaat gatcagccca ctgacgcgtt gcgcgagaag
Bsp1286I
BsiHKAI
4141 attgtgcacc gccgctttac aggcttcgac gccgcttcgt tctaccatcg acaccaccac
4201 gctggcaccc agttgatcgg cgcgagattt aatcgccgcg acaatttgcg acggcgcgtg
BpmI
4261 cagggccaga ctggaggtgg caacgccaat cagcaacgac tgtttgcccg ccagttgttg
4321 tgccacgcgg ttgggaatgt aattcagctc cgccatcgcc gcttccactt tttcccgcgt
Cfr10I
4381 tttcgcagaa acgtggctgg cctggttcac cacgcgggaa acggtctgat aagagacacc
BstAPI
4441 ggcatactct gcgacatcgt ataacgttac tggtttcaca ttcaccaccc tgaattgact
EarI
HaeII
4501 ctcttccggg cgctatcatg ccataccgcg aaaggttttg cgccattcga tggtgtccgg
EcoNI
4561 gatctcgacg ctctccctta tgcgactcct gcattaggaa gcagcccagt agtaggttga
Bsp1286I
BsiHKAI
SphI
HaeII
4621 ggccgttgag caccgccgcc gcaaggaatg gtgcatgcaa ggagatggcg cccaacagtc
DsaI
EaeI
HaeII
Bsp1286I
BanII
4681 ccccggccac ggggcctgcc accataccca cgccgaaaca agcgctcatg agcccgaagt
Bsp1286I
BanII
HaeII
4741 ggcgagcccg atcttcccca tcggtgatgt cggcgatata ggcgccagca accgcacctg
Cfr10I
SgrAI
HaeII
EaeI
Cfr10I
ClaI
4801 tggcgccggt gatgccggcc acgatgcgtc cggcgtagag gatcgagatc gatctcgatc
BsrBI
XbaI
4861 ccgcgaaatt aatacgactc actatagggg aattgtgagc ggataacaat tcccctctag
StyI
NcoI
DsaI
Eco72I
4921 aaataatttt gtttaacttt aagaaggaga tataccatgg cacatcacca ccaccatcac
Cfr10I
BsaWI
AgeI
KpnI
Acc65I
BstBI
BsrBI
SapI
EarI
Bsp1286I
BanII
AvaI
4981 gtgggtaccg gttcgaatga tgacgacgac aagaagagcg gctcgggctc tgggagtggt
BpmI
5041 agcggcagtg ggatgagtaa aggagaagaa cttttcactg gagttgtccc aattcttgtt
»»»..... GFPuv .....»
Bsp1286I
5101 gaattagatg gtgatgttaa tgggcacaaa ttttctgtca gtggagaggg tgaaggtgat
».....GFPuv.....»


StyI
NcoI
DsaI
5161 gcaacatacg gaaaacttac ccttaaattt atttgcacta ctggaaaact acctgttcca
GFPuv
MscI
EaeI
BsaWI
AccIII
5221 tggccaacac ttgtcactac tttctcttat ggtgttcaat gcttttcccg ttatccggat
GFPuv
NdeI
BsrGI
5281 catatgaaac ggcatgactt tttcaagagt gccatgcccg aaggttatgt acaggaacgc
GFPuv
MluI
AflIII
Eco57I
5341 actatatctt tcaaagatga cgggaactac aagacgcgtg ctgaagtcaa gtttgaaggt
GFPuv
DraI
5401 gataccctg ttaatcgtat cgagttaaaa ggtattgatt ttaaagaaga tggaaacatt
GFPuv
XhoI
AvaI
Bst1107I
5461 ctcggacaca aactcgagta caactataac tcacacaatg tatacatcac ggcagacaaa
GFPuv
BamHI
5521 caaaagaatg gaatcaaagc taacttcaaa attcgccaca acattgaaga tggatccgtt
GFPuv
MunI
5581 caactagcag accattatca acaaaatact ccaattggcg atggccctgt ccttttacca
GFPuv
SalI
HincII
BstBI
5641 gacaaccatt acctgtcgac acaatctgcc ctttcgaaag atcccaacga aaagcgtgac
GFPuv
SacI
Ecl136II
Bsp1286I
BsiHKAI
DrdI
BanII
5701 cacatggtcc ttcttgagtt tgtaactgct gctgggatta cacatggcat ggatgagctc
GFPuv
BsiEI
Cfr10I
EcoRI
HaeII
AscI
5761 tacaaataat gaattccaac tgagcgccgg tcgctaccat taccaacttg tggcgcgcct
GFPuv
SalI
Eco52I
SbfI
EaeI
PstI
HincII
BsiEI
XhoI
BspMI
HindIII
NotI
AvaI
5821 gcaggtcgac aagcttgcgg ccgcactcga gtctggtaaa gaaaccgctg ctgcgaaatt
StyI
PacI
BlnI
5881 tgaacgccag cacatggact cgtctactag cgcagcttaa ttaacctagg ctgctgccac
Bpu1102I
StyI
5941 cgctgagcaa taactagcat aacccccttgg ggcctctaaa cgggtcttga ggggtttttt
6001 gctgaaagga ggaactatat ccggat