# THE DEVELOPMENT OF A MODEL SYSTEM AND HIGH THROUGHPUT ASSAY FOR THE STUDY OF INTERACTIONS BETWEEN ZINC FINGER PROTEINS AND DNA.

**LEONIE JANE MORGAN**
Doctor of Philosophy

**ASTON UNIVERSITY**
September 1998

ASTON UNIVERSITY

# THE DEVELOPMENT OF A MODEL SYSTEM AND HIGH THROUGHPUT ASSAY FOR THE STUDY OF INTERACTIONS BETWEEN ZINC FINGER PROTEINS AND DNA.

Leonie Jane Morgan
Doctor of Philosophy
September 1998

## SUMMARY

It has been recognised for some time that a full code of amino acid-based recognition of DNA sequences would be useful. Several approaches, which utilise small DNA binding motifs called zinc fingers, are presently employed. None of the current approaches successfully combine a combinatorial approach to the elucidation of a code with a single stage high throughput screening assay. The work outlined here describes the development of a model system for the study of DNA protein interactions and the development of a high throughput assay for detection of such interactions. A zinc finger protein was designed which will bind with high affinity and specificity to a known DNA sequence. For future work it is possible to mutate the region of the zinc finger responsible for the specificity of binding, in order to observe the effect on the DNA / protein interactions. The zinc finger protein was initially synthesised as a His tagged product. It was not possible however to develop a high throughput assay using the His tagged zinc finger protein. The gene encoding the zinc finger protein was altered and the protein synthesised as a Glutathione S-Transferase (GST) fusion product. A successful assay was developed using the GST protein and Scintillation Proximity Assay technology (Amersham Pharmacia Biotech). The scintillation proximity assay is a dynamic assay that allows the DNA protein interactions to be studied in "real time". This assay not only provides a high throughput method of screening zinc finger proteins for potential ligands but also allows the effect of addition of reagents or competitor ligands to be monitored.

**Keywords:** Zinc fingers, DNA, DNA/Protein interactions, His tag, Glutathione S-Transferase (GST), Scintillation Proximity Assay (SPA).

## ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Dr A. V. Hine for her enthusiasm, encouragement, and guidance throughout this project. I would also like to thank Dr M. Piccardo, Dr I. Jessop and Mrs K. Howells, at Amersham Pharmacia Biotech for their guidance and generous contribution of reagents, advice and ideas.

I would also like to gratefully acknowledge the Biotechnological and Biological Science Research Council (BBSRC), for their financial support of this project.

I would like to take this opportunity to thank Prof. D. C. Billington and Dr P. A. Lambert for all their guidance and support throughout the past three years. Dr C Schwalbe for useful discussions and help with x-ray crystallography and Dr T McDevitt for her early help and support with the gene cloning. I would also like to express my gratitude to the technical staff in the Department of Pharmaceutical and Biological Sciences, particularly Mr K. Hughes for all his help with radiochemical work and Mr R. Tilling for his patience with my sometimes urgent reagent ordering.

Finally, I cannot thank enough my family and friends for their support and encouragement throughout my years of study.

# CONTENTS

## LIST OF FIGURES.

## LIST OF TABLES.

## ABBREVIATIONS.

Conventional three letter and single letter abbreviations are used for amino acids though out the text.

| | |
|---|---|
| Å | Ångström ($10^{-10}$m) |
| ATP | Adenosine triphosphate |
| BCIP | 5-bromo-4-chroro-3indoyl phosphate |
| BSA | Bovine serum albumin |
| dATP | Deoxyadenosine 5'-triphosphate |
| dCTP | Deoxycytidine 5'-triphosphate |
| dGTP | Deoxyguanosine 5'-triphosphate |
| DNA | Deoxyribonucleic acid |
| dTTP | Deoxythymidine 5'-triphosphate |
| DTT | Dithiothreitol |
| EDTA | Ethylene diamine tetra-acetic acid |
| *E. coli* | *Escherichia coli* |
| $K_d$ | Dissociation constant |
| hr (s) | Hour (s) |
| kDa | Kilodaltons |
| l | Litre (s) |
| LB | Luria broth |
| M | Mole (s) per litre |
| m | Milli |
| μ | Micro |
| min (s) | Minute (s) |
| MWCO | Molecular Weight Cut Off |
| g | Gram (s) |
| NBT | Nitro blue tetrazolium |
| nm | Nanometre (s) |
| OD | Optical density |
| PAGE | Polyacrylamide gel electrophoresis |

| | |
|---|---|
| PBS | Phosphate buffered saline |
| PCR | Polymerase chain reaction |
| PMSF | Phenylmethlysulphonylflouride |
| rpm | Revolutions per minute |
| SDS | Sodium dodecyl sulphate |
| sec (s) | Second (s) |
| SPA | Scintillation proximity assay |
| TBE | Tris borate EDTA |
| TEMED | N,N,N',N',-tetramethylethylenediamine |
| TF | Transcription factor (s) |
| Tris | Tris (hydroymethyl)aminomethane |
| X-gal | 5-Bromo-4-chloro-3-indoyl β-D-galactopyranoside |

# 1 INTRODUCTION.

## 1.1 BACKGROUND.

### 1.1.1 Historical perspective.

Since the central dogma of molecular biology (DNA → RNA → Protein) was proposed (Crick, 1958) it has become obvious that the flow of information cannot be unidirectional. In order to enable organisms to respond to their environment and for the differentiation of cell types in multicellular organisms there must be a system controlling the expression of genetic information. The first evidence for the control of the expression of genes came from metabolic studies in *E. coli*. In the presence of both glucose and lactose, *E. coli* will metabolize only glucose, indicating that all or some of the gene products responsible for lactose metabolism are missing. Only when there is no longer any glucose present will the *E. coli* start to utilize the lactose. This suggests that *E. coli* is capable of controlling the expression of its genes in response to environmental cues (Miller 1978).

Gene expression might be controlled at one or more levels. These include transcription control, control of RNA degradation and translational control (reviewed by Derman, 1981). The comparison of the species of RNA present in different cell types within an organism revealed that each cell type contained only a specific subset of RNA species. This implied that transcriptional control was the primary method of control of gene expression. Transcriptional control involves a series of events which may be triggered by a variety of intra-cellular inter-cellular or environmental cues. These "feedback systems" range from being as simple as the control of the lac operon in *E. coli* to the complex systems seen in eukaryotes. The final stage in transcriptional control of gene expression however always involves the recognition of specific "regulatory" sequences of DNA by a group of proteins known as transcription factors (TFs).

The first step towards understanding how a protein could recognize a specific DNA sequence was made by Seeman *et al.* (1976). These workers described the specific binding of protein side chains to DNA base pairs, and proposed a model for the interaction between Arg-G and Asn/Gln-A.

Initially only TFs from bacteria could be isolated. It was far more problematic to isolate TFs from eukaryotic cells as they are present at such low concentrations. One of the first TFs to be isolated and characterized by x-ray crystallography was the Lambda cro protein. The cro-protein contains two α-helices separated by a non helical turn. Residues within the α-helices are responsible for the binding of the protein to DNA (Ptashne, 1986).

As the number of transcription factors isolated increased it was recognized that many of them shared similar motifs. The first DNA binding motif to be described was Helix Turn Helix which is the motif seen in the cro-protein (Sauer *et al.*, 1982). This was followed by the discovery of Zinc Finger motifs (Miller *et al.*, 1985), and Leucine Zippers (Landscultz *et al.*, 1988). Many but not all (Raumann *et al.*, 1994) of these motifs use an α-helix to recognize bases in the major grove of the DNA (Pabo and Sauer, 1992)

The interactions between the α-helix of these proteins and the major grove of the DNA initially seemed quite simple and it was thought that there would be a periodic (one amino acid recognizing one base pair) code of interactions between the amino acids in TFs and DNA (Pabo *et al.*, 1990). But as more motifs have been discovered it seems likely that a syllabic code (a combination of amino acids recognizing several base pairs) rather that a periodic code may better explain some DNA protein interactions (Desjarlais and Berg, 1992; Pavletich and Pabo, 1993). Various methods have been used to attempt to decipher a code of interactions between transcription factors and DNA, many of these are based on the use of zinc finger motifs. Zinc finger motifs are ideal for the study of protein-DNA interactions as they recognize DNA in a simple modular manner (reviewed by Choo and Klug 1997).

### 1.1.2 Types of Zinc Binding Motif.

The term zinc finger was first used to describe the repeated motifs found in the *Xenopus* transcription factor IIIA (TFIIIA) (Miller *et al.*, 1985). Since the discovery of the zinc finger motifs in TFIIIA many related structures have been described (reviewed by; Berg, 1990; Berg, 1993; Schmiedeskamp and Klevit, 1994; Schwabe and Klug, 1994 Klug and Schwabe, 1995). All of these domains bind zinc ions in order to stabilize their structure and many are capable of interacting with DNA. Some of the better characterized zinc finger related motifs are described below.

The $Cys_2$ $His_2$ zinc fingers (such as those found in TFIIIA, reviewed by Rhodes and Klug 1993) are the largest and most widespread of the zinc binding domains. They are found in organisms ranging from yeast to humans (Berg, 1993), and are predicted to be present in 300-700 human genes which represents almost 1% of the human genome (Hoovers *et al.*, 1992). $Cys_2His_2$ motifs fold to form a β-sheet containing the two invariant cysteine residues and an α-helix containing the conserved histidine residues. The secondary structure of the protein is stabilized by the tetrahedral co-ordination of a zinc ion between the two cysteine and two histidine residues (structure described in further detail in 1.2.2).

The DNA binding domain of the GATA 1 protein represented the first member of the GATA motifs. GATA binding motifs fold into a secondary structure which is similar to that of the $Cys_2His_2$ motif. The structure is stabilized by the co-ordination of a zinc ion by four invariant cysteine residues. GATA1 contains two such DNA binding domains, only one of which is utilized for the binding of the protein to DNA (Omichinski *et al.*, 1993).

The closely related Glucocorticoid (GR) and Oestrogen receptors (ER) are both members of a group of proteins known as nuclear hormone receptors (reviewed by Freedman and Luisi, 1993). Both proteins are ligand activated, they must be associated with their respective hormone in order to bind to DNA. Each contains several domains that perform various functions. For instance, in the ER protein, region E is responsible for hormone binding, region C for DNA binding and region D links these two functional domains (Kumar *et al.*, 1987). X-ray crystal structures for both the GR (Lusi *et al.*, 1993) and the ER (Schwabe *et al.*, 1993) are available. Within each protein the DNA binding regions fold to form two zinc binding domains each containing an α-Helix, in a

structure which is similar to that seen in the $Cys_2His_2$ motifs. The structure is stabilized by the co-ordination of two zinc ions between two sets of four cysteine residues. Both ER and GR bind to DNA as a dimer, recognising palindromic DNA sequences that differ in only two bases. Each subunit recognises half of a palindromic site (6 base pairs) and each half of the site is separated by 3 base pairs.

The LIM domain was named after the first three proteins in which it was discovered, Lin-11, Isl-1 and Mec-3 (Fryed *et al.*, 1990; Karlsson *et al* 1990). The LIM motif forms a domain with two independent zinc binding sites. Three cysteines and one histidine bind the first ion and four cysteines the second. The secondary structure is essentially composed of β-sheets, and bears no particular resemblance to the structure of the $Cys_2His_2$ motifs (Petrez-Alvarado *et al.*, 1994). LIM domains do however bind to DNA, and seem to be capable of interacting in arrays to recognise a larger target (Sadler *et al.*, 1992).

The first binuclear cluster to be described was the DNA binding domain of yeast GAL-4 (Pan and Coleman 1989). Two zinc ions are co-ordinate by six cysteine residues in a central cluster around which the rest of the domain is wrapped. Binuclear clusters are structurally distinct from the zinc binding domains described above in so much as the two fingers are "co-joined" by the binuclear cluster (Kraulis *et al.*, 1992). The structure of the zinc fingers however is not dissimilar to the structure of $Cys_2His_2$ motifs. Binuclear clusters can also bind to specific DNA sequences as dimers (Baleja *et al.*, 1992).

Retroviral nucleocapsid proteins contain one or two copies of a $Cys_3His_1$ zinc binding motif. They appear to play a variety of roles in the viral life cycle all of which involve nucleic acid binding. Three cysteine and one histidine residues are responsible for the binding of a single zinc ion. The structure of the protein is significantly different to that found in the $Cys_2His_2$ motifs, and is more similar to the structure of rubredoxin (Summers *et al.*, 1990).

The Transcription factor IIS (TFIIS) plays a role in transcriptional elongation. The DNA binding domain of TFIIS folds into three antiparalel β-sheets stabilized by the co-ordination a zinc ion between four cysteine residues. The secondary structure again bears no resemblance to the $Cys_2His_2$ motifs (Quain *et al.*, 1993).

The zinc binding domains of Protein Kinase C and the RING finger are structurally distinct to other zinc binding motifs as they both use two zinc ions to stabilize one structural unit. The RING or $Cys_3HisCys_4$ domain binds two zinc ions. The first zinc ion is bound by one histidine and three cysteine residues, and the other by four cysteine residues. This motif is common in proteins found in the nucleus (Barlow *et al.*, 1994). Protein Kinase C contains a domain which binds two zinc ions in a similar manner to the RING finger. Its function as a domain is not clear, but it has now been found in several oncogenic products (Hommel *et al.*, 1994).

Some of the zinc binding motifs such as $Cys_2His_2$ motifs, the GATA DNA binding domain and the DNA binding domains found in the nuclear hormone receptors (ER and GR) may be evolutionarily related. The only similarity between all of the motifs described here is the ability to co-ordinate a zinc ion in order to stabilize their secondary structure. However, most of the motifs do seem to bind to nucleic acids, and therefore may have similar functions. For the purpose of the current study the term zinc finger will be used to describe only the $His_2Cys_2$ zinc motif.

## 1.2 $CYS_2HIS_2$ ZINC FINGER MOTIFS.

### 1.2.1 Function of zinc finger proteins.

Zinc fingers occur in many proteins and are thought to be utilized largely for the sequence specific recognition of DNA. Proteins containing as few as two zinc fingers and as many as thirty seven have been described (Jacobs, 1992). Although in many cases the zinc fingers are utilized for DNA recognition, there are likely to be exceptions. For example out of the nine motifs in TFIIIA only seven are responsible for base specific contacts and two bridge the minor grove acting as "spacers" (Nolte *et al.*, 1998).

Many proteins containing zinc fingers are involved in the regulation of expression. Among some of the first zinc finger proteins to be described were transcription factors such as TFIIIA (found at high abundance in *Xenopus* oocytes, Miller *et al*, 1985) and the human transcription factor Sp1 (Gidoni *et al.*, 1984). Many of the genes containing zinc finger motifs are involved in development, for example the mouse immediate early gene Zif268 (Christy *et al*, 1988). The aberrant expression or mutation of some zinc finger genes is important in the development of cancer. Proto-oncogenes containing zinc finger motifs include GL1 (Kinzler *et al.*, 1988) and the Wilms' tumor gene (Call *et al.*, 1990).

### 1.2.2 Structure of $Cys_2His_2$ zinc finger motifs.

$Cys_2His_2$ Zinc fingers are 28-30 amino acids long, with highly conserved cysteine, histidine and hydrophobic residues (Fig 1.1). The invariant cysteine and histidine residues were predicted to co-ordinate a single zinc ion tetrahedrally in order to stabilize the structure of the DNA binding domain (Miller *et al.*, 1985). The secondary structure of the $Cys_2His_2$ zinc finger was originally predicted by model building with comparison to other metal binding proteins (Berg, 1987). This structure was later confirmed by nuclear magnetic resonance spectroscopy on single zinc finger motifs from the *Xenopus* protein Xfn (Lee *et al*, 1989), yeast ADR1 (Klevit *et al.*, 1990), a consensus zinc finger peptide (Krizek *et al.*, 1991) and also a "minimal" zinc finger motif (Michael *et al.*, 1992).

The identity of the contacting residues (α. β and γ, Fig 1.1), which are largely responsible for the specificity of the interaction with DNA was predicted by mutagenesis studies (Nardelli *et al.*, 1991), data base analysis (Jacobs, 1992) and confirmed by crystal structure analysis (Pavletich and Pabo, 1991). In most zinc finger proteins these residues are responsible for the specificity of the interaction of the protein with the DNA although there are exceptions (Takatsuji, 1996)

Key

- DNA contacting residues
- cysteine residues
- histidine residues
- Conserved Hydrophobic residues
- DNA contacting residue which is occasionally involved in the specificity of the binding of the zinc finger motif to the DNA

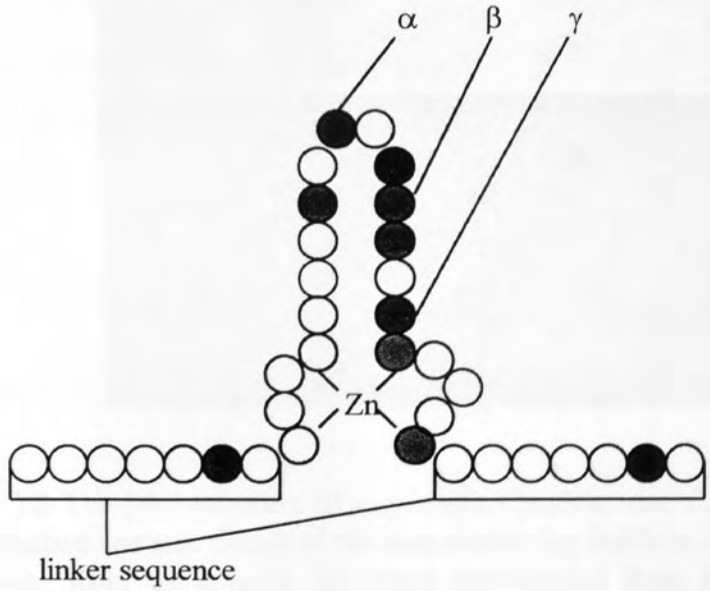



Fig 1.1 Structure of the zinc finger mini-domain. The consensus sequence = Tyr/Phe-X-Cys-$X_{2-4}$-Cys-$X_3$-Phe-$X_5$-Leu-$X_2$-His-$X_{3-5}$-His (X= any amino acid (Miller *et al.*, 1985)). Arrays of zinc fingers are usually connected by linkers (Berg, 1993).

When the $Cys_2His_2$ zinc finger is folded into its three dimensional structure (Fig 1.2) the zinc co-ordinating cysteine residues are within the hairpin of an irregular antiparallel β-sheet. The β-sheet extends towards the "tip" of the finger. The other side of the finger forms an α-helix containing the DNA contacting residues and the zinc co-ordinating histidine residues (Lee *et al*, 1989; Krizek *et al*., 1991; Michael *et al*., 1992). The residues which are thought to be largely responsible for the specificity of the DNA interactions (α, β and γ) are placed on the outside of the α helix (Klevit *et al*., 1990).


Figure 1.2 The ββα structure of a synthetic $Cys_2His_2$ zinc finger, showing the zinc ion co-ordination and side chains of the base contacting residues (α=Glu, β=Ala and γ=Arg) protruding from the α-helix. Structure downloaded from the Brookhaven data base (Brookhaven ref. pdb1znf.ent).

The ββα structure of the $Cys_2His_2$ zinc finger is extremely stable. It requires only the three conserved hydrophobic residues, the conserved cysteines and histidines to fold correctly. All the residues apart from these may be replaced by alanine and the ββα structure will still form in the presence of zinc (Michael *et al.*, 1992). The ββα structure seems to be a minimal structure for a functional $Cys_2His_2$ zinc finger, and in some cases additional structures may be present to stabilize it (Klug and Schwabe, 1995). The first zinc finger in SW15 interacts with residues that lie N terminal to the zinc finger motifs (Neuhaus *et al.*, 1992). This structure primarily enhances the stability but may also extend the potential DNA binding surface of the zinc finger (Dutnall, 1996).

Solution structures of proteins containing several zinc finger motifs revealed that each motif retained the structure observed in single motifs (Neuhaus *et al.*, 1992). Adjacent motifs formed what seemed to be structurally independent DNA binding domains joined by flexible linkers (Nakeseko *et al.*, 1992). The linkers not only act to join zinc finger motifs together but are also involved in the binding of the zinc finger proteins to DNA and space the motifs such that each motif can bind to a DNA subsite (Choo and Klug, 1993). The zinc finger linker was initially thought to be important because it tends to be highly conserved (Jacobs, 1992). It is so highly conserved that oligonucleotides representing its DNA sequence have successfully been used for the screening of libraries to identify novel zinc finger proteins (Becker *et al.*, 1995). Deviations of only one amino acid from the consensus linker sequence can significantly lower the binding of an entire zinc finger protein to DNA (Choo and Klug, 1993).

### 1.2.3 The Structure of $His_2Cys_2$ zinc finger / DNA Complexes.

Zinc fingers typically recognize 3 base subsites in DNA, arrays of zinc fingers being capable of recognition of larger sites. They are highly versatile DNA binding motifs, specifically recognizing both A+T rich sequences (CF2II, Gogos *et al.*, 1996) and G+C rich sequences (Zif268, Jamieson *et al.*, 1996). Proteins containing zinc finger motifs have been shown to interact with not only DNA but also RNA (Shi and Berg, 1995b) and DNA / RNA hybrids (Christy *et al.*, 1988). Certain motifs are even capable of discrimination between methylated and non-methylated DNA (Choo, 1998).

The interaction of zinc fingers with DNA was first demonstrated by the x-ray crystallography structure of fingers 1, 2 and 3 of Zif268 (Pavletich and Pabo, 1991).

There are now x-ray crystal structures available of 5 zinc fingers from GLI (Pavletich and Pabo 1993), 2 zinc fingers from Tramtrack (Fairall *et al.*, 1993) and 6 zinc fingers from Transcription factor IIIA (1998). The overall structure of Zif268 bound to DNA seems to be fairly typical of zinc finger proteins although the base specific contacts may vary. This structure has been solved on two separate occasions at 2.1 Å (Pavletich and Pabo, 1991) and at 1.6 Å (Elrod-Erickson *et al.*, 1998). It will therefore be used here to illustrate DNA / zinc finger interactions.

On binding of the zinc finger domains to DNA, the amino terminus of the α-helix angles itself down to fit into the major grove of the DNA. The helix aligns itself such that the residues primarily responsible for the specificity of binding of the protein to the DNA are aligned with a 3 base subsite. The β sheet lies behind the helix, away from the base pairs. The fingers arrange themselves around the major grove of the DNA in a C shape binding the DNA as discrete modules (Fig. 1.3).


Figure 1.3 Crystal structure of the Zif268 DNA interaction, showing the insertion of the α-helix into the major grove of the DNA. Structure downloaded from the Brookhaven data base (Brookhaven ref. pdb1zaa.ent).

The binding of the zinc finger to the strand it recognizes is anti-parallel with respect to the 5'-3' and N-C conventions (Nardelli *et al.*, 1991). Zinc finger 1 recognizing the most 3' DNA subsite (Fig. 1.4).

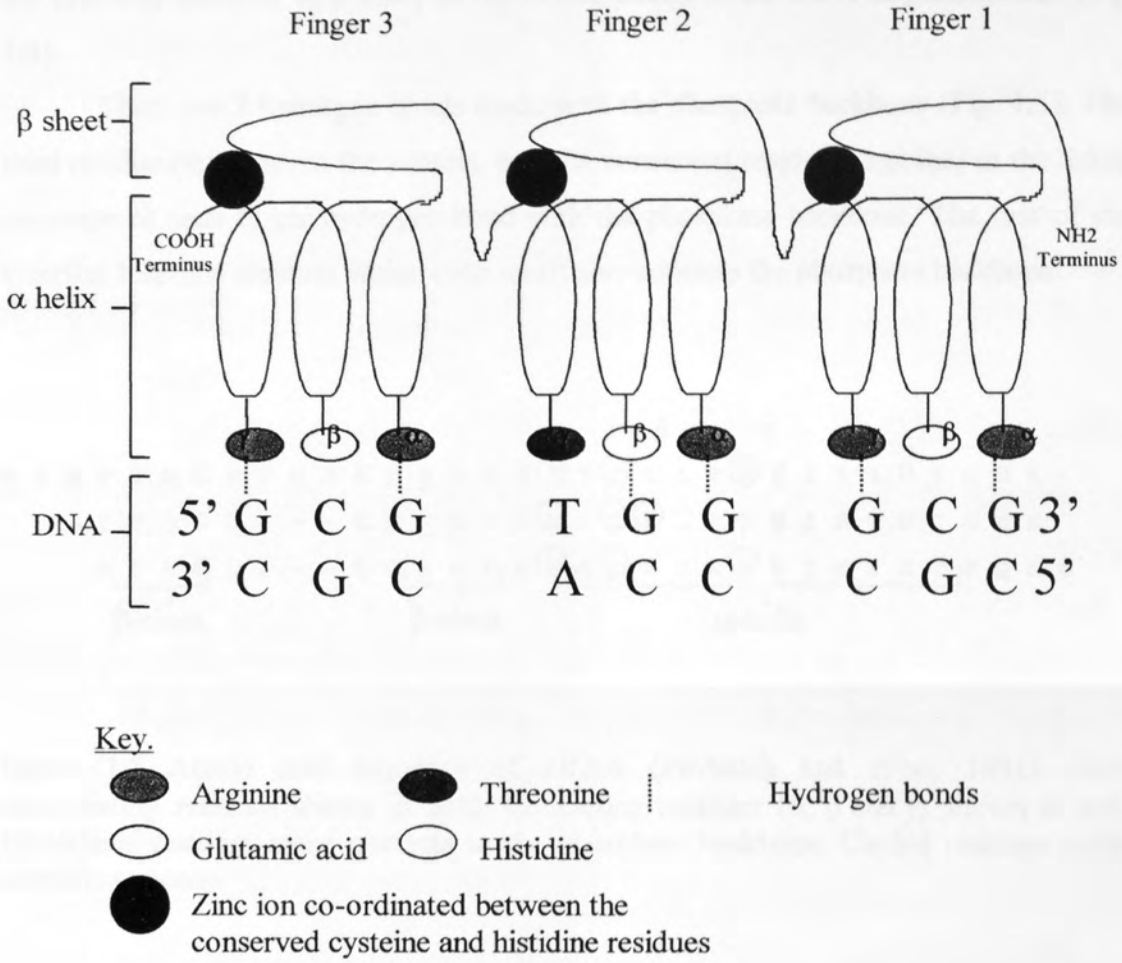



Figure 1.4 Anti - parallel binding of zinc finger motifs of Zif268 to its target DNA (Pavletich and Pabo, 1991).

The Zif 268 / DNA interaction is stabilized by a relatively simple network of hydrogen bonds. Some of these (such as the contact made by the first zinc co-ordinating histidine) are non-specific with regards to the DNA sequence and involve the phosphate back bone rather than the bases. Contacts responsible for the specificity of DNA binding are primarily made by α, β and γ to the G rich strand of the DNA target sequence (Fig 1.4).

There are 7 hydrogen bonds made with the phosphate backbone (Fig. 1.5). The third residue (agrinine) in the protein, and the conserved residue (arginine) in the linker sequence of each finger hydrogen bond with the phosphate backbone. The first of the invariant histidine residues within each motif also contacts the phosphate backbone.

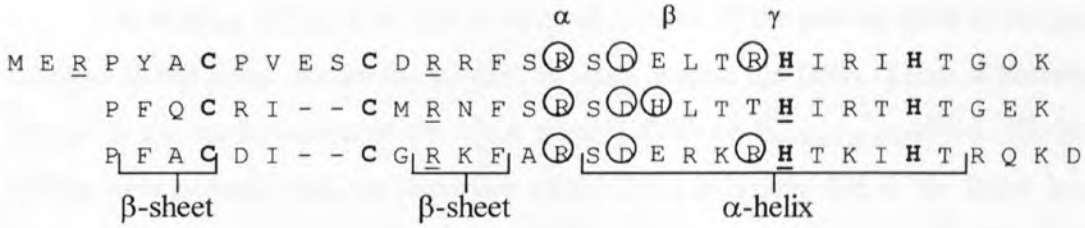



Figure 1.5 Amino acid sequence of Zif268 (Pavletich and Pabo, 1991). Zinc coordinating residues shown in bold. Contacting residues (α, β and γ) shown in red. Underlined residues make contacts to the phosphate backbone. Circled residues make contacts to bases.

Nine contacts are made by amino acids to DNA bases (Fig 1.5). Six of these contacts are primarily responsible for the specificity of the DNA binding, as illustrated in Fig. 1.4. The remaining three contacts are made by the asparagine residues indicated in Fig. 1.5. These are made to the to base at the 3' end of the preceding subsite, on the opposite (C rich) strand of the DNA to the other base specific contacts.

The Asp in finger 1 does not contribute to the specificity of the binding between the protein and the DNA as the contact that it makes is to a base outside of the binding site. The Asp in finger 2 contacts the C of the CG base pair at the 3' end of the subsite recognized by finger 1. The identity of the base at this position is already determined by the Arg (at position γ) in finger 1. The Asp contact made by finger 2 is therefore unlikely

to contribute to the specificity of the binding. The asparagine contacts made by fingers 1 and 2 will strengthen the overall binding of the protein to the DNA but are unlikely to contribute to the specificity of the binding of the protein to the DNA.

It is possible that the contact made by the Asp in finger 3 contributes to the recognition of the TA base pair at the 3' end of the finger 2 subsite, as finger 2 does not make a specific contact to this base pair. But as Asp can recognize both A and C the specificity of this contact is likely to be "degenerate" (Elrod-Erickson *et al*, 1996).

The structure also contains a number of water mediated contacts (Elrod-Erickson *et al.*, 1996). Such contacts may be important where no hydrogen bond is formed between an amino acid and a base, but where a definite preference for a specific base exists. They are also important in the overall contribution to the affinity of the protein for the DNA. (Schwabe, 1997).

On binding of Zif268 to DNA the conformation of the protein itself is not greatly changed as the linker allows the protein to wrap around the DNA. There is however a change in the conformation of the DNA from B-DNA to $B_{\text{enlarged groove}}$-DNA. Modelling studies have revealed that the distortion of the DNA is largely due to the linker length. Shortening of the linker would probably allow the protein to bind to B-DNA (Elrod-Erickson *et al.*, 1996). It is not necessarily the case that on the binding of zinc finger proteins to the DNA the conformation will change to $B_{\text{enlarged groove}}$-DNA. Conformations of DNA which are close to both A-DNA and B-DNA have been observed in zinc finger / DNA complexes (Nekludova and Pabo 1994).

## 1.3 ZINC FINGER DNA CODE OF INTERACTIONS, DESIGN AND USES OF ZINC FINGER PROTEINS.

### 1.3.1 Present Approaches to the Elucidation of a Zinc Finger / DNA interaction code.

The amino acids at positions α, β and γ in a $Cys_2His_2$ zinc finger determine the base triplet to which a zinc finger domain will bind. It is known for example that if α= Arg, β= His, γ= Arg the domain will preferentially bind to the sequence 5'G (A/G) G3' (Desjarlais and Berg, 1994). With the possibility of 20 different amino acids at α, β and γ there are 8000 different zinc fingers, of which each might bind to one or more of the 64 possible DNA target sequences. It has been recognized for some time that a full code of combinations of amino acids at α, β and γ in relation to target sequences in DNA would be useful (Desjarlais and Berg 1992b). Two approaches are presently employed in an attempt to decipher the code for zinc finger / DNA interactions; phage display and length encoding.

Berg and co-workers utilize a three zinc finger DNA binding protein in which the two outer zinc fingers are fixed and the central one is altered at the critical base contacting residues. These proteins are then individually screened with a mixture of the 64 target sequences, which are all of different lengths. The target sequences to which the zinc fingers have bound are identified by a two stage process. Involving gel shift analysis followed by electrophoresis of the bound target sequences, in order to identify the lengths of the DNA the protein has preferentially bound to. In this way the preferential target sequences of a given zinc finger may be identified (Desjarlais and Berg, 1994).

Several groups use phage display methods for the selection of zinc finger proteins that bind to specific DNA sequences. This approach requires libraries of zinc finger constructs to be made and displayed on phage coats. Pabo and coworkers simply screen zinc finger proteins displayed on the coats of phage for ability to bind specific DNA sequences (Rebar and Pabo, 1994; Rebar *et al.*, 1996).

Klug and coworkers have taken this approach further synthesizing libraries of proteins containing three zinc finger motifs in which the central motif contains randomized amino acids at the positions responsible for DNA recognition. The libraries

of zinc finger genes are fused to a phage coat gene and therefore the proteins displayed on the coats of the phage. Zinc finger phage with the ability to bind to a given DNA sequence are isolated affinity via purification of libraries of phage using a specified immobilized DNA sequence (Choo and Klug, 1994a). This selects one or several zinc finger phage capable of binding to a given DNA sequence.

In order to identify the optimal binding site for a given zinc finger protein, a DNA binding assay is then performed. This demonstrates whether or not the zinc finger protein selected by the previous method is the most suitable for binding the specified sequence. Libraries of target DNA sequences fixed at one base and randomized at the other two within the recognition subsite of the central zinc finger are fixed to microtiter plates. The zinc finger phage are then allowed to bind to the DNA. The plates are washed to remove unbound phage and then bound phage is detected using an immunological assay. The favoured binding sites of the protein can then be predicted by which of the groups of DNA libraries "light up" (Choo and Klug, 1994b).

As the structure of more zinc finger domains bound to DNA have been solved it has become obvious that zinc fingers in naturally occurring proteins such as Tramtrack (Fairall *et al.*, 1993) GL1 (Pavletich and Pabo, 1993) and TFIIIA (Nolte *et al.*, 1998) do not always follow the simple recognition rules exhibited by Zif268 (Pavletich and Pabo, 1991; Elrod-Erickson *et al.*, 1996). In some structures the amino acid residue immediately to the amino side of DNA recognition residue β (Fig 1.3) can play an important role in DNA recognition (Isalan *et al.*, 1997). Irregularities in DNA recognition by naturally occurring zinc finger proteins have cast doubt on whether a general code could be used to predict the function of existing proteins (Elrod-Erickson *et al.*, 1998). Consensus based zinc finger proteins however seemed to follow the same basic rules of DNA recognition exhibited by Zif268 (Kim and Berg, 1996). Therefore it should be possible to use a consensus zinc finger frame along with a code for interactions for the design of proteins to recognize specific DNA sequences.

**1.3.2 Design of Zinc Finger Peptides to interact with specific sequences.**

Various methods are presently employed for the design of proteins which specifically bind to given DNA sequences. These methods can be divided into two groups: those which screen libraries of proteins in order to identify a protein with a particular binding specificity and those which involve modelling or data base analysis to design a protein with the required DNA specificity. The techniques used for deciphering a code of interactions discussed above (1.3.2) are also used to screen for proteins with given DNA binding specificities. Modelling techniques and sequence comparison have also been used with some degree of success (Nardelli *et al.*, 1991; Rebar and Pabo, 1994; Choo *et al.*, 1994; Jamieson *et al.*, 1996). Zinc finger motifs with known recognition sites can then be "mixed and matched" to recognize DNA sequences larger that 3bp (Desjarlais and Berg, 1993; Choo *et al.*, 1994).

These techniques have achieved some success but are based on the assumption that the zinc fingers act as completely independent DNA binding modules. Mutagenesis studies have shown that the amino acid sequence of one zinc finger motif in a protein can effect the DNA binding specificity of neighboring motifs (Isalan *et al.*, 1997). The x-ray crystal structures (Fairall *et al.*, 1993; Pavletich and Pabo, 1993; Nolte *et al.*, 1998; Pavletich and Pabo, 1991; Elrod-Erickson *et al.*, 1996) also show that there is a certain amount of "cross talk" between one finger and another finger's binding subsite. It unlikely to be possible to generate a truly optimized zinc finger protein / DNA interaction from prediction based on knowledge of other interactions acquired from study of isolated, single, finger domains (Elrod-Erickson *et al.*, 1998). If an "optimal" protein is to be designed it is necessary to consider the effect of neighbouring fingers whilst optimizing the binding of the motif in question. A technique that adds and optimizes one zinc finger motif at a time allowing the DNA binding protein to "grow" over the target DNA sequence has been developed (Greisman and Pabo, 1997).

Alterations to areas of the protein not directly involved with base recognition may also effect the binding of a designed protein to DNA. A zinc finger protein with a femtomolar dissociation constant has been constructed in which the three zinc fingers of Zif268 were fused to a designed zinc finger protein. The lengthening of the linker sequence between motifs in this protein was show to contribute to an increased affinity for the DNA (Kim and Pabo, 1998). The study of individual parts of zinc finger proteins

in isolation demonstrates that in order to optimize the binding of an entire protein to DNA there may be many more factors to take into consideration than the DNA contacting residues.

### 1.3.3 Potential importance of designed zinc finger peptides.

The ability to design zinc finger proteins to recognize specific DNA sequences has applications in research, biotechnology and medicine. There are two ways in which zinc finger proteins may be used in these fields. Firstly zinc finger proteins alone may be utilized solely for their DNA-binding capacity. Also zinc finger proteins may be linked to functional domains of other proteins in order to create proteins with novel combinations of sequence specificity and function.

Zinc finger proteins alone are used largely for the repression of gene expression. This may be achieved either by selecting a protein which specifically recognizes a sequence within the gene of interest, or by the insertion of the DNA target sequence of a known zinc finger protein into the gene of interest (Kim and Pabo, 1997). Specific repression of genetic expression by zinc finger proteins may have medical applications in the treatment of diseases, both in the repression of aberrant expression of genes and in the inactivation of viruses. Some examples of such uses of zinc finger proteins are discussed below.

In many cases cancer is caused by the activation of an oncogene. The specific repression of an oncogene using a zinc finger based protein has been demonstrated in *vitro*. A zinc finger peptide was selected which was capable of specifically recognising a region in the BCR-ABL fusion gene (Philadelphia chromosome). Cell lines containing the Philadelphia chromosome were shown to revert to growth factor dependence in the presence of the zinc finger protein, indicating that such proteins may provide effective therapies for cancer caused by the activation of oncogenes (Choo *et al.*, 1994).

It has been demonstrated that zinc finger proteins could be designed which may combat viral diseases, this was demonstrated by the anti-influenza activity of a zinc finger peptide. A peptide containing one zinc finger motif, which is part of an influenza structural protein, was synthesized. This peptide was capable of inhibiting the viral polymerase and was therefore able to repress viral replication in mice. This seemed to improve the survival of mice treated with the peptide (Judd *et al*, 1997). This treatment

may eventually be used in humans as a therapy for viral infections such as the common cold or even HIV.

Zinc finger proteins can also be produced with novel combinations of sequence specificity and function. Uses for such proteins include both the development of novel sequence specific enzymes and the regulation of genetic expression.

The isolation of restriction enzymes capable of recognizing novel sequences is lengthy and expensive, a structure based design approach would be more practical. It has been shown that it is possible to modify the recognition sequence of the restriction enzyme *fok* I using zinc fingers (Kim *et al,* 1996).

Several attempts have been made to make sequence specific transcription factors. Structure based design has been successfully used to fuse zinc fingers 1 and 2 from Zif268 to the homeodomain of Oct-1 (Pomerantz *et al*, 1995). A zinc finger protein based on Zif268 capable of binding to a given DNA sequence has been selected and fused to the activation domain from VP16 (Choo *et al.*, 1997). Also a zinc finger peptide has been fused to the TATA box binding protein (Kim *et al.*, 1997). All approaches yielded transcription factors with functions that are reasonably predictable. The potential uses of such proteins are discussed below.

Zinc finger domains combined with functional domains of proteins capable of controlling expression can be used as a tool to manipulate the expression of genes. For example it would be useful in the deciphering of developmental pathways to be able to regulate genes of interest. Unfortunately, naturally occurring regulation factors will effect the expression of several genes (Kel *et al*, 1995). Zinc finger based proteins may be designed that are targeted at a more specific sequence within the gene of interest (Choo *et al*, 1997).

It would be useful to be able to control the expression of genes in genetically modified organisms. The expression of a protein of interest in bacteria, plants or animals may be harmful to the organism. The ability to control the expression of the proteins of interest in bacteria has already been shown to be important.

There are already systems available for the controlled expression of genes in bacteria but many of them rely on the lac system which is slightly "leaky", and causes the induction of host genes as well as the gene of interest. Although expression systems

available such as those utilizing phage DNA polymerases are adequate for most uses systems with increased specificity and control would still be useful.

Control of genes in plants and animals tends to rely on naturally occurring systems which are nonspecific and effect endogenous genes as well as those of interest (Kel *et al*, 1995). Systems for the specific expression of genes in plants have been developed which rely on an environmental stimulus such as a chemical (Caddick *et al.*, 1998). It is unlikely however that such a system could be used for the control of gene expression in an animal.

The design of zinc fingers for the control of expression of genes within organisms would allow the development of "non-leaky", and specific expression systems. It is also likely that zinc finger proteins could be used to control the level of the expression of exogenous genes. This would be an important advancement that would in particular be useful in gene therapy (Kim and Pabo, 1997).

## 1.4 AIMS AND OBJECTIVES.

The aim of this project was to synthesize a zinc finger protein suitable for the study of DNA protein interactions and to develop a high throughput assay for the detection of such interactions.

The zinc finger protein must be capable of binding with high affinity and specificity to a known DNA sequence. It should be possible for future study to alter the gene encoding the zinc finger, in order to determine the effect of the alteration of the DNA contacting residues on DNA binding.

It should be possible to synthesize large quantities of the zinc finger protein by the controlled expression of the gene in *E. coli*.

The interaction between the zinc finger protein should be characterized as a model system which can then be used to develop a high throughput assay to detect the interaction between the zinc finger proteins and potential ligands. The ideal assay format might be both dynamic and "real time" such as Scintillation Proximity Assay (SPA) (Amersham Pharmacia Biotech).

# 2. MATERIALS AND METHODS.

## 2.1. RECIPES FOR BUFFERS.

All water was de-ionised using an Elgastat option 4 water purifier (Elga). Water and solutions for use with DNA were sterilised by autoclaving at 121°C, 15 p.s.i. for 15 mins. In the following recipes, $Na_2EDTA$ pH8.0 is abbreviated to EDTA.

| Ref. No. | Buffer | Tris-HCl | NaCl | EDTA | Glycine | Other. |
|---|---|---|---|---|---|---|
| 1 | 1x TAE | 400mM pH 8.0 | | 1mM | | 1.142% Glacial acetic acid |
| 2 | 1x TBE | 87mM pH 8.0 | | 2mM | | 87mM Boric acid |
| 3 | 1x Glycerol tolerant sequencing buffer | 87mM pH 8.0 | | 2mM | | 3.6% Taurine. |
| 4 | 10x SDS-PAGE | 250mM (no HCl) | | | 1.92M | 1% SDS |
| 5 | TE pH8.0 | 10mM pH 8.0 | | 1mM | | |
| 6 | STE | 10mM pH 7.5 | 100mM | 1mM | | |
| 7 | TBS | 10mM pH 8.0 | 150mM | | | |
| 8 | PBS pH7.4 | | 150mM | | | 4mM $NaH_2PO_4$ 16mM $Na_2HPO_4$ |
| 9 | Glucose | 25mM pH 8.0 | | 10mM | | 50mM Glucose |
| 10 | 20x SSC | 3M | | | | 0.3M $Na_3$Citrate |

| Ref. No. | Loading buffer | Xylene cyanol | Bromophenol blue | Other |
|---|---|---|---|---|
| 11 | Agarose | 0.025% | 0.025% | 15% Ficoll-400 |
| 12 | Formamide | 0.025% | 0.025% | 95% de-ionised formamide |
| 13 | SDS-PAGE | 0.0125% | | 6.25% SDS, 250mM Tris-HCl pH6.9, 31.25% (v/v) glycerol, 312.5mM DTT |
| 14 | Sequencing | 0.05% | 0.05% | 95% Formamide, 20mM $Na_2EDTA$ pH8.0 |

### 2.1.15 1.0M TEAB

Triethylamine (13.93ml) was diluted to 100ml in $H_2O$. The pH of the solution was adjusted to 7.4 by bubbling through $CO_2$.

### 2.1.16 Protein purification buffers (Van Dyke *et al.*, 1992).

2x BC; 20mM Tris-HCl pH7.9
20% Glycerol
0.2mM EDTA
10mM Mercaptoethanol
0.5mM PMSF

BC 100/40; 1x BC
100mM KCl
40mM Imidazole

BC 1000/100; 1x BC
1M KCl
100mM Imidazole

BC 1000/2000;1x BC

1M KCl

2M Imidazole

**2.1.17 Solutions for developing antibody blots.**

NBT solution - 1ml 70% DMF containing 30mg NBT.

BCIT solution -1ml DMF containing 15mg BCIT.

Carbonate Buffer - 0.1M Na $HCO_3$, 1mM $MgCl_2$ pH adjusted to 9.8 with NaOH.

2.2 RECIPES FOR MEDIA.

**2.2.1 LB Medium.**

Tryptone (Difco) (1%), 0.5% Yeast extract (Difco) and 1% NaCl dissolved in de-ionised $H_2O$ and the pH adjusted to 7.0 with NaOH. For LB agar 1.6% agar (Oxiod) was also added. All media were sterilised by autoclaving at 121°C, 15 p.s.i. for 15 mins.

**2.2.2 M9 Salts;**

| | |
|---|---|
| 0.230M | $NaHPO_4 \cdot 7H_2O$ |
| 0.044M | $KH_2PO_4$ |
| 0.017M | NaCl |
| 0.037M | $NH_4Cl$ |

**2.2.3 Minimal Medium.**

Sterile $H_2O$ or $H_2O$ agar (containing 1.6% agar with respect to the final volume) was combined with sterile M9 salts (20%), 10% glycerol (2.5%), 200mM $MgSO_4$ 10mM $CaCl_2$ (1%) and Thiamine (0.1%).

## 2.3 BACTERIAL STRAINS.

| | |
|---|---|
| *E. coli* TG2: | *supE hsd Δ thi Δ (lac-proAB) Δ (srl-recA)*306 : Tn10(*tet*$^r$) F'[*tra*D36 *pro*AB$^+$ *lacI*$^q$ *lacZ*ΔM15] |
| *E. coli* HB101: | *supE*44 *hsd*S20($r_B^- m_B^-$) *rec*A13 *ara*-14 *pro*A2 *lac*Y1 *gal*K2*rsp*L20 *xyl*-5 *mtl*-1 |
| *E. coli* DH5α: | *supE*44 (ϕ80 *lac* ZΔM15) *hsd* R17 *recA1 end* A1 *gyr* A96 *thi*-1 *rel* A1 |
| *E. coli* HMS 174 (DE3): | *rec*A1 *hsd*R *rif*$^r$ |
| *E. coli* BL21(DE3): | *rec*A1 *hsd*R *rif*$^r$ S *gal* (λcI*ts*857 *ind*1 Sam7 *nin*5 gene 1)*lac*UV5-T7 |

## 2.4 GENERAL TECHNIQUES.

### 2.4.1 Ethanol precipitation.

A one tenth volume of 3M tri-sodium acetate pH5.5 and 2 volumes of ice-cold ethanol were mixed with the DNA by vortexing. The mixture was placed at -20°C for 20 mins and the centrifuged in a microfuge at 13000 rpm for 15 mins. The ethanol was removed using an aspirator and the pellet dried at room temperature for 10 mins.

### 2.4.2 Phenol / chloroform extraction.

One volume of phenol / chloroform (1:1 v/v) was added to the DNA and mixed by vortexing. The mixture was centrifuged in a microfuge at room temperature for 5 mins. The resulting upper (aqueous) layer was removed and the lower organic layer discarded. One volume of chloroform was added to the DNA and the mixture was vortexed. The mixture was centrifuged and the aqueous layer (containing the DNA) removed. "Chloroform" refers to a 24:1 (v/v) mixture of chloroform and isoamyl alcohol.

### 2.4.3 Dialysis.

Visking (MWCO: 12,000-14,000) dialysis tubing (Fisher Scientific) was boiled for 10 mins in a solution containing 2% $NaHCO_3$ and 1M $Na_2EDTA$ pH8.0. The tubing was rinsed in de-ionised $H_2O$, boiled as above, allowed to cool, rinsed, and stored in solution at 4°C. Spectra/Por (MWCO: 1000) dialysis tubing (Spectrum) was pre-soaked in de-ionised $H_2O$. Before use the tubing was rinsed thoroughly in de-ionised $H_2O$ and clipped at one end. DNA or protein was placed inside the tubing and the open end clipped allowing room for expansion of the liquid inside. The tubing was then stirred in an appropriate buffer at 4°C.

### 2.4.4 Autoradiography.

Gels were dried onto chromatography paper (1 Chr, Whatman) using a gel drier (model 583 Bio-Rad) under vacuum at 80°C for 2 hrs, then exposed to Kodak Biomax MR film (Amersham). Membranes were enclosed in Saran wrap, or dried in an oven at 80°C, and exposed to Kodak X-Omat AR film (Amersham). All films were developed according to the manufacturer's instructions.

### 2.4.5 Scintillation counting.

Scitillation counting of radio-labeled compounds was achived by placing an aliquot of the sample into scintillation "cocktail" (OptiPhase Hisafe II). The addition of scintillant to SPA was not required. Programs for scintillation counting on a Packard 1600 TR liquid scintilation analyser are given below.

| Sample | Region | Lower limit | Upper limit | Count time |
|---|---|---|---|---|
| $^{32}P$ | A | 5.0 | 1700.0 | 1min |
| | B | 50.0 | 1700.0 | |
| $^{33}P$ | A | 0.0 | 2000.0 | 1min |
| | B | 0.0 | 2000.0 | |
| $^{35}S$ | A | 0.0 | 167.0 | 1min |
| | B | 2.0 | 18.6 | |
| $^{3}H$ | A | 0.0 | 18.6 | 4min |
| | B | 2.0 | 18.6 | |
| SPA | A | 0.0 | 2000 | 1min |
| | B | 0.0 | 2000 | |

## 2.5 GEL ELECTROPHORESIS.

### 2.5.1 Agarose gel electrophoresis.

Gels were made from molecular biology grade agarose (0.8-4.0% agarose, Gibco BRL), 1xTAE (2.1.2) and 1μg/ml Ethidium bromide. Gels for the purification of DNA fragments were made from SeaPlaque agarose (1.0% agarose, FMC Bio products), 1xTAE and 1μg/ml Ethidium bromide. DNA was mixed with one fifth volume of agarose loading buffer (2.1.11) and loaded onto the gel. All gels were electrophoresed in 1x TAE (2.1.1) at 10 volts/cm length of gel for one hour. Gels were viewed on a transilluminator and images captured using a UVP system. Pictures were printed on a digital graphic printer (UP-D890, Sony).

**2.5.2 Sequencing gels.**

Sequencing plates (Bio-Rad Sequi-Gen sequencing cell) were cleaned with detergent, rinsed thoroughly with $H_2O$ and wiped with ethanol. The bottom plate (without the buffer well) was treated with Dimethydichorosilane solution, then polished with ethanol. The spacers were glued to this plate which was then supported horizontally. TBE gels were purchased pre-mixed (Severn Biotech Ltd.). Glycerol tolerant gels (50ml) containing 6% Acrylamide (Severn Biotech Ltd.) 21g Urea and 1x Glycerol tolerant buffer (2.1.3) were freshly mixed. TEMED (25µl) and 0.25ml of 10% AMPS solution were then added to 50ml of sequencing gel mixture. The gel was then introduced between the plates using a 50ml syringe and gradual sliding of the top plate over the bottom plate. The plates were secured with large clips for 30 mins whilst the gel was allowed to polymerise. The buffer 1x TBE (2.1.2) or 1x glycerol tolerant buffer (2.1.3) as appropriate, was pre-warmed to 60°C in a microwave oven. The apparatus was then assembled and the gel electrophoresed according to the manufacturer's instructions.

**2.5.3 Polyacrylamide gels for protein analysis.**

Plates (Bio-Rad Mini protean II) were cleaned and assembled according to the manufacturer's instructions. A 14.25ml resolving gel was made containing acrlyamide diluted to an appropriate concentration from a 30% stock solution (Severn Biotec Ltd.), 1.88ml 3M Tris-HCl pH8.8 and 0.15ml 10% SDS. Immediately prior to introducing the gel between the plates 7.5µl TEMED and 75µl 10% AMPS solution was added. The gel was overlaid with $H_2O$ and allowed to polymerise for 30 mins.

A 4.75ml stacking gel mix was made, containing 0.625ml 30% Acrylamide solution, 1.25ml 0.5M Tris-HCl pH6.8 and 50µl 10% SDS. The water overlay was removed from the polymerised resolving gel. TEMED (3.75µl) and 35µl 10% APS were added to the stacking gel mix which was then poured between the plates over the resolving gel. A comb was inserted and the gel allowed to polymerise for 30 mins.

Samples were resuspended in, or combined with a one tenth volume of SDS-PAGE loading buffer (2.1.13) and heated to 95°C for 5 mins immediately before loading. Markers were combined with 5µl SDS-PAGE loading buffer and heated to 65°C for 5min prior to loading. Electrophoresis was carried out in 1x SDS-PAGE buffer (2.1.4) at 100 volts for 2.5-3 hrs. Gels were then removed from the apparatus and stained with coomassie blue (2.5.4).

**2.5.4 Coomassie Blue staining.**

Gels were soaked overnight in a solution containing 50% methanol, 7.5% acetic acid and 0.12% coomassie blue. They were then soaked in a solution containing 5% methanol, and 0.75% acetic acid with agitation until the protein bands could be clearly viewed. Images of the gels were captured using a UVP system. Pictures were printed on a digital graphic printer (UP-D890, Sony).

**2.5.5 Gel Retardation Assay.**

Plates (Bio-Rad Protean II xi Cell) were cleaned and assembled according to the manufacturer's instructions. A gel mixture (50ml) containing 7% Acrylamide diluted from a 30% stock solution (Severn Biotech Ltd.), 0.5x TBE (2.1.2), 0.25ml 10% APS solution and 25µl TEMED was poured between the plates. A comb was then inserted and the gel allowed to polymerise for 30 mins.

DNA and protein samples were combined in a buffer containing 25mM HEPES pH7.5 and 10% glycerol. Binding was allowed to take place at room temperature for 30 mins. The gel apparatus was assembled and pre-electrophoresed at 200 volts for 1 hr in 0.5x TBE buffer (2.1.2).

Samples were loaded directly onto the gel and electrophoresed at 150-200 volts for 2.5-3 hrs at 4°C. Gels were removed from the apparatus and prepared for auto-radiography (2.4.4).

## 2.6 OLIGONUCLEOTIDES.

### 2.6.1 Preparation of oligonucleotides.

Oligonucleotides were synthesised (Oligo 1000 DNA synthesiser, Beckman) according to the manufacturer's instructions then removed from the synthesis column using an ultra fast cleavage and deprotection kit (Beckman). Samples were lyophilised by centrifugation under vacuum (Speed vac SVC100, Savant). Those to be used immediately were resuspended in 200µl of sterile de-ionised $H_2O$. Oligonucleotides requiring purification were resuspended in 200µl of Formamide loading buffer (2.1.12).

The apparatus and method of assembly used for purification of oligonucleotides was the same as that used for gel retardation assays (2.5.5), except that the gel mixture (50ml) contained 21g urea and 1x TBE (2.1.2) and was combined with 25µl TEMED and 0.25ml 10% AMPS. A comb was inserted and the gel allowed to polymerise for 30 mins.

Samples were heated to 95% for 5 mins prior to loading on the gel. The gel was electophoresed at 35mA for 1.5-2 hrs, then removed from the apparatus and placed on Saran wrap. The DNA bands were visualised by placing the gel on a TLC plate and illuminating with an overhead U.V. source. The band corresponding to the largest product was removed from the gel using a scalpel. The gel slice was cut into small strips and placed in a sterile 1.5ml tube. The DNA was eluted from the gel by soaking overnight at 4°C in 1ml of 0.1M TEAB (2.1.15).

A 0.2ml column (Promega Mini prep column) containing DE52 (Whatman) pre-equilibrated in 0.1M TEAB (2.1.15) was prepared and washed with 1ml of 0.1M TEAB. The elutate from the gel slice was applied to the column using a 2ml syringe. The column was the washed with 2-3ml of 0.1M TEAB. The oligonucleotide was then eluted from the column with 1ml 1.0M TEAB. The elutate was lyophilised by centrifugation under vacuum (Speed vac SVC100, Savant).

Measuring the absorbance of an aliquot of the oligonucleotide at both 260nm and 280nm assessed the concentration and purity of the samples. The concentration of the oligonucleotide was estimated by assuming that a solution containing 3.3μg/ml oligonucleotide would have an absorbance of 1 at 260nm. The purity was demonstrated that the absorbance at 280nm divided by that at 260nm was close to 1.8 (Sambrook *et al.*, 1989).

### 2.6.2 Annealing of oligonucleotides.

Unless otherwise stated complementary oligonucleotides were combined at 1pmol/μl and placed in a thermo-cycler (Hybaid OmniGene) at 95°C for 5min. The oligonucleotides were then cooled slowly to room temperature. This was achieved by placing the tubes containing the oligonucleotides in a beaker of water heated to 95°C which was allowed to cool on the bench for 2hrs.

## 2.7 CELL CULTURE.

### 2.7.1 Preparation and transformation of competent cells.

A single colony was used to inoculate a small volume of LB (2.2.1), containing 35μg/ml choramphenicol when using pLysS strains, and incubated overnight in a 37°C shaking incubator. LB (30ml), containing 35μg/ml choramphenicol when using pLysS strains, was then inoculated with 0.3ml of the overnight culture and incubated at 37°C for 1.5-2 hrs until the absorbance at 600nm was between 0.4 and 0.6. The cells were centrifuged (Beckman J2-21 centrifuge, JA20 rotor) at 6,000 r.p.m, for 5 mins at 4°C. The growth medium was discarded and the pellet re-suspended in 6ml ice cold, sterile 50mM $CaCl_2$ solution. The cells were placed on ice for 5 mins before centrifugation and re-suspension were repeated. The cells were placed on ice for a further 5 mins, the centrifugation was repeated the $CaCl_2$ discarded and the cells resuspended in 1.2ml of ice cold 50mM $CaCl_2$. The suspension was then incubated on ice for 20 mins.

Transformation: 100μl of the cells were combined with the DNA of interest. The mixtures were incubated on ice for 30 mins. The cells were placed at 37°C for 30

seconds and returned to ice for 2 mins. LB (0.5ml) was added to the transformed cells and they were incubated at 37°C for 40 mins.

When transforming pUC19 constructs, 200µl of transformed cells were then combined with 50µl 2% X-gal and 10µl 1M IPTG. The transformants (200µl) were spread on LB plates containing 50µg/ml ampicillin (and 35µg/ml chloramphenicol, when using strains containing the pLysS plasmid). To ensure that *E. coli* strains used were not ampicillin resistant, cell viability was tested by spreading 200µl of the competent cells onto an LB plate containing ampicillin. Plates were incubated overnight at 37°C

## 2.8 DNA PREPARATIVE METHODS.

Plasmid preparations were routinely performed using Wizard (Promega) or Qiagen preparation kits according to the manufacturer's instructions.

### 2.8.1 Large scale preparations of plasmid DNA (Sambrook *et al.*, 1989)

LB (2.3.1) supplemented with 50µg/ml ampicillin (and 50µg/ml chloramphenicol where strains contain the pLysS plasmid were used), was inoculated with a single colony and incubated overnight with shaking at 37°C.

Cells were harvested by centrifugation at 6,000 r.p.m for 10 mins, at 4°C (Beckman J2-21 centrifuge, JA20 rotor). Cells were resuspended in 100ml of ice cold STE (2.1.6), and the centrifugation was repeated then the STE discarded. The bacterial pellet was resuspended in 10ml of glucose buffer (2.1.9). Lysozyme solution (10mg enzyme in 1ml 10mM Tris-HCl pH 8.0) and 20ml of 0.2M NaOH, 1% SDS were added to the cell suspension. The cells were allowed to lyse at room temperature for 10 mins. Ice cold 5M potassium acetate pH4.8 (15ml) was added and mixed by inversion. The mixture was placed on ice for 10 mins and centrifuged at 4,500 r.p.m, for 15 mins at 4°C (Beckman J2-21 centrifuge, JA20 rotor). The supernatant was decanted and filtered, 0.6 volumes of isopropanol was added and the DNA allowed to precipitate at room temperature for 20 mins. DNA was recovered by centrifugation at 10,000 r.p.m.

for 15 mins at room temperature (Beckman J2-21 centrifuge, JA20 rotor). The DNA pellet was washed with 70% ethanol, the tube inverted and the pellet allowed to dry at room temperature for 30 mins. The pellet was then resuspended in 4ml of sterile $H_2O$.

A sample of CsCl was dried for 30 mins at 65°C and an equivalent weight dissolved in the DNA solution. Ethidium bromide (0.5ml of 10mg/ml) was added and mixed by inversion, then centrifuged at 8,000 r.p.m. for 5 mins at room temperature (Beckman J2-21 centrifuge). The supernatant was loaded into a quick seal polyallomer centrifuge tube (capacity 3.9ml, Beckman), the samples topped up with 1mg/ml CsCl solution, balanced using weighing scales and sealed according to the manufacturer's instructions. The tubes were centrifuged at 100,0000 r.p.m. for 16 hrs at 15°C (Beckman TL-100 Ultra-centrifuge). The tubes were carefully removed from the rotor and the band visualised with short wavelength U.V. radiation. The lower band was removed into a sterile tube using a 19-gauge hypodermic needle and 1ml syringe. The DNA was shaken with one volume of CsCl/$H_2O$ saturated *n*-butanol and the lower phase removed. Extraction was repeated until the DNA no longer fluoresced under U.V. illumination.

CsCl was removed from the DNA by dialysis in visking tubing (2.4.3) against 3x 1l TE pH8.0 (2.1.5) for 2x 1 hr followed by overnight dialysis. The DNA was removed from the dialysis tubing, ethanol precipitated (2.4.1) and resuspended in 1ml sterile $H_2O$. Concentration and purity of the DNA was determined by measuring the absorbance of an aliquot of it at 260 and 280nm.

## 2.9 ENZYMATIC TECHNIQUES.

Unless otherwise stated, all enzymatic techniques were performed according to the manufacturer's instructions. Where one enzymatic technique is followed by a second requiring a different reaction buffer the DNA was ethanol precipitated (2.4.1) and re-suspended in the buffer required for the following manipulation. Unless otherwise stated, all DNA sequencing was carried out on both strands of DNA.

### 2.9.1 Polymerase chain reaction (PCR)

PCRs were carried out in 1x PCR buffer (Perkin Elmer) with, 10μM dNTPs (Perkin Elmer) and 1μM oligonucleotide primers. Half a unit of Taq polymerase (Perkin Elmer) was used per 10μl final volume. The final volume of each reaction unless otherwise stated was 20μl. In reactions used to amplify purified DNA, 0.5-2μl of the DNA was used. When PCR was used to screen bacterial colonies directly, a small amount of cells was transferred to the reaction tube using a pipette tip. The tip was then used to inoculate a gridded, numbered LB ampicillin plate in order to allow clones to be retrieved later. Within each set of PCRs a negative control was set up containing no template DNA. Reactions were overlaid with 20-30μl of mineral oil (molecular biology grade, Sigma). Tubes were placed in a thermo-cycler (Hybaid OmniGene) and the following program executed;

Denaturing- 95°C for 30 secs,

Annealing- 50°C for 30 secs,

Elongation- 72°C for 1 min,

Cycle repeated 30 times.

PCR products were then examined by electrophoresis on a 2% agarose gel (2.5.1).

### 2.9.2 "Blunt ending" reaction.

PCR product (40μl) was placed in a 0.5ml sterile tube and mixed with;

| | |
|---|---|
| 6μl | 20mM ATP |
| 1μl/10 units | T4 polynucletide kinase (MBI Fermentas) |
| 1μl/10units | DNA polymerase I large fragment (Klenow, New England Biolabs) |
| 1μl | PCR buffer (Perkin Elmer) |
| 1μl | Sterile $H_2O$ |

The mixture was incubated at 37°C for 30 mins, then the enzymes inactivated at 95°C for 5 mins.

**2.9.3 Ligation.**

Ligations were performed in 1x ligase buffer (MBI Fermentas) with 1mM ATP, 5 units of T4 DNA ligase, approximately 20ng of plasmid DNA and (unless otherwise stated) a 10 fold molar excess of insert. Reaction were carried out in a 20µl volume, and incubated overnight at 16°C.

**2.9.4 Agarose gel purification.**

DNA was electrophoresed in a 1% SeaPlaque gel (2.6.1). A gel slice containing the fragment of interest were removed using a scalpel and placed in a sterile 0.5ml tube. The gel was then digested using Beta-Agarase (New England Biolabs) according to the manufacturer's instructions. The recovered DNA was re-suspended in sterile $H_2O$. To estimate the concentration and purity of the DNA, an aliquot was electrophoresed on an agarose gel (2.5.1).

## 2.10 HYBRIDISATION.

**2.10.1 Preparation of colony blots.**

A grid was marked on a piece of nitro-cellulose membrane (Bio-Rad) in pencil. The membrane was placed on a LB plate containing 50µg/ml ampicillin. Divisions were inoculated with a small quantity of cells from a single colony, which were also inoculated onto a similarly marked LB ampicillin plate, allowing clones to be retrieved. The plates were incubated overnight at 37°C. The membranes were removed from the plates and floated on 1.5M NaCl, 0.5M NaOH for 30 secs, then submerged in this solution for 1 min. The membranes were then removed and submerged in 1.5M NaCl, 0.5M Tris-HCl pH8.0 for 5 mins. They were washed in 2x SSC (2.1.10) and air-dried. The membrane was then subjected to hybridisation (2.10.2).

**2.10.2 Membrane hybridisation.**

Twenty pmol of oligonucleotide was incubated at 37°C for 30 mins in a 20µl reaction containing 1x kinase buffer (MBI Fermentas), 10µCi [$\gamma$-$^{32}$P]ATP and 0.5 units of T4 polynucleotide kinase (MBI Fermentas). Hybridisation was carried out in a hybridisation oven (HB1D hydridiser, Techne) at 65°C. Membranes were pre-hybridised for 4 hrs in a solution containing 0.5M sodium phosphate pH7.2, 7% SDS, 1mM EDTA (0.35ml/cm$^2$ membrane). The solution was replaced (0.17ml/cm$^2$ membrane), the labelled probe was denatured at 95°C for 5 mins and added. Membranes were hybridised overnight, washed twice in 2x SSC (2.1.10) for 30 mins once in 2x SSC, 0.1% SDS for 15 mins and once in 0.1x SSC for 10 mins. Membranes were then autoradiographed (2.4.4).

## 2.11 PROTEIN ANALYSIS.

**2.11.1 Protein expression.**

A single colony (freshly transformed, 2.7.1) was inoculated into a small quantity of LB supplemented with 50µg/ml ampicillin and when using strains containing the pLysS plasmid, 35µg/ml chloramphenicol. The culture was incubated overnight with shaking at 37°C. The culture was diluted 100 fold into fresh media and incubated for 3-4 hrs until an absorbance of 0.6-0.65 at 600nm was reached. IPTG was then added to the culture to a final concentration of 1mM. When expressing genes in pT7-7 vectors rifampicin was added to a final concentration of 35µg/ml, five mins after induction. The culture was then incubated at 37°C for a further 3 hrs.

The cells were harvested by centrifugation at 5,000 r.p.m. for 5 mins at 4°C (In a microfuge for samples less than 1ml, and in a Beckman J2-21 (JA20 rotor) for larger samples). Samples to be analysed by SDS-PAGE gel electrophoresis (2.5.3) were re-suspended in 30µl of SDS-PAGE loading buffer (2.1.13). Samples for purification were resuspended in $H_2O$ containing 0.25mM PMSF when preparing His tagged protein, or

PBS (supplied with the GST purification kit (Amersham Pharmacia Biotech) when purifying GST protein. Cell were resuspended in 10ml of the appropriate buffer/100ml original culture size. Cells were lysed by several cycles of freezing in liquid nitrogen and thawing at 37°C. Cellular debris was removed by centrifugation at 15,000 r.p.m for 10 mins at room temperature (Beckman J2-21 centrifuge, JA20 rotor). The cleared lysate was stored at -70°C.

**2.11.2 Column chromatography of zinc finger proteins.**

A column (Wright scientific, 1cm diameter/15cm length) was rinsed with de-ionised water and filled with either Chelex-100 metal affinity resin (Bio-Rad) suspended in 0.5M $NiCl_2$ or DEAE-Sepharose (Amersham Pharmacia Biotec). The resin was allowed to settle under gravity. Columns were eluted at 0.5ml/min using a peristaltic pump (P1, Pharmacia).

Metal chelating columns were charged with two column volumes of 0.5M $NiCl_2$, then equilibrated with 10 column volumes of BC 100/40 (2.1.16). Cleared lysates (2.11.1) were combined with an equal volume of 2x BC 100/40 (2.1.16) to a total volume of no more than 50ml. The protein was loaded onto the column which was then washed with 2 column volumes BC 100/40 and 1 column volume of BC1000/100 (2.1.16). The protein was eluted in one column volume of BC 1000/2000 (2.1.16).

DEAE-Sepharose columns were equilibrated with 10 column volumes of 1x BC (2.1.16) adjusted to pH8.8 with NaOH. Cleared lysates (2.11.1) were combined with an equal volume of 2x BC pH8.8 (2.1.16) to a total volume of no more than 50ml. The cleared lysate was applied to the column which was then washed with 2 column volumes of 1x BC pH8.8. Under these conditions the His-ZF protein (unlike the majority of the cellular protein) does not bind to the anion exchange media, and flows through the column. Elutates were typically collected in 5ml aliquots, which were analysed by polyacrylamide gel electrophoresis (2.5.3).

Samples containing the protein of interest were dialysed (2.4.3) in Spectra/Por membrane MWCO 1000 (Spectrum). His-tagged proteins were dialysed against 2x 1l 20mM Tris-HCl pH7.9, 10mM Mercaptoethanol, 0.5mM PMSF, 20% Glycerol, for 2x 1 hr, then against 1l 20mM Tris-HCl pH 7.9, 10mM Mercaptoethanol, 0.5mM PMSF,

50% Glycerol overnight. Protein purity was determined by polyacrylamide gel electrophoresis (2.5.3), and concentration by Bio-Rad protein assay (2.11.4).

**2.11.3 Purification of GST fused protein.**

For the purification of the GST and GST fused to the zinc finger protein (GST-ZF) the protocols supplied with the vector and purification kit (Amersham Pharmacia Biotech) were adhered to unless otherwise stated.

Samples containing the protein of interest were dialysed (2.4.3) in Spectra/Por membrane MWCO 1000 (Spectrum). GST proteins were dialysed against 2x 1l 50mM Tris-HCl pH 8.0, 1mM DTT, 10mM NaCl 20% Glycerol, for 2x 1 hr. Then against 1l 50mM Tris-HCl pH 8.0, 1mM DTT, 10mM NaCl, 50% Glycerol overnight. Protein purity was determined by polyacrylamide gel electrophoresis (2.5.3), and concentration by Bio-Rad protein assay (2.11.4).

**2.11.4 Bio-Rad protein assay.**

Solutions of BSA (0.8ml) were diluted from a stock solution (Sigma) at the following concentrations; 25µg/ml, 20µg/ml, 15µg/ml, 10µg/ml, 5µg/ml, 2.5µg/ml 1µg/ml and $H_2O$ only. Samples (0.8ml) were combined with Bio-Rad protein assay reagent (0.2ml), mixed by vortexing and the colour allowed to develop at room temperature for 10 mins. The absorbance of the samples was measured at 595nm using the $H_2O$ sample as a reference. A calibration curve was plotted using the known BSA standards (which were measured alongside each assay). The concentration of the unknown sample was then extrapolated from the calibration curve.

**2.11.5 Antibody blot.**

Proteins of interest (5µl) were pipetted onto nitro-cellulose membrane. The membrane was then washed in 1x TBS pH 7.5 (2.1.7) for 5min followed by incubation in 20ml blocking solution (1x TBS pH 7.5, 3% BSA) for 1hr. The membrane was then washed twice for 5min in 20ml 1x TBS. The membrane was then incubated with 20µg/ml antibody in blocking solution for 1hr. The membrane was washed twice for 5min in 20ml 1x TBS, then incubated with 20µg/ml conjugate (Anti-goat IgG alkaline phosphatase conjugate, Sigma) in blocking solution for 1hr. Unbound conjugate was removed by two final washes in 20ml 1x TBS.

The blot was developed by incubation in a freshly prepared solution containing 1ml NBT solution (2.1.17), 1ml BCIT solution (2.1.17) in 100ml of carbonate buffer (2.1.17). The development was halted by washing the membrane in water.

**2.11.6 Protein-DNA blot.**

The membrane for DNA blots was prepared in the manner described above (2.11.5 the antibody was replaced with 10nM $^{33}P$ labelled DNA. The unbound DNA was removed by gentle rinsing in 1x TBS pH 7.5. The membrane was then allowed to dry and analysed by autoradiography (2.4.4).

**2.11.7 Scintillation proximity assay.**

The conditions used in assays unless otherwise stated were 1x TBS pH 7.5 (2.1.7), 0.5mg of protein A coated SPA beads, 1.5µg Anti-GST antibody, 10nM $^{33}P$ labelled target sequence DNA and 16.4nM protein. Target sequence DNA was made by T4 polynucleotide kinase labelling (MBI Fermentas, 2.9) of one oligonucleotide. This was then annealed to its complementary oligonucleotide (2.6.2), and un-incorporated label removed using a spin 10 column (Sigma, according to manufacturer's instructions). Negative controls contained the GST protein instead of GST-ZF protein.

Order of addition A:

DNA was pre-coupled to the zinc finger protein in a 20µl reaction for 30mins then combined with the SPA beads and antibody to a final volume of 50µl. The mixture was equilibrated by incubation at room temperature with gentle agitation for 1hr. Samples were then centrifuged in order to pellet the SPA beads the and microfuge tubes were placed directly into scintillation vials. Samples were analysed in a scintillation counter (2.4.5).

Order of adittion B:

Antibody and proteins were incubated together in a volume of 48µl with gentle agitation for 30min at room temperature. The DNA, was then added, and the mixture was then incubated for a further 30min at room temperature with gentle agitation. Assay volumes used were 50µl and 100µl (both in 1x TBS pH 7.5) the quantities of all reagents added remained the same. Samples were then centrifuged in order to pellet the SPA beads the and microfuge tubes place directly into scintillation vials. Samples were analysed in a scintillation counter (2.4.5).

All data was analysed using Microsoft excel 5.0, unless otherwise stated two tailed *t*-tests were used for data analysis.

## 2.12 SUPPLIES.

Unless otherwise stated enzymes were supplied with accompanying buffers. Restriction enzymes were obtained from either ***MBI Fermentas, (Immunogen Intl. (UK) Ltd., Sunderland, UK.*** or ***New England Biolabs, Hitchin, Herts., UK.*** Unless otherwise stated chemical reagents were supplied by ***Sigma Chemical, Dorset, UK***

***Amersham Pharmacia Biotech. Bucks., UK.***

Sequenase kit.

Thermosquenase kit.

[$^{3}$H] Amino acid mix.

[$\alpha$-$^{33}$P] Terminators.

Kodak Biomax MR film.

Kodak X-Omat AR film.

Scintillation proximity assay technology.

DEAE-Sepharose CL-6B.

GST Fusion vector (pGEX-2TK).

GST bulk purification modules.

GST detection modules (CNDB assay ant anti-GST antibody).

Hi-Trap metal chelating column.

***Amicon LTD, Gloucs., UK.***

Microcon concentrators.

Centriplus concentrators.

***BDH Ltd. Poole, UK.***

Dimethyldichlorosilane solution.

***Beckman Instruments, Inc., Palo Alto, CA., USA.***

Oligonucleotide synthesis reagents.

Ultra fast cleavage and deprotection kit.

Quick seal polyallomer tubes, capacity 3.9ml.

***Bio-Rad Laboratories Ltd, Herts., UK.***

Bio-Rad protein assay.

Chelex 100 (analytical grade chelating resin).

Nitrocellulose membrane.

***Difco Laboratories, Detroit, USA.***

Yeast extract.

Tryptone.

***DuPont NEN, Meyvis, Belgium.***

Nucleoside, [$^{35}$S] thiophosphate.

Nucleotides, [$^{32}$P].

***Fisher Scientific UK. Ltd., Loughborough, UK.***

Sodium chloride.

Sodium hydroxide.

Glycerol.

Visking dialysis tubing MWCO: 12,000-14,000.

***Fisons laboratory supplies, Loughborough, UK.***

Glacial acetic acid.

Methanol.

Sulphuric acid.

Analytical grade urea.

***FMC Bio products, Rockland USA.***

Sea Plaque GTG agarose.

***Gibco BRL Life Technologies, Paisley, Scotland, UK.***

Agarose (electrophoresis grade).

Max efficiency *E. coli* DH5α cells.

***MBI Fermentas, (Immunogen Intl. UK.) Ltd., Sunderland, UK.***

T4 polynucleotide Kinase.

T4 DNA Ligase.

***New England Biolabs, Hitchin, Herts., UK.***

Beta-Agarase.

DNA polymerase 1, large (Klenow) fragment.

Alkaline phosphatase calf intestinal (CIAP).

***OXOID, Basinstoke, Hamps., UK.***

Agar technical.

***PE-Applied Biosystems UK, Cheshire, UK.***

Oligonucleotides.

***Perkin Elmer Ltd, Warrington, Cheshire, UK.***

Taq polymerase (Amplitaq)

Amplitaq Gold.

dNTP's.

***Promega UK., Southampton, UK.***

Wizard Plus Maxi prep kit.

Wizard Plus Midi prep kit.

Wizard Plus Mini prep kit.

***Qiagen, Crawley, West Sussex, UK.***

Qiagen plasmid mini kits.

Anti-His antibodies and anti-His antibody selection kit.

***Severn Biotech Ltd., Kidderminster, Worcs., UK.***

30% Acrylamide / Bisacrylamide ratio 29:1.

6% Acrylamide sequencing solution.

***Spectrum laboratory products, LosAngles, CA., USA.***

Spectra/Por (cellulose ester) membrane, MWCO: 1000.

***Whatman Intl., Maidstone, Kent, UK.***

Chromatography paper 1Chr.

Pre-swollen microgranular (anion exchanger) DE52.

# 3 DESIGN AND CONSTRUCTION OF A ZINC FINGER GENE.

## 3.1 DESIGN OF THE ZINC FINGER GENE.

### 3.1.1 Introduction

The project is concerned with the development of a model system for the elucidation of a zinc finger / DNA interaction code. The system requires a zinc finger protein which will bind with high affinity and specificity to a known DNA sequence. For future work it must be possible to mutate the region of the zinc finger responsible for the specificity of binding, in order to observe the effect on the DNA / protein interactions.

The model zinc finger protein contains three identical zinc finger motifs, (Fig 3.1) each joined by a linker sequence. The outer two motifs are intended to strengthen the interaction with the DNA ligand (Shi and Berg, 1995a), and ensure that the central zinc finger is aligned with its target trinucleotide (Choo and Klug, 1994a). For future study the section of the central zinc finger responsible DNA binding specificity can be altered by cassette mutagenesis. Alteration of the central zinc finger such that it does not recognise the target DNA sequence should considerably lower the binding of the whole protein (Desjarlais and Berg, 1993).

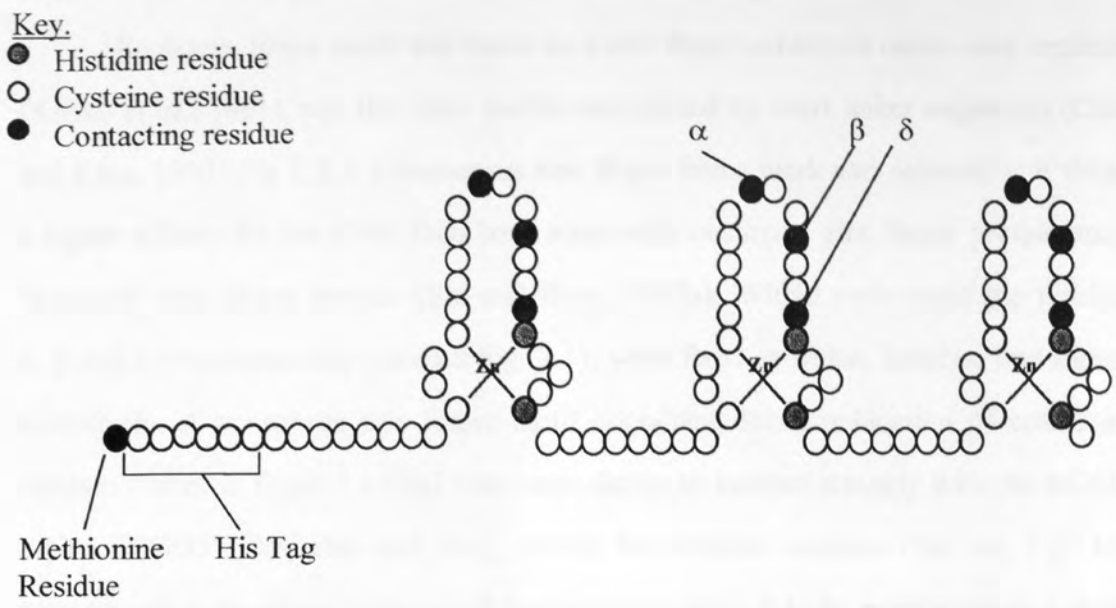



Figure 3.1 Design of the zinc finger protein

The zinc finger protein may be synthesised by the controlled expression of a synthetic gene in *E. coli* cells. Synthetic genes are typically constructed from single stranded oligonucleotides, which may be up to 100 bases long. There are two methods commonly used to assemble these into a double stranded segment of DNA. Fully overlapping oligonucleotides may be annealed together in one or several stages then ligated together. The resulting constructs can be cloned. Large genes can be cloned as several discrete fragments that are later spliced together (Englels and Uhlmann, 1989). Alternatively partially overlapping oligonucleotides can be assembled by recursive PCR (Prodromou and Pearl, 1992), again large genes may be assembled from smaller fragments, cloned separately and subsequently spliced together.

### 3.1.2 Amino acid sequence encoded by the zinc finger gene.

Each zinc finger motif was based on a zinc finger consensus amino acid sequence (Krizec *et al.*, 1991), and the three motifs were joined by short linker sequences (Choo and Klug, 1993)(Fig 3.2.). A consensus zinc finger frame work was selected as it shows a higher affinity for the DNA than both a naturally occurring zinc finger protein and a "minimal" zinc finger protein (Shi and Berg, 1995a). Within each motif the residues α, β and δ (the contacting residues Fig. 3.1), were fixed as lysine, histidine and alanine receptively. A consensus zinc finger motif containing this combination of contacting residues (taken in finger 1 of Sp1) has been shown to interact strongly with the dsDNA triplet 5'GGG3' (Desjarlais and Berg, 1993). Six histidine residues (His tag, Fig. 3.2) were placed at the amino terminus of the protein to allow it to be purified using a single step metal affinity chromatography method (Van Dyke *et al.*, 1992).

```
                                                  α     β     δ
M H H H H H H T G E K P Y K C P E C G K S F S (K) K S (H) L V (A) H Q R T H
                T G E K P Y K C P E C G K S F S (K) K S (H) L V (A) H Q R T H
                T G E K P Y K C P E C G K S F S (K) K S (H) L V (A) H Q R T H
```

Figure 3.2 Amino acid sequence encoded by the gene. Linker sequence (Choo and Klug, 1993); underlined, Zinc co-ordinating residues; bold type. Contacting residues; encircled with α, and δ indicated above.

### 3.1.3 DNA sequence of the zinc finger gene.

The amino acid sequence was translated into a DNA sequence (Fig 3.3) using the genetic code and incorporating the optimal codon bias for maximal expression in *E. coli* (Wada, *et al.,* 1992). Repetition of the DNA sequenced was minimised by alternating the codons used between motifs. This reduces the possibility of inter-genic recombination. Unique restriction enzyme sites were incorporated at each end of the gene (5' *Nde* 1 and 3' *Cla* 1), allowing the gene to be inserted into the expression vector pT7-7 (appendix III). Flanking the central zinc finger domain there are two *Age* 1 sites, which enable the removal or replacement of the central motif. Within the central zinc

finger either side of the area containing the contacting residues there are two unique restriction sites (5' *Bsi* W1 and 3' *Hind* III). These sites are intended to be used for cassette mutagenesis of this region.

```
   Nde 1
5'
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC

AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC

                                    Age 1
CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT

                               Hind III
AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT

                          Bsi W1     Age 1
CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT

AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT

                                         Cla 1
CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG AT 3'
```

Fig 3.3 DNA sequence (sense strand) of the zinc finger gene. Restriction enzyme sites are shown in bold type (Full restriction enzyme map is shown in appendix I).

## 3.2 CONSTRUCTION OF THE ZINC FINGER GENE - METHOD 1.

### 3.2.1 Gene construction method 1.

The gene sequence was divided into partially overlapping oligonucleotides (Fig 3.4). The addition of $T_8$ at the 5' end and $A_8$ at the 3' end was intended to facilitate restriction of the constructed gene with *Nde* 1 and *Cla* 1 if necessary.

```
    LJM01
5'TT TTT TTT CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA
3'AA AAA AAA GTA TAC GTG GTA GTG GTA GTG GTA TGG CCG CTT
                                  LJM02

AAA CCG TAC AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA
TTT GGC ATG TTT ACA GGC CTC ACG CCG TTT AGG AAA AGA TTT

    LJM03
AAA AGC CAC CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG
TTT TCG GTG GAC CAC CGT GTA GTT GCG TGG GTG TGG CCA CTC

AAA CCG TAT AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA
TTT GGC ATA TTT ACG GGC CTT ACG CCG TTT TCG AAG TCG TTT
                                  LJM04

AAA TCC CAT CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA
TTT AGG GTA GAC CAA CGG GTG GTC GCA TGA GTA TGG CCA CTT

LJM05
AAG CCG TAT AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG
TTC GGC ATA TTC ACG GGC CTT ACG CCT TTT AGA AAG AGG TTC

AAA TCT CAT CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG ATA
TTT AGA GTA GAC CAC CGC GTA GTC GCG TGC GTA ATT AGC TAT
                                  LJM06

AAA AAA A 3'
TTT TTT T  5'
```

Figure 3.4 Illustration of oligonucleotides used in the first method of assembly. The underlined bold bases indicate the sequences of the oligonucleotides used for gene construction.

The oligonucleotides were assembled using a method similar to recursive PCR (Prodomou and Pearl, 1992). This two stage process is similar to a method used for molecular evolution (Stemmer, 1994a; 1994b). In the first stage all the oligonucleotides were combined in a PCR reaction to produce a range of products. In the second stage a PCR was carried out using the product of the first stage as a template and oligonucleotides LJM01 and LJM06 as primers in order to amplify only the full length gene.

### 3.2.2 Method 1 gene construction, attempt 1.

The oligonucleotides were synthesised but not gel purified according to the method outlined in 2.6.1.

Stage1: the oligonucleotides LJM01-LJM06 (6pmol of each) were combined in a 50µl PCR reaction (2.9.1). The products were examined by agarose gel electrophoresis (2.5.1). A smear of fragments resulted ranging from 100 to over 300bp (Fig. 3.5).

Stage 2: a second 50µl PCR mix (2.9.1) was made containing 1µl of the product from stage 1 as a template. LJM01 (1mM) and 1mM LJM06 were used as primers to amplify only the "full length" gene product. A 55°C annealing temperature was employed. When the product was examined by agarose gel electrophoresis the fragment corresponding to the zinc finger gene was barely visible (Fig. 3.5).

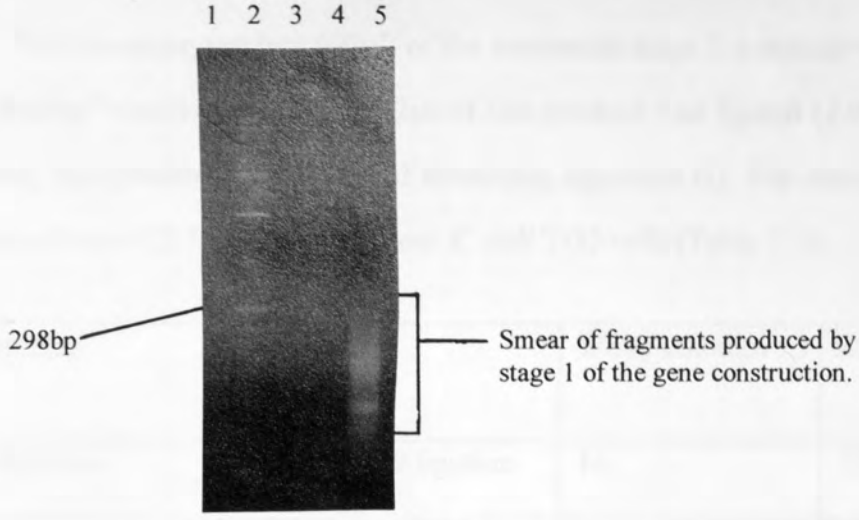



Figure 3.5 Products of gene assembly method 1, stages 1 and 2 PCR reactions were performed as described (55°C annealing temperature) and analysed by electrophoresis on a 2% agarose gel (2.5.1). Lanes: 1) pBR322/*Hin* F1 digest (0.1μg); 2) Stage 2 amplification (20% of PCR reaction); 3) Stage 2 negative control (20% of PCR reaction); 4) Stage 1 amplification (20% of PCR reaction); 5) Stage 1 negative control(20% of PCR reaction).

Stage 2 of the gene synthesis was repeated using a lower annealing temperature of 50°C. This resulted in an increased yield of product (Fig. 3.6)

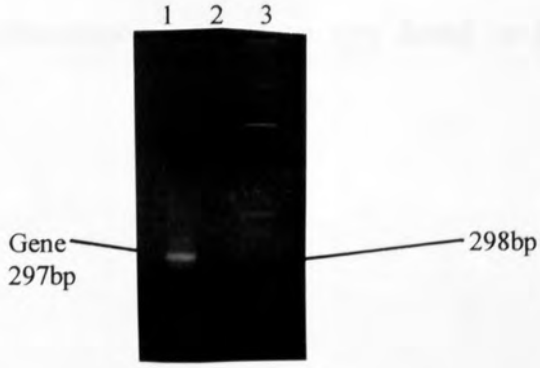



Figure 3.6 Products of gene assembly method 1, stage 2 PCR reaction was performed as described (50°C annealing temperature) and analysed by electrophoresis on a 2% agarose gel (2.5.1). Lanes: 1) Stage 2 amplification (20% of PCR reaction); 2) Stage 2 negative control (20% of PCR reaction); 3) pBR322/*Hin* F1 digest (0.1μg).

The remaining product (40µl) of the successful stage 2 synthesis was placed in a "blunt ending" reaction (2.9.2), and 2µl of this product was ligated (2.9.3) into *Sma* 1 restricted, phosphatased pUC19 (MBI fermentas, appendix II). The resulting constructs were transformed (2.7.1) into competent *E. coli* TG2 cells (Table 3.1).

| Experiment | Components transformed | White colonies | Blue colonies |
|---|---|---|---|
| Gene ligation | Gene / pUC19 ligation | 16 | 129 |
| Control ligation | pUC19 self ligation | 2 | 22 |
| Control transformation | Uncut pUC19 | None | Lawn |
| Antibiotic resistance | No DNA | None | None |

Table 3.1 Transformation results obtained from gene assembly method 1 attempt 1.

Four of the recombinant colonies obtained were screened by PCR using LJM01 and LJM06 as primers (2.9.1). Two colonies containing the gene were identified. Plasmid DNA from one clone (ZF-Gene 1.1) was prepared using a Maxi prep kit (Promega, 2.8). The DNA was sequenced from LJM01 and LJM06 by automated sequencing (Alta-Biosciences), the gene was found to have three single base pair deletions (Fig. 3.7).

```
5'
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC

                  Δ
AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC

CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT

          Δ                                           Δ
AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT

CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT

AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT

CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG AT 3'
```

Figure 3.7 Diagram of the sequence (sense strand) of ZF-Gene 1.1. Deleted bases are shown in bold with Δ symbol directly above.

### 3.2.3 Method 1 gene construction, attempt 2.

Oligonucleotides LJM02 and LJM04 were re-synthesised, then the gene re-constructed and ligated into pUC19 (3.2.2). The pUC19 constructs were transformed into competent *E. coli* TG2 cells (2.7.1). Results are shown in Table 3.2.

| Experiment | Components transformed | White colonies | Blue colonies |
|---|---|---|---|
| Gene ligation | Gene, pUC19 ligation. | 11 | 30 |
| Control transformation | Uncut pUC19 | None | Lawn |
| Antibiotic resistance | No DNA | None | None |

Table 3.2 Results of transformation of gene construct attempt 2, method 1.

Recombinant colonies were screened by PCR (2.9.1) using LJM01 and LJM06 as primers. Three clones containing the gene were identified. Plasmid DNA from one clone (ZF-Gene 1.2), was prepared using a Maxi prep kit (Promega, 2.8). The construct was sequenced from pUC 19 forward and reverse primers (appendix V) using a Sequenase sequencing kit (Amersham Pharmacia Biotech, 2.9), the gene was found to have three single base pair deletions (Fig. 3.8).

```
5'CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC
                                                  Δ
  AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC

  CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT

  AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT
                                              Δ
  CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT

  AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT
        Δ
  CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG AT3'
```

Figure 3.8 Diagram of the sequence (sense strand) of ZF-Gene 1.2. Deleted bases shown in bold with Δ symbol directly above.

Since the deletions in ZF-Gene 1.2 were in different locations to those in the previous version of the gene (ZF-Gene 1.1), it was considered unlikely that specific deletions were present in the oligonucleotides. As synthesis of new oligonucleotides did not eliminate the problem, it was possible that the method of construction of the gene was resulting in the mutations. An attempt was made therefore to correct ZF-Gene 1.2.

## 3.3 ATTEMPT TO CORRECT ZF-GENE 1.2.

### 3.3.1 Method of correction.

Correction of the gene required several stages, replacing initially the two bases deleted towards the 3' end of the gene. The mutant gene (ZF-Gene 1.2) was sub-cloned into the expression vector pT7-7 (appendix III). A fragment corresponding to the 3' end of the gene (Fig. 3.8) was then synthesised and cloned in pUC19 (appendix II). This fragment was then sub-cloned into the pT7-7/ZF-Gene 1.2 construct, to correct the two 3' mutations.

### 3.3.2 Sub-cloning of ZF-Gene 1.2 into pT7-7.

The expression vector pT7-7 and construct ZF-Gene 1.2 were restricted with *Nde* 1 and *Cla* 1 (2.9). Both the vector (pT7-7) and the insert (ZF-Gene 1.2) DNA were gel purified (2.9.4). The excised gene was then ligated into the expression vector (2.9.3). The product of the ligation reaction was transformed into competent *E. coli* TG2 (2.7.1) cells (Table 3.3).

| Experiment | Components transformed | Colonies |
|---|---|---|
| Gene ligation | Gene, pT7-7 ligation | 59 |
| Control ligation | pT7-7 self ligation | 38 |
| Control transformation | Uncut pT7-7 | Lawn |
| Antibiotic resistance | No DNA | None |

Table 3.3 Results of transformation of sub-cloned ZF-Gene 1.2.

Colonies containing the gene were identified by PCR (2.9.1) using pT7-7 forward and reverse primers (appendix V). Plasmid DNA from one clone (pT7-7/ZF-Gene 1.2) was prepared using a Maxi prep kit (Promega, 2.8). Confirmation that a pT7-7 clone had been isolated (rather than the original pUC19 clone) was obtained by *Nde* 1 digestion and electrophoresis of the plasmid DNA (data not shown).

### 3.3.3. Construction and cloning of the gene correction fragment.

The 3' end of the gene containing two mutations (between the *Spl* 1 and *Bsi* W1 sites) was re-synthesised. LJM05 was annealed to LJM06 by combining them at 1pmol/μl in 20μl 1x DNA polymerase I, (Klenow) fragment buffer (New England Biolabs). Oligonucleotides LJM05 and LJM06 were annealed by incubation at 95°C for 5 mins then allowed cool slowly to room temperature (2.6.2). The annealed construct was made double stranded by incubation at 37°C, for 30 mins, in a 30μl 1x DNA polymerase 1, large (Klenow) fragment buffer (New England Biolabs), 1mM dNTPs (Perkin Elmer) and 5 units of DNA polymerase 1 large (Klenow) fragement (New England Biolabs).

```
LJM07
5'
CGT ACG CAT ACC GGT GAA AAG CCG TAT AAG TGC CCG GAA TGC
GCA TGA GTA TGG CCA CTT TTC GGC ATA TTC ACG GGC CTT ACG
3'

                                    LJM05
GGT AAA TCT TTC TCC AAG AAA TCT CAT CTG GTG GCG CAT CAG
CCT TTT AGA AAG AGG TTC TTT AGA GTA GAC CAC CGC GTA GTC


CGC ACG CAT TAA TCG ATA AAA AAA A 3'
GCG TGC GTA ATT AGC TAT TTT TTT T 5'
               LJM06
```

Figure 3.9 Diagram showing gene correction fragment. LJM05 and LJM06 are underlined and LJM07 is shown in bold.

The annealed double stranded DNA (1μl) was used as a template for a 50μl PCR (2.9.1) using oligonucleotide primers LJM06 and LJM07 (Fig. 3.8). The PCR product was made blunt ended (2.9.2), and ligated (2.9.3) into *Sma* 1 restricted, phosphatased pUC19 (MBI Fermentas). The resulting constructs were transformed (2.7.1) into competent *E. coli* TG2 cells (Table 3.4).

| Experiment | Components transformed | White colonies | Blue colonies |
|---|---|---|---|
| Gene ligation | Gene correction fragment, pUC19, ligation | 162 | 46 |
| Control transformation | Uncut pUC19 | None | Lawn |
| Antibody resistance | No DNA | None | None |

Table 3.4 Results of transformation of gene correction fragment ligated into pUC19 *Sma*1.

Colonies containing inserts the size of the gene correction fragment were identified by PCR (2.9.1) using pUC19 forward and reverse primers (appendix V). Plasmid DNA from one clone containing the gene correction fragment was prepared by Maxi prep (Promega, 2.8).

### 3.3.4 Sub-cloning of gene correction fragment into pT7-7/ZF Gene 1.2 construct.

Both pT7-7/ZF-Gene 1.2 and pUC19 containing the gene correction fragment were restricted with *Bsi* W1 and *Cla* 1 (2.9). The vector (pT7-7/ZF-Gene 1.2) and the gene correction fragment were gel purified (2.9.4). The gene correction fragment was ligated (2.9.3) into the pT7-7/ZF-Gene 1.2 construct, and the products transformed (2.7.1) into competent *E. coli* TG2 cells (Table 3.5).

| Experiment | Components transformed | Colonies |
|---|---|---|
| Gene ligation | Gene correction fragment, pT7-7/ZF-Gene 1.2 ligation | 81 |
| Control ligation | pT7-7 self ligation | 8 |
| Control transformation | Uncut pT7-7/ZF-Gene 1.2 | Lawn |
| Antibiotic resistance | No DNA | None |

Table 3.5 Result of transformation of corrected pT7-7/ZF-Gene 1.2.

Possible recombinant colonies were screened by PCR (2.9.1) using LJM01 and LJM06 as primers, to detect those containing the whole gene. Four clones containing the gene were detected, plasmid DNA from all of these clones was prepared using a Midi prep kit (Promega, 2.8). One of the two mutations which should have been corrected was within the 3' *Age*1 restriction site. Correction of the mutation restores the *Age* 1 site. Restoration of the site could be detected by restriction of the plasmid DNA with *Age* 1 and *Ssp* 1. The plasmid DNA from all four constructs was restricted with *Age* 1 and *Ssp* 1 (2.9), and the products electrophoresed on a 1% agarose gel (2.5.1). One clone in which the 3' *Age* 1 restriction site had been restored was identified (ZF-Gene 1.3).

Plasmid DNA was prepared from this clone using a Maxi prep kit (Promega, 2.8). The gene was sequenced from pT7-7 forward and reverse primers (appendix V) using a Sequenase kit (Amersham Pharmacia Biotech, 2.9). The two mutations in ZF-Gene 1.2 had been corrected but ZF-Gene 1.3 had gained a new deletion in the region of the gene correction fragment (Fig. 3.10).

```
5'
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC

                                                  Δ
AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC

CTG GTG GCA CAT CAA CGC ACC CAC  ACC GGT GAG AAA CCG TAT

AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT

                                        Δ
CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT

AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT

CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG AT 3'
```

Figure 3.10 Diagram showing the sequence (sense strand) of ZF-Gene 1.3. Deleted bases shown in bold with Δ symbol directly above. Gene correction fragment is underlined.

The plasmid DNA of a second pUC19 construct containing the gene correction fragment (3.3.3) was prepared using a Wizard Maxi prep kit (Promega, 2.8). The DNA (ZF-Gene 1.4) was sequenced form pUC19 forward and reverse primers (appendix V) using a Sequenase kit (Amersham Pharmacia Biotech, 2.9). The gene fragment contained a novel single base deletion (Fig 3.11).

```
5'
CGT ACG CAT ACC GGT GAA AAG CCG TAT AAG TGC CCG GAA TGC

                                                  Δ
GGT AAA TCT TTC TCC AAG AAA TCT CAT CTG GTG GCG CAT CAG

CGC ACG CAT TAA TCG AT 3'
```

Figure 3.11 Diagram showing the sequence (sense strand) of ZF-Gene 1.4. The deleted base is shown in bold with Δ symbol above.

The single base deletions were not a result of the method of synthesis. If the deletions were present in the oligonucleotides themselves, then a mixed population of complete and deleted molecules must be present. In order to investigate this possibility oligonucleotides were obtained from an alternative source.

## 3.4 CONSTRUCTION OF THE ZINC FINGER GENE , METHOD 2.

### 3.4.1 Construction method 2.

In order to investigate the potential existence of deletions in the oligonucleotides, the gene was divided into two sections; the front (5') end, and the back (3') end. The front end of the gene was synthesised from gel purified oligonucleotides which were prepared as described in 2.6.1. The back end was synthesised from oligonucleotides purchased from PE-Applied biosystems. The separate synthesis for the two ends of the gene in the absence of an enzymatic step would indicate whether or not the deletion were associated with the preparation of the oligonucleotides. The division of the gene into two segments also limits the possibility of inter-genic recombination due to repetition of the gene sequence.

The front end was segmented into 4 overlapping oligonucleotides. LJM03 was incorporated into the design which required three new oligonucleotides LJM08, LJM09 and LJM10 (Fig 3.12). The construction method involved no PCR stage, as the fully overlapping oligonucleotides could be assembled by annealing and made double stranded using DNA polymerase I, large fragment (Klenow). The front end of the gene was initially ligated into pUC19 (appendix II). The gene contains a *Nde* 1 site at its 5' end and a *Bsi* W1 site at its 3' end. The construct could then be excised from pUC19 by restriction with *Nde* 1 and *Bsi* W1 and ligated to the back end of the gene.

The back end was assembled from 2 new overlapping oligonucleotides AVH1 and AVH2 (Fig. 3.12). The complementary oligonucleotides when annealed have overhangs at each end, the 5' one corresponding to sticky end produced by *Bsi* W1 restriction digestion. The 3' overhang corresponds to the sticky end produced by a *Cla* 1 restriction digestion. The back end of the gene could be ligated on to the front end of the gene at the *Bsi* W1 site. The whole construct could then be ligated into pT7-7 (appendix III).

```
    Nde 1       LJM08
5'
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC
GTA TAC GTG GTA GTG GTA GTG GTA TGG CCG CTT TTT GGC ATG
3'              LJM09


                                                  LJM03
AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC
TTT ACA GGC CTC ACG CCG TTT AGG AAA AGA TTT TTT TCG GTG


CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT
GAC CAC CGT GTA GTT GCG TGG GTG TGG CCA CTC TTT GGC ATA
                   LJM10


AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT
TTT ACG GGC CTT ACG CCG TTT TCG AAG TCG TTT TTT AGG GTA


                       Bsi W1       AVH1
CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT
GAC CAA CGG GTG GTC GCA TGC GTA TGG CCA CTT TTC GGC ATA
                                AVH2


AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT  CAT
TTC ACG GGC CTT ACG CCT TTT AGA AAG AGG TTC TTT  AGA GTA


                                      Cla 1
                                ("restricted" site)
CTG GTG GCG CAT CAG CGC ACG CAT TAA T   3'
GAC CAC CGC GTA GTC GCG TGC GTA ATT ACG 5'
```

Figure 3.12 Illustration of oligonucleotides used in the second method of assembly. Oligonucleotides shown in colour were prepared as described in 2.6.1. Oligonucleotides purchased from PE-Applied biosystems are underlined. Restriction sites are shown in bold.

### 3.4.2 Construction of the front end of the gene.

Oligonucleoties LJM08, LJM09 and LJM10 were prepared as described in 2.6.1. Oligonucleotides LJM03, LJM08, LJM09 and LJM10 were purified (2.6.1), removing any oligonucleotide that was not full length.

Phosphate groups were attached to LJM09 (30pmol) and 30pmol LJM03 in a 30µl T4 polynucleotide kinase reaction (2.9).

The construct was prepared as follows:

Annealing stage 1: LJM08 (30pmol in 30µl) was annealed (2.6.2) to LJM09 (30pmol in 20µl), and LJM03 (30pmol in 30µl) annealed (2.6.2) to LJM10 (30pmol in 20µl).

Annealing stage 2: LJM08/09 (25pmol) and 25pmol LJM03/10 were combined and annealed together by incubation at 70°C for 5 mins followed by cooling slowly to room temperature. The annealed product was then ligated in a 110µl reaction containing, 3.4µl ligation buffer (MBI Fermentas), 10 units of ligase (MBI Fermentas), and 2µl 20mM ATP. The mixture was incubated at room temperature for 2 hrs, then heated to 65°C for 10 mins, 25µl was reserved and electrophoresed (2.5.1) on a 3% agarose gel (Fig. 3.13).

The remaining ligated product of the second annealing stage was made blunt ended by combining with, 1µl of each 10mM dNTP, 2mM ATP, 10 units DNA polymerase large fragment (klenow, MBI Fermentas), and 10 units T4 polynucleotide kinase (MBI Fermentas). This was placed at 37°C for 15 mins then heated to 65°C for 10 mins to inactivate the enzymes. The product of the "blunt ending" reaction (30µl) was examined by agarose gel electrophoresis (Fig. 3.13).

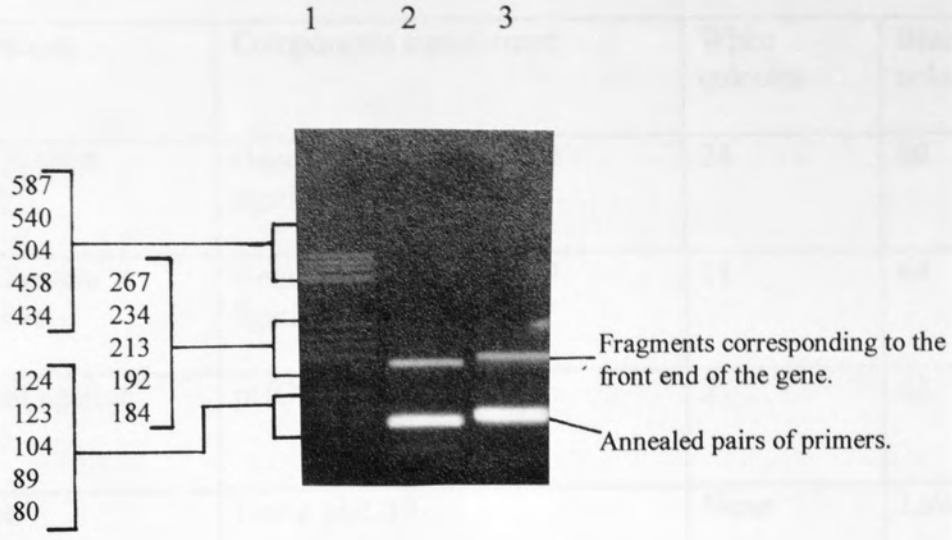



Figure 3.13 Agarose gel electrophoresis showing ligated, and blunt ended products of front end gene synthesis, method 2. Lanes:1) Molecular weight marker pBR322 *Hae* III (0.3μg); 2) Annealed oligonucleotides LJM03, LJM08, LJM09 and LJM10 (22.7% of the product); 3) Constructed front end of gene (33.3% of the product).

The product of the blunt end reaction (5μl) was ligated into pUC19 (2.9.3). The ligated products were transformed (2.7.1) into competent *E. coli* TG2 and *E. coli* HB101 cells (Table 3.6).

| Experiment | Components transformed | White colonies | Blue colonies |
|---|---|---|---|
| Gene ligation (TG2) | Gene (front end), pUC19 ligation | 26 | 89 |
| Gene ligation (HB101) | Gene (front end), pUC19 ligation | 11 | 64 |
| Control ligation (TG2) | pUC19 self ligation | 2 | 62 |
| Control transformation (TG2) | Uncut pUC19 | None | Lawn |
| Antibiotic resistance (TG2) | No DNA | None | None |
| Antibiotic resistance (HB101) | No DNA | None | None |

Table 3.6 Results of transformation of the front end of the gene assembled by method 2.

Recombinant colonies from each strain were screened by PCR (2.9.1) using pUC19 forward and reverse primers. Two *E. coli* TG2 colonies (ZF-Genes 2.1 and 2.2), and one HB101 colony (ZF-Gene 2.3) containing the gene were isolated. Plasmid DNA from all three clones was prepared using Qiagen plasmid mini kits (2.8). The DNA was sequenced with pUC19 forward and reverse primers (appendix V) using a Sequenase kit (Amersham Pharmacia Biotech, 2.9). ZF-Genes 2.1, 2.2 and 2.3 all contained mutations within the zinc finger gene (Fig. 3.14).

```
5'   Δ(2.3) Δ(2.1)     Δ(2.2)
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC


AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC


                                                    Δ(2.3)
CTG GTG GCA CAT CAA CGC ACC CAC  ACC GGT GAG AAA CCG TAT

         Substitution
           2.2&2.3
AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT


CTG GTT GCC CAC CAG CGT ACG 3'
```

Figure 3.14 Diagram showing the sequence (sense strand) of ZF-Genes 2.1, 2.2 and 2.3. Mutations shown in bold type, Δ indicates deletions, substitution = G conversion to A. Numbers above the mutated bases indicate in which clone the mutation occurred.

### 3.4.2 Synthesis of the back end of the gene.

Since the construction of the front end of the gene had been unsuccessful, the back end was constructed, made blunt ended and cloned directly into pUC19. It would not be possible to subclone such a construct as the restriction sites were not complete. This experiment was intended only to asses whether or not oligonucleotides prepared independently would result in a construct containing mutations.

AVH1 and AVH2 were purchased from PE-Applied biosystems (ABI ReadyPure DNA). They were annealed together at 1pmol/μl in a 40μl 1x PCR buffer (Perkin Elmer), by heating to 95°C for 5 mins and allowing them to cool slowly to room temperature. The annealed product was placed in a blunt end reaction (2.9.2) and 5μl of the blunted product was ligated (2.10.3) into *Sma* 1 restricted, phophatased pUC19 (MBI Fermentas, appendix II). The constructs were transformed into competent *E. coli* TG2 cells (2.7.1).

The resulting recombinant (white) colonies were screened by PCR (2.9.1) using pUC 19 forward and reverse primers (appendix V) to identify constructs containing an insert. The plasmid DNA (ZF-Gene 2.4) from one of these colonies was prepared using a Qiagen plasmid mini kit (2.8). The construct was sequenced from pUC19 forward and reverse primers (appendix V) using a Sequenase kit (Amersham Pharmacia Biotech, 2.9). The sequence was found to be correct (data not shown). This implied that the deletions were a product of the oligonucleotides synthesised "in house" as oligonucleotides purchased from PE applied biosystems had not produced a cloned product containing mutations.

3.5. GENE CONSTRUCTION METHOD 3.

### 3.5.1 Method of construction 3.

The synthesis of the back end of the gene remained the same as this had previously been successful (3.4.2). The front end of the gene was divided into four new overlapping oligonucleotides, AVH3, AVH4, AVH5 and AVH6 (Fig. 3.15). These were assembled using two stages of annealing. Annealed AVH1/2 was then ligated onto the back end of the gene and the construct inserted directly into pT7-7 (appendix III).

```
   Nde 1
5'("restricted" site)          AVH3
AT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC
   TAC GTG GTA GTG GTA GTG GTA TGG CCG CTT TTT GGC ATG
3'                             AVH5

AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC
TTT ACA GGC CTC ACG CCG TTT AGG AAA AGA TTT TTT TCG GTG

                                      AVH4
CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT
GAC CAC CGT GTA GTT GCG TGG GTG TGG CCA CTC TTT GGC ATA
                AVH6

AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT
TTT ACG GGC CTT ACG CCG TTT TCG AAG TCG TTT TTT AGG GTA

                            Bsi W1                AVH1
CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT
GAC CAA CGG GTG GTC GCA TGC GTA TGG CCA CTT TTC GGC ATA
                                                  AVH2

AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT
TTC ACG GGC CTT ACG CCT TTT AGA AAG AGG TTC TTT AGA GTA

                                        Cla 1
                                   ("restricted" site)
CTG GTG GCG CAT CAG CGC ACG CAT TAA T    3'
GAC CAC CGC GTA GTC GCG TGC GTA ATT ACG  5'
```

Figure 3.15 Diagram showing oligonucleotides used in method of gene synthesis 3. Oligonucleotides are shown in colours indicated, restriction sites are shown in bold.

**3.5.2 Construction of gene method 3.**

Oligonucleotides AVH3, 4, 5 and 6 were purchased from PE applied biosytems (ABI ReadyPure DNA). Phosphate groups were attached to 40pmol of AVH1, AVH4, and AVH5 in a 20µl reactions using T4 polynucleotide kinase (MBI Fermentas, 2.9). AVH6 (40pmol) was phosphatased in a 10µl reaction using T4 polynucleotide kinase (MBI Fermentas, 2.9).

The construct was prepared as follows:

Annealing stage 1: AVH1+AVH2, AVH3+AVH5, AVH4+AVH6 (40pmol each oligonucleotide), were each combined in 30µl of 1x T4 polynucleotide kinase buffer (MBI Fermentas). The pairs of oligonucleotides were annealed by heating the samples to 95$^{o}$C for 5 mins the cooling slowly to room temperature (2.6.2).

Annealing stage 2: AVH3/5 was then annealed to AVH4/6 by warming to 70$^{o}$C combining, and allowing to cool slowly to room temperature (2.6.2). The resulting construct was examined by agarose gel electrophoresis (Fig. 3.16).

The remaining 51.5µl was combined with 30µl AVH1/2, 10µl 10mM ATP, 10 units of T4 DNA ligase (MBI Fermentas), 1.85µl 10x T4 DNA ligase buffer (MBI Fermentas), and 1.65µl sterile $H_2O$. The ligation reaction was incubated overnight at room temperature and the product analysed by agarose gel electrophoresis (Fig. 3.16).

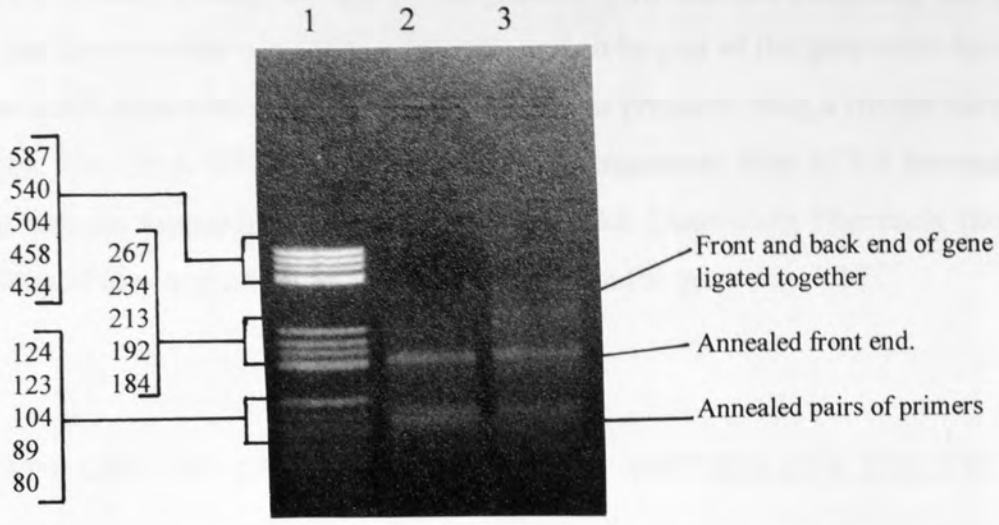



Figure 3.16 Agarose gel electrophoresis (3% gel, 2.5.1) of annealed constructs AVH3/5/4/7 and the annealed construct with AVH1/2 ligated (3.5.2). Lanes: 1) Molecular weight marker pBR322 *Hae* III (0.3μg); 2) Annealed front end (14% of the product); 3) Front and back end ligated together (17.4% of the ligation reaction).

The constructed gene (1μl of the ligated product) was ligated (2.9.3) into pT7-7 pre-restricted with *Nde* 1 and *Cla* 1 (2.9). The ligations were transformed (2.7.1) into competent *E. coli* TG2 cells (Table 3.7).

| Experiment | Components transformed | Colonies |
|---|---|---|
| Gene ligation | Gene, pT7-7 ligation | 8 |
| Control ligation | pT7-7 self ligation | 10 |
| Control transformation | Uncut pT7-7 | Lawn |
| Antibody resistance | No DNA | None |

Table 3.7 Results of transformation of gene constructed by method 3.

Colonies resulting from the transformation of the gene ligation were screened by PCR (2.9.1) using LJM01 and LJM06 as primers. Two colonies containing the whole gene, and four colonies containing what appeared to be part of the gene were detected. Plasmid DNA from both of the full length clones was prepared using a Qiagen mini prep kit (2.8). The DNA (ZF-Gene 3.1 and 3.2) was sequenced from pT7-7 forward and reverse primers (appendix V) using a Sequenase kit (Amersham Pharmacia Biotech, 2.9). Both of the constucts contained deletions within the gene (Fig. 3.17).

```
5'
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC

                            Δ
AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC

CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT

                Δ
AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT

CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT

AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT

CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG AT 3'
```

Figure 3.17 Diagram of the sequences (sense strand) of ZF-Gene 3.1 and 3.2. Deletion found in ZF-Gene 3.1 shown underlined, deletions found in ZF-Gene 3.2 shown in bold with Δ directly above deleted base.

The gene preparation was re-synthesised as descried above and ligated (2.9.3) into freshly prepared pT7-7 restricted with *Nde* 1 and *Cla* 1 (2.9). In order to increase the efficiency of the ligation, the amount of gene construct ligated was increased to 5µl. To improve the efficiency of transformation the constructs were transformed (2.7.1) into MAX Efficiency DH5α competent cells (Gibco BRL). Results are shown in Table 3.8.

| Experiment | Components transformed | Colonies |
|---|---|---|
| Gene ligation | Gene, pT7-7 ligation | 680 (approximately) |
| Control ligation | pT7-7 self ligation | 8 |
| Antibiotic resistance | No DNA | None |

Table 3.8 Results of transformation of gene constructed by method 3 into MAX Efficiency DH5α competent cells.

Colonies (200) were screened by hybridisation (2.10) using AVH3 labelled with $^{32}P$ using T4 Kinase (2.9). Seventeen clones containing the gene were identified. A large scale plasmid DNA preparation (2.8.1) was made from 6 of these clones. The DNA of five of these clones (ZF-Gene 3.3, 3.4, 3.5, 3.6 and 3.7) was sequenced using a sequenase kit (Amersham Pharmacia Biotech, 2.9) from pT7-7 forward and reverse primers (appendix V). Four of the constructs (ZF-Gene 3.3, 3.4, 3.5 and 3.6) contained deletions (Fig. 3.18), ZF-Gene 3.7 however contained no mutations (Fig. 3.19). Sequencing results consistently showed a compression in the 3' end of the gene. The compression was present in the results from sequencing of both strands. In order to confirm that the sequence of the gene was correct within this region, the plasmid DNA was analysed using dITP sequencing. This method confirmed that a gene of the required sequence had been cloned (Fig. 3.19).

As a correct synthetic gene had been difficult to isolate it was necessary to check that the construct, now cloned, was stable. In order to check the stability of the ZF-Gene 3.7 sequence, the plasmid DNA was re-transformed (2.7.1) into competent *E. coli* DH5α cells. It was also necessary to check the stability of the sequence in the strains in which the gene would be expressed. ZF-Gene 3.7 was therefore also transformed (2.7.1) into competent *E. coli* BL21 [DE3] pLysS and *E. coli* HMS [DE3] 174 pLysS cells. A large scale Plasmid DNA preparation (2.8.1) was made from each of the strains. The DNA prepared was sequenced as before from pT7-7 forward and reverse primers. No mutations were detected in any of the DNA preparations (data not shown).

5'
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC

Δ(3.5) Δ(3.4) Δ(3.3)
AAA TGT CCG G**A**G TGC GGC AAA TCC TTT TC**T** AAA AAA **A**GC CAC

Δ(3.5)
CTG GTG GCA CAT CAA CGC ACC **C**AC ACC GGT GAG AAA CCG TAT

AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT

Δ(3.3)
CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG T**A**T

AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT

CTG GTG GCG C*AT CAG* CGC ACG CAT TAA TCG AT 3'

Figure 3.18 Diagram showing the sequence (sense strand) of ZF-Genes 3.3, 3.4, 3.5 and 3.6. Single base deletions are shown in bold, with a Δ symbol directly above the deleted base. Numbers above the mutation indicate the clone in which the mutations were found. Multiple base deletion in 3.3 shown in red, multiple base deletion in 3.4 is shown in green. Underlined sequence is the large deletion found in ZF-Gene 3.6.

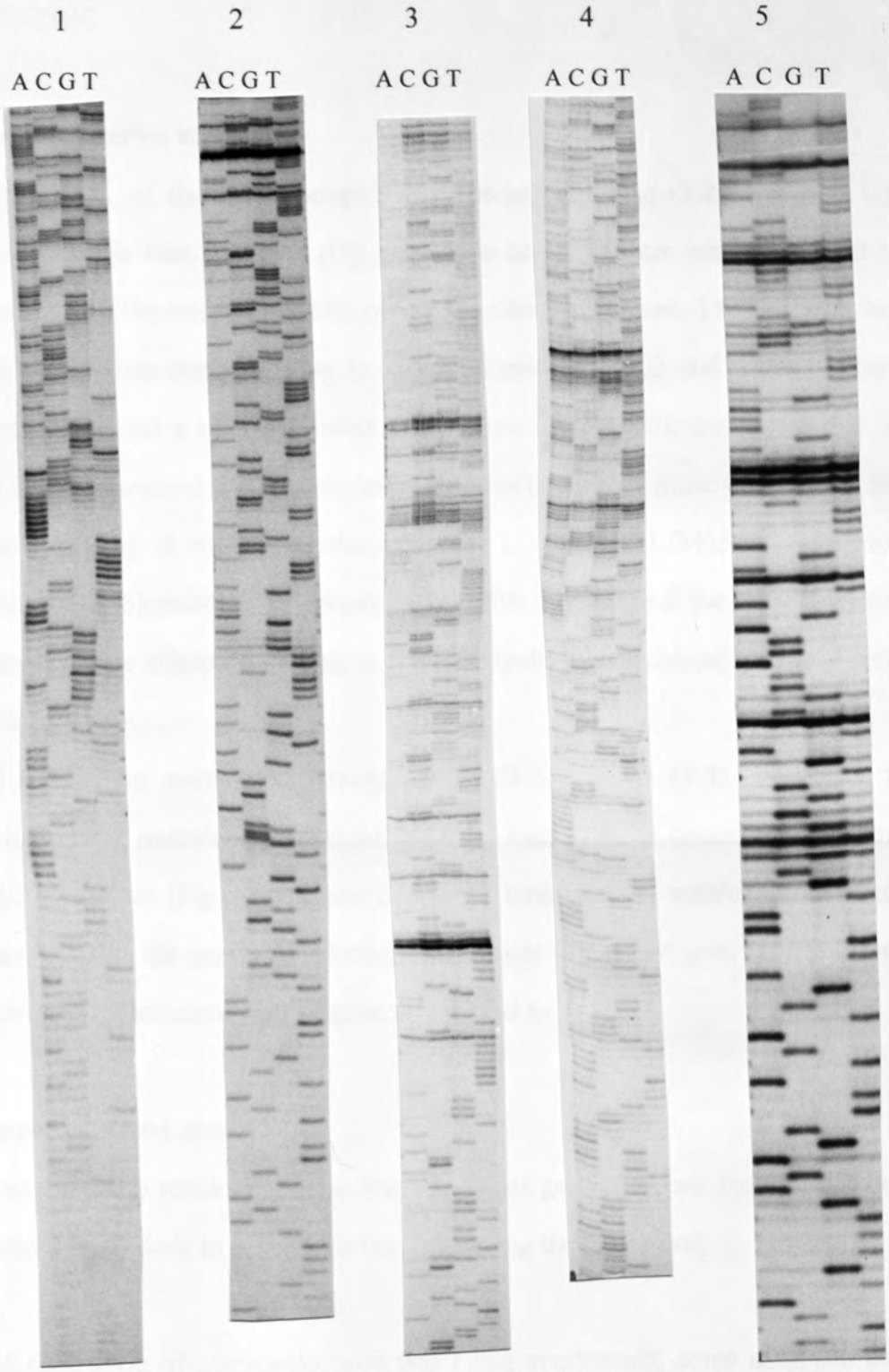



Figure 3.19 Sequence of ZF-Gene 3.7. 1) Sequence from pT7-7 forward primer (short run). 2) Sequence from pT7-7 reverse primer primer (short run). 3) Sequence from pT7-7 forward primer (long run). 4) Sequence from pT7-7 reverse primer (long run).5) Sequence from pT7-7 reverse primer, performed using dITP reagents and protocol (Amersham Pharmacia Biotech, 2.9).

## 3.6 DISCUSSION.

### 3.6.1 Gene construction method 1.

The product of the first attempt to synthesise the gene (3.2) ZF-Gene 1.1, contained three single base deletions (Fig. 3.7). The deletions were initially thought to have been present in the oligonucleotides used to synthesise the gene. The deletions had occurred in the regions corresponding to oligonucleotides LJM02 and LJM04. These were re-synthesised and a second attempt at the gene construction was made (3.2.3). ZF-Gene 1.2 also contained three single base deletions (Fig. 3.8), these were within the regions corresponding to oligonucleotides LJM02, LJM04 and LJM05. No mutation had previously been identified in the region of LJM05. Therefore if the mutation were indeed present in the oligonucleotides, a mixed population of deleted and complete molecules was present.

An attempt to correct the mutations in ZF-Gene 1.2 (3.3) succeeded in correcting the two 3' mutations. But in doing so, on two separate occasions (ZF-Gene 1.3 (Fig. 3.10) and 1.4 (Fig. 3.11)) new mutations were created within the fragment synthesised to correct the gene. Sub-cloning of the front end of the gene did not cause any new mutations, once cloned the sequence appeared to be stable.

### 3.6.2 Gene construction method 2.

Considering the results from the first method of gene synthesis there seemed to be three factors most likely to be responsible for causing the mutations;

- A mixed population of oligonucleotides was being synthesised, some complete and some containing deletions.

- The Taq polymerase (lacking any proof reading ability) was incorporating mutations and was unable to correct them.

- The sequence of the gene despite the best attempts to prevent this might be too repetitive and the cause of inter-genic recombination.

The later was unlikely because a similar gene containing three direct repeats had been successfully cloned (Desjarlais and Berg, 1993). But as a precaution the synthesis of the gene was attempted in two sections. The two portions of the gene were synthesised in the same way using different oligonucleotides in order to determine if the origin of the oligonucleotides was important. The constructs were also transformed into two different strains of *E. coli* in to determine whether or not there may be a strain specific problem.

Synthesis of the front end of the gene: Three new oligonucleotides were synthesised, LJM07, LJM08 and LJM09. These along with LJM03 were gel purified so that only DNA of one length was present, and any potentially deleted oligonucleotides would be removed. Normally deletions would only be present at the 5' end of an oligonucleotide. If "unextended" oligonucleotides are not capped (to prevent nucleotides from being added) deleted molecules may be present. The gene was assembled using two annealing stages, and the construct made completely double stranded using DNA polymerase 1, large (Klenow) fragment (3.4.2). This approach eliminated the possibility that the Taq polymerase was causing the mutations. The resulting constructs were transformed into both *E. coli* TG2, and HB101 cells (This was intended to eliminate any strain specific problem). Two *E. coli* TG2, and one HB101 clones containing the gene were sequenced ZF-Genes, 2.1, 2.2 and 2.3 (Fig 3.14) all contained single base deletions. This indicated that the deletions were unlikely to be present in the oligonucleotides.

Synthesis of the back end of the gene: The two new oligonucleotides required to synthesise the back end of the gene were purchased from PE-Applied Biosystems. These were complementary and incorporated overhangs for direct ligation on to the back end of the gene. To determine whether the deletions were specific to the oligonucleotides synthesised on the Oligo 1000 DNA synthesiser (Beckman), these were cloned and

sequenced. No mutations were found within the cloned fragment of gene. This implied that the problem was associated with the oligonucleotides synthesised on the Oligo 1000 DNA synthesiser.

### 3.6.3 Gene construction method 3.

Due to the success of the method used to construct the back end of the gene, four new oligonucleotides (AVH03, AVH04, AVH05 and AVH06) covering the region of the front end of the gene were purchased from PE-Applied Biosystems. These could be assembled by two annealing stages which were then ligated to the back end of the gene (AVH01 and AVH02)(3.5). The gene was ligated directly into pT7-7 to eliminate any possibility of expression of a toxic product, which might promote selection for constructs containing mutations. Two clones containing the gene were initially identified. ZF-Gene 3.1 and 3.2. ZF-Gene 3.1 contained three single base deletions and ZF-Gene 3.2 contained a 17 base pair deletion (Fig. 3.17).

In order to identify a clone which did not contain a deletion, or several genes which could be spliced together to construct the correct sequence, 17 clones containing the gene were identified. Initially five if these were sequenced. ZF-Genes 3.3, 3.4, 3.5 and 3.6 all contained deletions (Fig 3.18). ZF-Gene 3.7 however contained no mutations (Fig. 3.19). The stability of the ZF-Gene 3.7 was tested by re-transformation and re-sequencing. The gene once the correct sequence had been identified seemed to be stable.

### 3.6.4 Summary.

The factors which did not seem to affect the frequency of the mutations occurring in the gene include;

- Whether the oligonucleotides were synthesised on the Oligo 1000 DNA synthesiser Beckman) or by PE-applied bio-systems.

- Whether or not the oligonucleotide were purified to remove product which may have bases omitted during synthesis.

- The method by which the oligonucleotides were assembled into a gene, and whether or not enzymes were used.

- What vector the gene was ligated into. The direct use of pT7-7 eliminates the possibility that the gene was being expressed and the toxicity of its product causing selection for clones containing mutant copies of the gene.

- What strain of bacteria the constructs were transformed into. None of the strains used seemed to effect the frequency of mutations.

- The transformation process, and sequence stability alone. Once the gene was cloned it seemed to be stable.

The multiple base deletions are unique to the third method of synthesis, the mechanism by which they occurred is not known, and the isolation of only two such clones could have happened by chance. The single base deletions were unlikely to be random events as 13 clones containing similar deletions were isolated. There do not seem to be any mutational "hot spots" or preference for the deletion of a particular base (Fig. 3.20).

```
5′  2         2   2                          3
CAT ATG CAC CAT CAC CAT CAC CAT ACC GGC GAA AAA CCG TAC

              3   1        3                 1 3
AAA TGT CCG GAG TGC GGC AAA TCC TTT TCT AAA AAA AGC CAC

                                  3    3                    2
CTG GTG GCA CAT CAA CGC ACC CAC ACC GGT GAG AAA CCG TAT

            1 2 3                                    1  3
AAA TGC CCG GAA TGC GGC AAA AGC TTC AGC AAA AAA TCC CAT

                                     1     1               3
CTG GTT GCC CAC CAG CGT ACG CAT ACC GGT GAA AAG CCG TAT


AAG TGC CCG GAA TGC GGT AAA TCT TTC TCC AAG AAA TCT CAT

  1     1                 3
CTG GTG GCG CAT CAG CGC ACG CAT TAA TCG AT 3′
```

Figure 3.20 Diagram showing distribution of mutations in the zinc finger gene. Single base deletions shown in bold, multiple base deletions shown in red. The numbers above the bases indicate the method by which the clone mutations were found in was made.

The reason for the consistent occurrence of single base deletions remains unclear. Similar problems in cloning a synthetic gene have not, as far as the author is aware, been reported. The factor causing the mutations, if there is a single cause, seems to be unique to this gene sequence or methods of synthesis used. The single base deletions are unlikely for the reasons outlined above, to be present in the gene prior to cloning. As the gene has been stable since cloning, if the sequence does contribute to causing the mutations, it must be in conjunction with a factor unique to the synthetic DNA.

One possible cause is the modification of the synthetic DNA by the bacteria subsequent to transformation because it is un-methylated. However this must be in conjunction with a sequence-specific factor as no such effect has previously been

described. A second possible cause of the deletions is that the bacteria remove bases due to disruption of base pairing from nucleotides that are left protected. All three methods of synthesis utilise oligonucleotides that were cleaved and deprotected using the ultra fast cleavage and deprotection kit. The ultra fast method of deprotection is unlikely to be as effective as the traditional method. More protected bases would typically be present in the synthetic gene than in oligonucleotides deprotected by traditional methods. This could cause an increased incidence of deletion by the bacterial DNA repair systems because unprotected bases are likely to cause distortion in the base pairing.

# 4 EXPRESSION PURIFICATION AND CHARACTERISATION OF THE HIS TAGGED ZINC FINGER PROTEIN.

## 4.1 INTRODUCTION.

In order to synthesise the His tagged zinc finger protein (His-ZF) the gene encoding it (ZF-Gene 3.7, 3.5.2) must be expressed in *E. coli* cells. His-ZF may be capable of binding to the host genome and if so, might be toxic to the cells. For this reason was the gene was ligated directly into the expression vector pT7-7 (appendix III). This vector is capable of expressing inserted genes at a high level from a T7 promoter. Most of the *E. coli* strains used in cloning do not contain a gene for T7 DNA polymerase, and cannot express genes cloned into pT7-7. In order to express the gene from the T7 promoter the pT7-7 construct must be transformed into strains such as *E. coli* BL21[DE3] or HMS174[DE3], which contain copies of the T7 DNA polymerase gene. In these strains the T7 DNA polymerase is under the control of a lac promoter and expression of it is induced using IPTG. Strains containing the pLysS plasmid have an extra level of control over expression of cloned genes, as the plasmid expresses a T7 DNA polymerase inhibitor. The inhibitor inactivates small quantities of T7 polymerase present as a result of low level ("leaky") expression from the lac operator. Rifampicin can be added to these cultures after induction to prevent further expression of the *E. coli* genes. The protein of interest should then be the major product.

Purification of the protein from the *E. coli* cell relies on properties unique to the zinc finger protein. Two methods of purification have been used. The first utilises the His tag at the carboxyl end of the protein. The second relies on the fact that the protein is positively charged.

## 4.2 OPTIMISATION OF EXPRESSION OF THE ZINC FINGER GENE.

### 4.2.1 Selection of *E. coli* strain for expression.

The *E. coli* strains suitable for expression of genes inserted in pT7-7 were HMS174[DE3] and BL21[DE3], both containing or without the pLysS plasmid. ZF-Gene 3.7 was transformed into competent *E. coli* BL21[DE3] pLysS, *E. coli* BL21[DE3] *E. coli* HMS 174[DE3] pLysS and *E. coli* HMS174[DE3] cells (2.7.1). Since any "leaky" expression of the zinc finger gene might prove toxic to the cells, growth curves were measured for each strain containing ZF-Gene 3.7. A single transformed colony (from each strain) was inoculated into a small amount of LB containing 50μg/ml ampicillin (and 35μg/ml chloramphenicol for strains containing the pLysS plasmid). Cultures were incubated at 37°C, overnight with shaking. The overnight culture (0.5ml) was inoculated into 50ml of LB containing 50μg/ml ampicillin (and 35μg/ml chloramphenicol for strains containing the pLysS plasmid), and incubated at 37°C with shaking. The absorbance of the culture a 600nm was measured immediately after inoculation, then every half-hour for 8.5 hrs (Fig. 4.1).

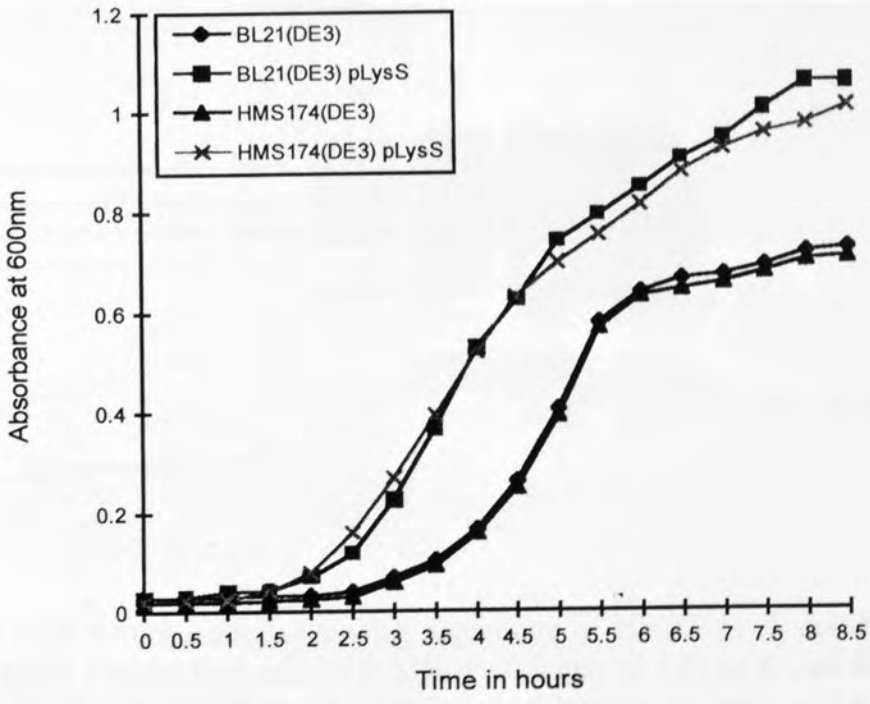



Figure 4.1 Graph showing growth of various strains of *E. coli* containing ZF-Gene 3.7. Growth was measured in absorbance at 600nm at 30min time intervals from inoculation.

Growth curves of the various strains revealed that the strains containing the pLysS plasmid grew faster than those without the extra control over the expression of the gene (Fig. 4.1). This is likely to be due to a slight toxic effect caused by a low level of expression of the gene. The strains containing the pLysS strains of the plasmid were therefore selected for expression of the zinc finger gene.

The zinc finger protein was then expressed in both *E. coli* BL21 [DE3] pLysS, and *E. coli* HMS174 [DE3] pLysS cells (30ml cultures, 2.11.1). Samples (0.5 ml) were removed from the culture before induction, then at, 1hr 2hrs and 3hrs after induction. The samples were examined by electrophoresis on a SDS-PAGE gel (Fig. 4.2).

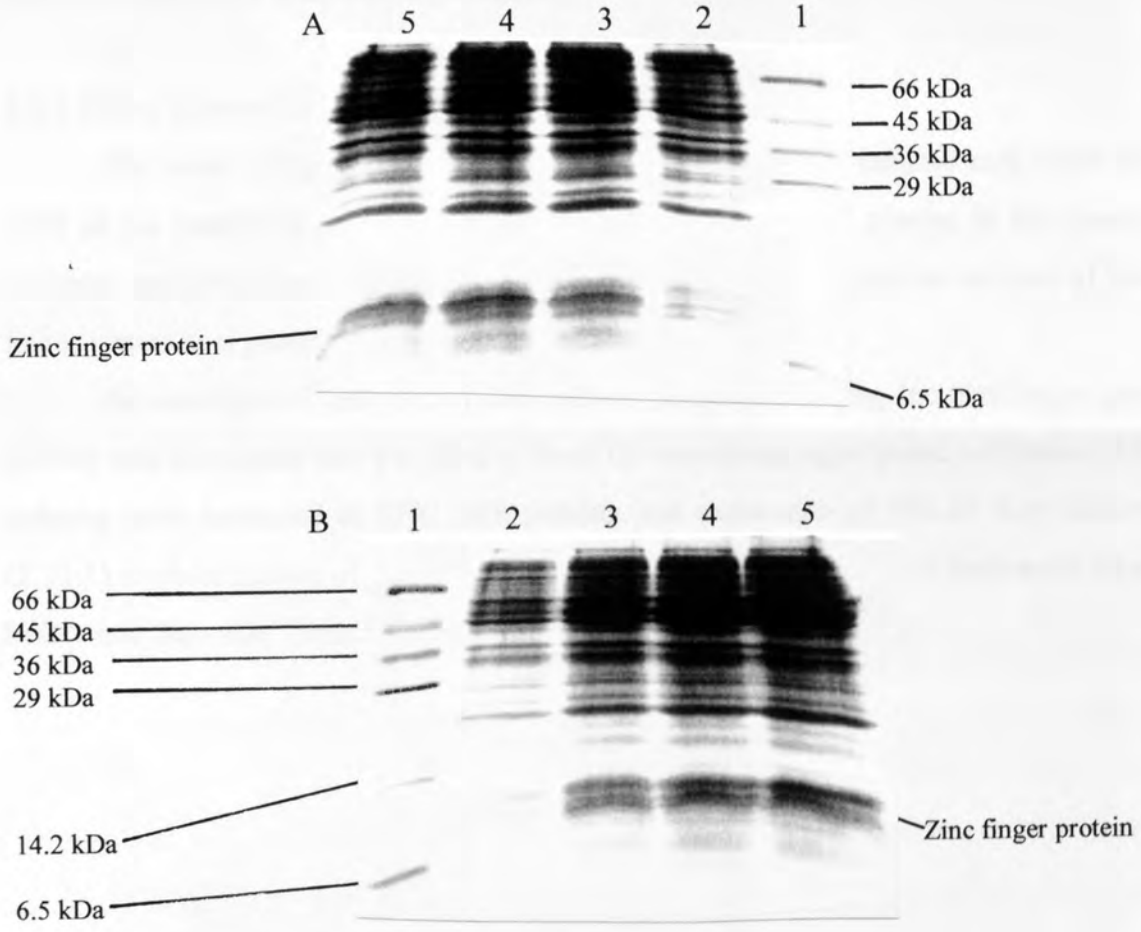



Figure 4.2 SDS PAGE analysis showing expression of His-ZF in *E. coli* BL21[DE3] pLysS. B: HMS 174[DE3] pLysS. 18% SDS-PAGE gels (2.5.1) A: *E. coli* BL21 [DE3] pLysS. B: HMS 174 [DE3] pLysS. The order of loading on both gels is the same. Lanes: 1) Low molecular weight marker (10μg, Sigma); 2) Un-induced lysate; 3-5) induced lysates, time from induction; 3) 1hr; 4) 2hrs; and 5) 3hrs.

The zinc finger protein was expressed in both *E. coli* Bl21[DE3] pLysS and HMS174[DE3] pLysS cells The levels of expression seen in both strains were similar (Fig. 4.2). The protein of interest is easier to identify in the lysates from the BL21 cells as there is less interference from similar sized host proteins. At this point *E. coli* BL21[DE3] pLysS cells were selected for the expression of His-ZF.

The time required to accumulate maximum quantities of protein was also identified during these initial experiments. There seems to be no significant accumulation of protein after 2hrs after induction (Fig 4.2). His-ZF was therefore expressed for 2hrs after induction in all following experiments.

### 4.2.2 Effect of point in growth that cultures are induced.

The point of the growth curve at which the cultures are induced may effect the yield of the protein of interest and the amount of host protein present in the lysates. Cultures were therefore induced at various points during growth and the amount of zinc finger versus host protein monitored.

An overnight *E. coli* BL21[DE3] pLysS culture containing the zinc finger gene (0.5ml) was inoculated into 4 x 50ml of fresh LB containing appropriate antibiotics. The cultures were incubated at 37°C with shaking and expression of His-ZF was induced (2.11.1) at absorbancies of 0.2, 0.4, 0.6 and 0.8 at 600nm. Samples (0.5ml) were taken before and 2hrs after induction and examined by SDS PAGE (Fig. 4.3).

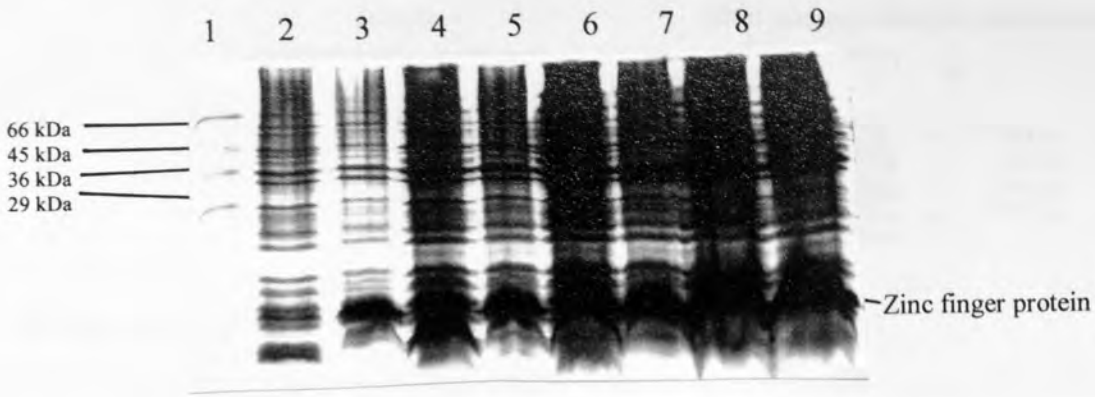



Figure 4.3 Gel showing effect on expression of the point in growth at which the culture is induced. 18% SDS-PAGE gel (2.5.3), lanes: 1) Low molecular weight marker (Sigma 10μg); Points of induction during growth in absorbance at 600nm, 2) 0.2 Un-induced; 3) 0.2 Induced; 4) 0.4 Un-induced; 5) 0.4 Induced; 6) 0.6 Un-induced; 7) 0.6 Induced; 8) 0.8 Un-induced; 9) 0.8 Induced.

Although the culture induced at an absorbance at 600nm of 0.8 yielded the most zinc finger protein (Fig. 4.3). An induction point between 0.6 and 0.7 was selected (and used in subsequent experiments) as it seemed to provide the optimum balance between the amount of host and zinc finger proteins produced.

### 4.2.3 Effect of point at which Rifampicin is added.

The point after induction at which rifampicin is added will effect the yield of the protein of interest. If it is added to soon after induction there will not be enough T7 polymerase available for efficient transcription of ZF-Gene 3.7. If the rifampicin is added too late the amount of host protein present in the lysates will be greater than necessary. Rifampicin was therefore added to cultures at various points after induction, the amount of zinc finger and host proteins produced were examined.

The zinc finger gene was expressed in *E. coli* BL21[DE3] pLysS (2.11.1) and the point at which the Rifampicin was added was varied. The gene was expressed in four 30ml cultures, the Rifampicin was added at 5 min, 10min, 15 min and 20min after induction. Samples (0.5ml) of the culture were removed before induction, 30 mins, 1 hr and 2 hrs after induction. The samples were examined by gel electrophoresis (Fig. 4.4).

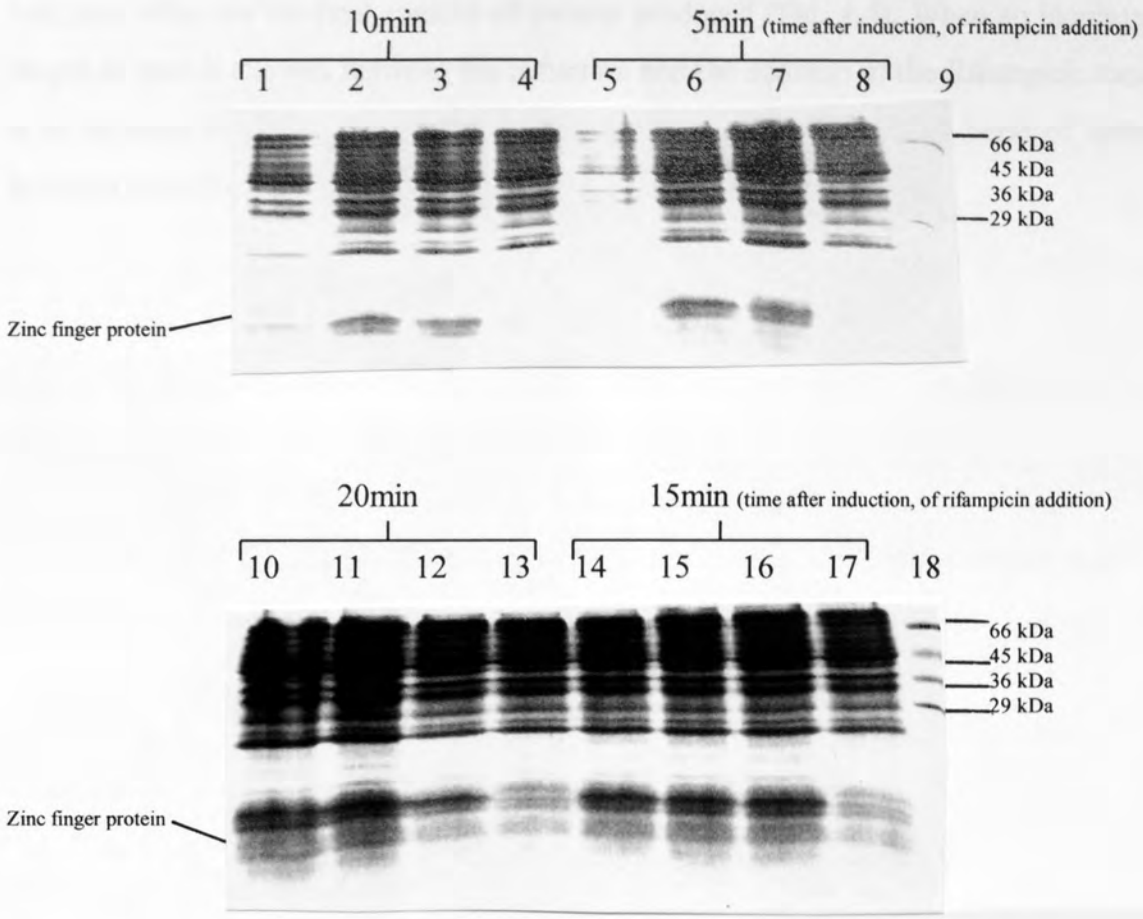



Figure 4.4 SDS-PAGE analysis to examine the effect on expression of the time at which rifampicin is added. Total cellular protein was prepared as described and electrophoresed on a 18% SDS-PAGE gel (2.5.3). Lanes 1-4 rifampicin added 10 min after induction. 4) Un-induced lysate; 3) 0.5hrs after induction; 2) 1hr after induction; 1) 2hrs after induction. Lanes 5-8 rifampicin added 5 min after induction. 8) Un-induced lysate; 7) 0.5hrs after induction; 6) 1hr after induction; 5) 2hrs after induction. 9) Low molecular weight marker (10μg, Sigma). Lanes 10-13 rifampicin added 20 min after induction. 13) Un-induced lysate; 12) 0.5hrs after induction; 11) 1hr after induction; 10) 2hrs after induction. Lanes 14-17 rifampicin added 15 min after induction. 17) Un-induced lysate; 16) 0.5hrs after induction; 15) 1hr after induction; 14) 2hrs after induction. 18) Low molecular weight marker (10μg, Sigma).

The time allowed between induction and the addition of Rifampicin to a culture had little effect on the final amount of protein produced (Fig. 4.4). When an increased length of time is allowed between the induction and the addition of the Rifampicin there is an increase in the amount of cellular proteins in the lysates. A time lapse of 5mins between induction and the addition of Rifampicin was therefore selected.

## 4.3 PILOT PROTEIN PURIFICATIONS.

### 4.3.1 Pilot Purification of protein by metal affinity.

After expressing the gene it was necessary to purify the zinc finger protein. His-ZF was initially purified by metal affinity chromatography. The histidine tag on the N-terminus of the protein has a high affinity for divalent cations such as $Ni^{2+}$. The zinc finger protein should therefore bind specifically to a metal chelating column that has been charged with a suitable ion. Pilot experiments were carried out using a 1ml HiTrap chelating column (Amersham Pharmacia Biotech).

His-ZF was expressed in a 30ml culture (2.11.1). Protein from 10ml of this culture was purified (2.12.2). The purity of the zinc finger protein was also determined by gel electrophoresis. The zinc finger protein was purified to apparent homogeneity, although on close inspection the sample was contaminated with small quantities of large proteins (Fig. 4.5). Due to low yield a larger scale purification method was required.

### 4.3.2 Pilot Protein purification by Anion Exchange.

The zinc finger protein appears to be larger on the SDS gel than would be expected (Fig. 4.5). This is because it contains an unusually high number of amino acids with a high pKa and therefore carries a positive charge which retards its migration through the gel. Due to its amino acid composition His-ZF is likely to be positively charged even at a high pH. At high pH the majority of the proteins in the cell lysate will be un-charged or negatively charged. This purification method relies on the high probability that at pH 8.8 the majority of the cellular proteins in the lysate will bind to the anion exchange media. The zinc finger protein however does not bind to the anion exchange column at this pH. The zinc finger protein therefore flows though during column washing.

The protein was expressed in a 100ml culture (2.11.1). The culture (20ml) was purified (2.12.2). A sample of the purified protein was examined by SDS PAGE. The zinc finger protein was purified to apparent homogeneity, but, again, on close inspection small quantities of contaminating large proteins were present in the sample (Fig 4.5).

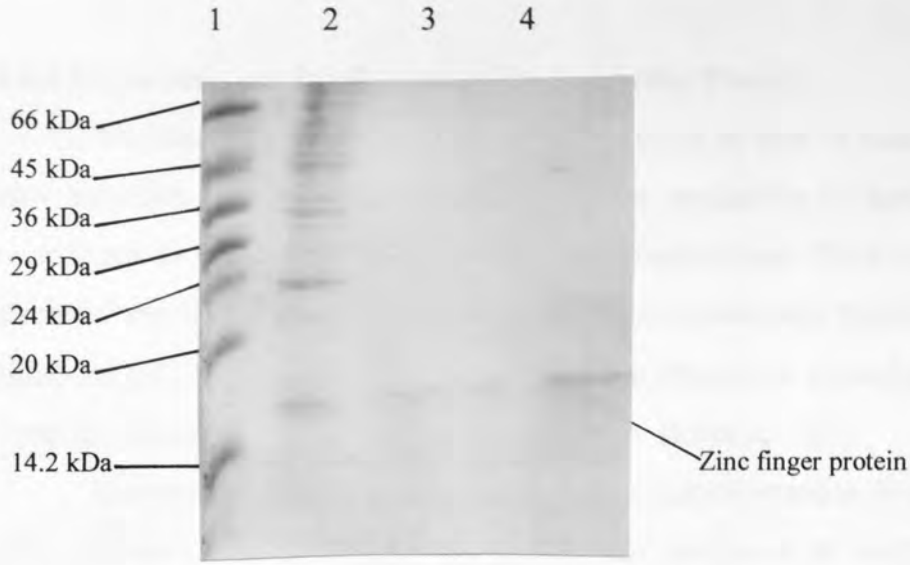



Figure 4.5 Gel showing samples of the zinc finger protein purified by metal affinity and anion exchange. 18% SDS-PAGE gel (2.5.3), lanes: 1) Low molecular weight marker (Sigma 10μg. 2) Induced crude lystate (30μl); 3) Zinc finger protein purified by metal affinity (3% of the column elutate); 4) Zinc finger protein purified by anion exchange (0.6% of the column elutate).

### 4.3.3 Removal of contaminating proteins purified zinc finger protein.

Zinc finger protein purified by either method consistently contained a small quantity of contaminating larger proteins. These were faintly visible and photographed poorly. To remove contaminating proteins the purified material was passed through a Microcon microconcentrator (Amicon) with a 50kda molecular weight cut off. The protein purified by anion exchange could not be detected in the liquid that passed through the microconcentrator. The protein purified by the metal affinity method however passed through the microconcentrator and when examined on a SDS-PAGE gel (2.5.3) the larger proteins had apparently been removed. Data not shown, as low concentrations of proteins were obtained and the SDS PAGE gels photographed poorly. However purified labelled protein could be analysed by SDS PAGE followed by auto-radiography (Fig. 2.5).

Protein for use in determination of protein function was therefore purified using the metal affinity method followed by the removal of contaminating larger proteins using a microconcentrator.

4.4 SYNTHESIS OF PROTEIN FOR USE IN ASSAYS.

### 4.4.1 Expression and Purification of Radiolabelled Protein.

The zinc finger gene was expressed (2.11.1) in four 1l cultures. Immediately after induction those cultures to be used for the production of labelled protein were centrifuged at 6,000 rpm for 5 mins at room temperature (Beckman). The LB was discarded and replaced with 100ml of minimal media containing 50µg/ml Rifampicin and either 200µCi [$^{3}$H] Amino acid mix (Amersham Pharmacia Biotech) or 200µCi [$^{35}$S] Cysteine (Dupont NEN). The cleared lysates were stored at -70°C.

Samples of the cultures were taken before induction and at 30 mins 1hr and 2hrs after induction. The labelling of the protein was monitored by electrophoresis of the [$^{35}$S] Cysteine labelled protein, followed by autoradiography of the dried gel (Fig 4.6).

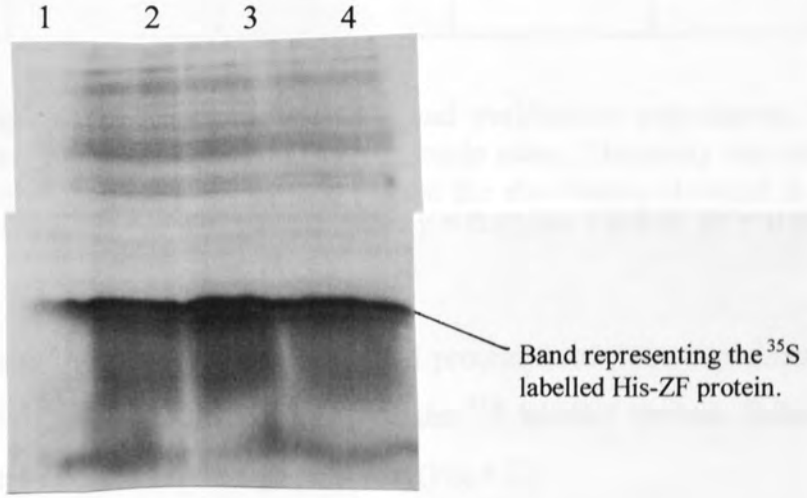



Figure 4.6 Autoradiograph (2.4.4) of a SDS PAGE gel (2.5.3) showing the $^{35}$S pulse labelling of His-ZF. Lanes; 1) Un-induced (30µl of lysate); 2) 0.5hrs after induction (30µl of lysate); 3) 1hr after induction (30µl of lysate); 4) 2hr after induction (30µl of lysate).

The cleared lysates were purified by metal affinity purification (2.11.2), and passed through a centriplus 50 concentrator (Amnicon) to remove the large contaminating proteins. The concentration (Table 4.1) of the protein was determined using a Bio-Rad protein assay (2.11.3).

In order to calculate the specific activity (Table 4.1) 1pmol of protein was placed in 0.5ml of scintillation fluid and counted on a scintillation analyser (2.4.5).

| Protein sample | Amount of protein in the crude lysate | Concentration of purified protein. | Final volume of protein obtained | Specific activity of protein |
|---|---|---|---|---|
| [$^{35}$S] labelled protein | 7,600µg | 5.25pmol/µl | 1ml | 49.19cpm/pmol |
| [$^{3}$H] labelled protein | 5,900µg | 3.53pmol/µl | 2.65ml | 19.87cpm/pmol |
| Cold protein | 3,900µg | 3.44pmol/µl | 2.5ml | N.A. |

Table 4.1 Showing results of protein labelling and purification experiments. Protein concentrations were estimated using a Bio-Rad protein assay. The assay was calibrated using known concentrations of BSA plotted against the absorbance obtained at 595nm. The BSA calibration provided a straight-line graph $y = 0.0324x + 0.028$. $R^2 = 0.9854$.

The success of isolation of a single labelled protein from these experiments was monitored by SDS-PAGE of a small quantity of the $^{35}$S labelled protein, followed by drying the gel and exposing it to photographic film (Fig 4.7).

Single band representing the $^{35}S$ labelled purified His-ZF protein.

Figure 4.7 Autoradiograph (2.4.4) of a SDS PAGE gel (2.5.3) showing the $^{35}S$ labelled preparation of His-ZF.

## 4.5 CHARACTERISATION OF THE HIS TAGGED PROTEIN.

### 4.5.1 Preparation of zinc finger protein and target sequence DNA.

Target sequence (TS) DNA was synthesised by the annealing of a pair of complementary oligonucleotides TS1A (5' biotinylated) and TS1B (Fig. 5.1). For successful synthesis of the TS DNA, full length TS1A and TS1B must be the major oligonucleotide products present. The presence of incompletely synthesised oligonucleotides on annealing may create a mixed population of dsDNA products complicating the binding of DNA to protein. The oligonucleotides (1pmol) were labelled with $^{33}P$ using T4 polynucleotide Kinase (2.9). In order to assess the amount of incomplete oligonucleotide present the labelled oligonucleotides were electrophoresed on a 10% TBE sequencing gel (2.5.2). This demonstrated that there was very little oligonucleotide that was not full length present in the preparations (data not shown).

In order to demonstrate that the oligonucleotides annealed to each other, $^{33}P$ TS1A (1pmol) was electrophoresed on a non-denaturing gel along side 1pmol of both $^{33}P$ TS1A annealed to TS1B and $^{33}P$ TS1B annealed to TS1A. As expected, a shift in the band position was observed on annealing (Fig 4.8).

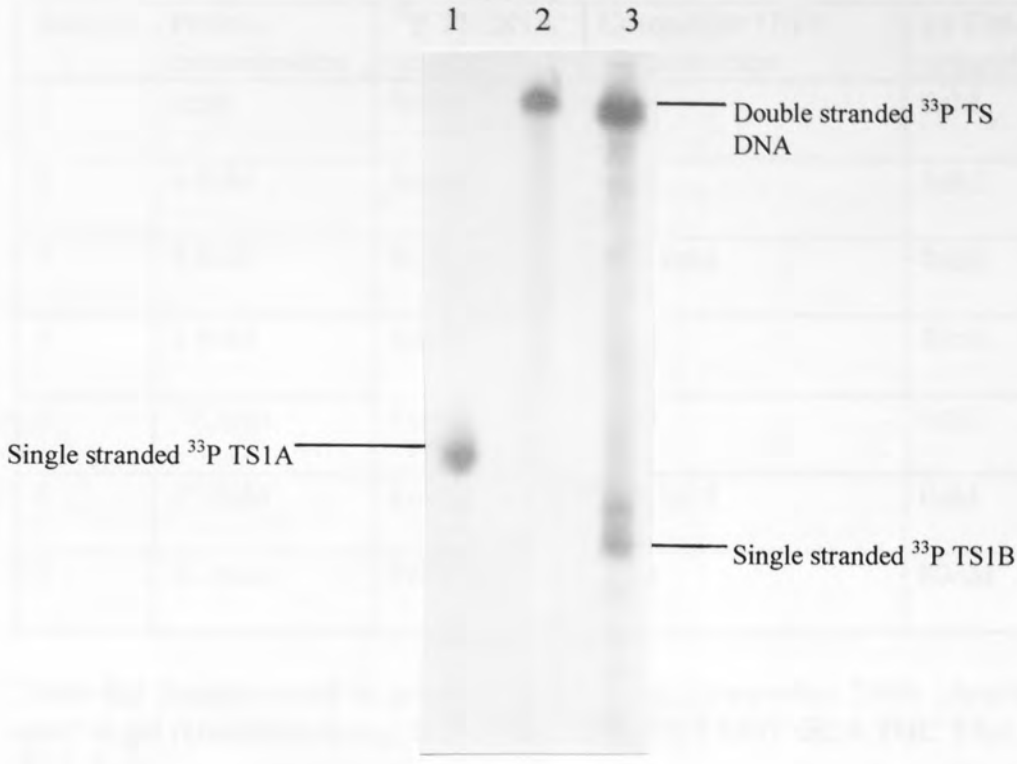



Figure 4.8 Autoradiograph of a non-denaturing PAGE gel (2.5.2). Showing results of oligonucleotide annealing experiment. Lanes: 1) $^{33}$P TS1A; 2) $^{33}$P TS1A annealed to TS1B; 3.)$^{33}$P TS1B annealed to TS1A

The preparation of the double stranded target sequence DNA was assessed by running the labelled DNA on a native acrylamide gel (Fig. 4.8). TS1B has labelled more efficiently than TS1A. This is probably due inhibition of the T4 Kinase by the 5' biotin on TS1A. There seems to be little oligonucleotide present that is not full length. and the oligonucleotides anneal to each other causing their migration though the gel to be retarded. The double stranded target sequence DNA seems to be sufficiently complete for use in the detection of protein binding.

### 4.5.2 Demonstration of DNA binding by His-ZF.

In order to determine whether or not the zinc finger bound the DNA sequence to which it was designed to bind, a gel shift assay was carried out (Fig. 4.9). Samples electrophoresed on the gel contained the reagents shown in Table 4.2.

| Sample | Protein concentration | $^{33}$P TS DNA concentration | Competitor DNA concentration | TS DNA concentration |
|---|---|---|---|---|
| 1 | 0nM | 8nM | 0nM | 0nM |
| 2 | 8.6nM | 8nM | 0nM | 0nM |
| 3 | 8.6nM | 8nM | 229.5nM | 0nM |
| 4 | 8.6nM | 8nM | 0nM | 80nM |
| 5 | 17.2nM | 8nM | 0nM | 0nM |
| 6 | 17.2nM | 8nM | 229.5nM | 0nM |
| 7 | 17.2nM | 8nM | 0nM | 80nM |

Table 4.2 Samples used in gel retardation assay. Competitor DNA (double stranded) used in gel retardation assay; 5' C TAG AGA TCT CGT GCA TGC TAA TGA TAT TCT T 3'.

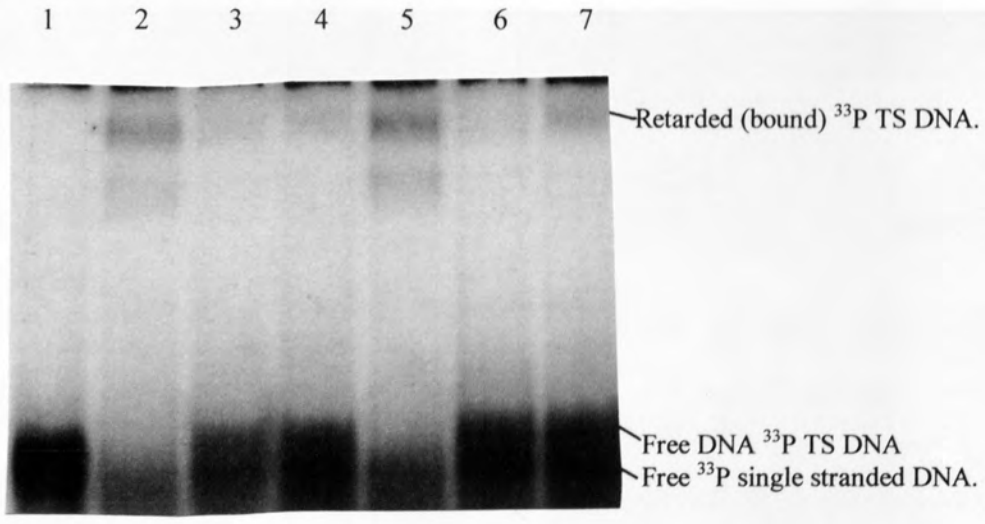



Figure 4.9 Autoradiograph showing results of gel shift experiment (2.5.5). Content of samples 1-7 shown in Table 4.2.

The gel retardation assay demonstrates that His-ZF binds to the target sequence DNA causing a band shift (Fig. 4.7). A band shift is observed in both the samples containing $^{33}P$ TS DNA and protein (8.7nM and 17.2nM) as expected. The band shift is not seen at 8.7nM on the addition of a 10x molar excess of non-labelled TS DNA, this is also expected as the cold DNA will compete with the hot DNA for protein. There is still a faint shift however at 17.2nM in the presence of cold TS DNA, due to the fact that there is more protein present. The band shift is also lost in the presence of a unlabelled competitor sequence DNA, this would not be expected if the binding is specific.

## 4.6 DISCUSSION.

### 4.6.1 Expression of ZF-Gene 3.7 and purification of the protein product (His-ZF).

A system for the expression of the ZF-Gene 3.7 was selected. *E. coli* strains containing the pLysS plasmid (*E. coli* BL21[DE3] pLysS and *E. coli* HMS174[DE3]) pLysS were selected as their growth was more rapid than those not containing the pLysS plasmid (Fig. 4.1). The slow growth of *E. coli* strains not containing the plasmid probably results from low level of expression ZF-gene 3.7, indicating that the His-ZF is toxic to the host cells.

The level of expression of the protein in the strains *E. coli* BL21[DE3] pLysS and *E. coli* HMS174[DE3]) pLysS was similar (Fig. 4.2). However the zinc finger protein is more readily visible in the lysates obtained from *E. coli* BL21 [DE3] pLysS. The expression of the gene was therefore optimised using this BL21 [DE3] pLysS strain of *E. coli*. Selected conditions were: induction of the culture at an absorbance of 0.6 at 600nm (4.2.2), addition of rifampicin 5mins after induction (4.2.3) and expression of the gene for 2hrs after induction (4.1.2).

The resulting zinc finger protein was purified using two techniques, anion exchange and metal affinity chromatoghraphy. Both methods of purification yielded what seemed to be reasonably pure protein. Unfortunately the purified zinc finger protein was recovered at low concentrations.

In the case of the metal affinity purified protein small amounts of contaminating protein can be removed by the use of microconcentrators. The anion exchange purified protein failed to pass through the microconcentrators. This could be an effect of the high pH of the buffer containing the protein. It is possible that at the high pH the membrane or protein were altered such that the protein is unable to pass though the membrane.

Using the expression and purification method developed here, His-ZF labelled by the incorporation of both $^{35}S$ cysteine and $^{3}H$-labelled amino acid mix was synthesised (4.4). Labelled protein is necessary in order to detect DNA / protein interactions using SPA.

### 4.6.2 Characterisation of the His tagged zinc finger protein.

The oligonucleotides used for the synthesis of the TS DNA and the TS DNA itself were analysed by gel electrophoresis. This demonstrated that DNA suitable for the detection of DNA protein interactions had been synthesised by the annealing of the complementary oligonucleotides.

The binding of the protein to the TS DNA was demonstrated by gel retardation assay. The gel shift showed that the protein bound to the labelled TS DNA. The protein also apparently binds to a competitor DNA sequence which was used in this assay. However the competitor is present at approximately 28.5x molar excess rather than 10x, and it is possible that the protein can bind to some extent to this sequence. It is therefore not too surprising although a little disappointing that the protein binds to this sequence.

## 4.6. SUMMARY.

- A suitable host for the expression of the zinc finger gene has been selected.

- The expression system was optimised.

- The zinc finger protein was purified using two methods. Using the metal affinity method it was possible to produce a reasonably pure protein product.

- Radiolabelled zinc finger protein was synthesised and purified for use in assay development.

- The protein was shown to be capable of binding to its DNA target sequence.

# 5 DEVELOPMENT OF AN ASSAY FOR THE DETECTION OF DNA / HIS-ZF INTERACTIONS.

## 5.1 INTRODUCTION.

The scintillation proximity assay (SPA, Amersham Pharmacia Biotech) is a dynamic assay allowing interactions to be studied in "real time". SPA beads are supplied with various coatings allowing (in this instance) either DNA or the zinc finger protein to be bound to the bead. The beads contain scintillant. Therefore, binding of a ligand, (tagged with a short emission length radiolabel) to the DNA or protein immobilised on the bead, can be detected in a scintillation counter. Two types of SPA bead (with the various coatings) are available: yttrium silicate and PVT. Where possible it is advantageous to use the PVT beads rather than the yttrium silicate beads, as there is generally a higher level of non-specific binding of both DNA and proteins to yttrium silicate beads (Amersham Pharmacia Biotech, Personal Communication).

It is possible to link biotinylated DNA to streptavidin coated beads. This assay method requires the protein to be radio-labelled in order to detect its binding to the DNA/SPA bead complex (Fig. 5.1)

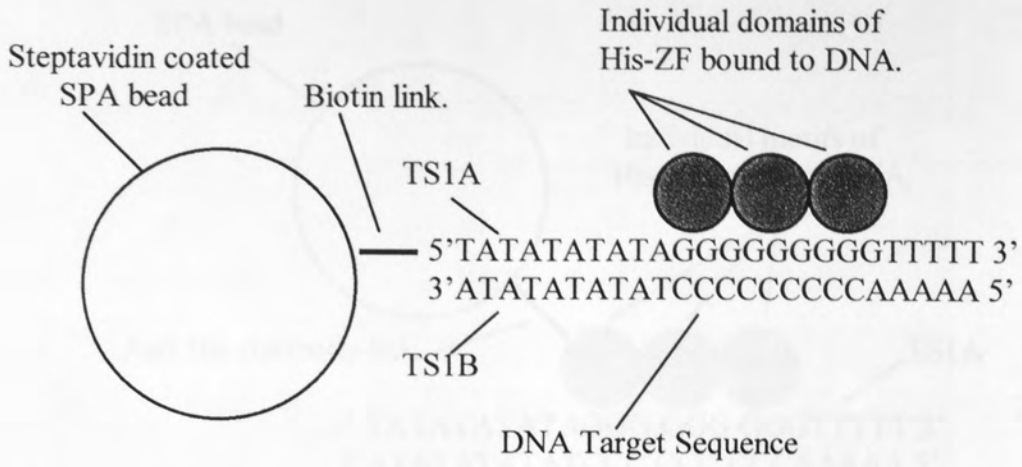



Figure 5.1 SPA method 1. The Biotinylated DNA is linked to streptavidin coated SPA beads. The zinc finger protein may be labelled with either $^{3}H$ or $^{35}S$. On binding the target DNA sequence, the labelled amino acids contained within the protein will be brought into sufficiently close contact to the SPA bead to induce a scintillation event.

Alternatively the zinc finger protein can be linked to protein A-coated SPA beads preferably using an antibody. The binding of labelled DNA to the protein/SPA bead complex could then be detected by scintillation counting (Fig 5.2).

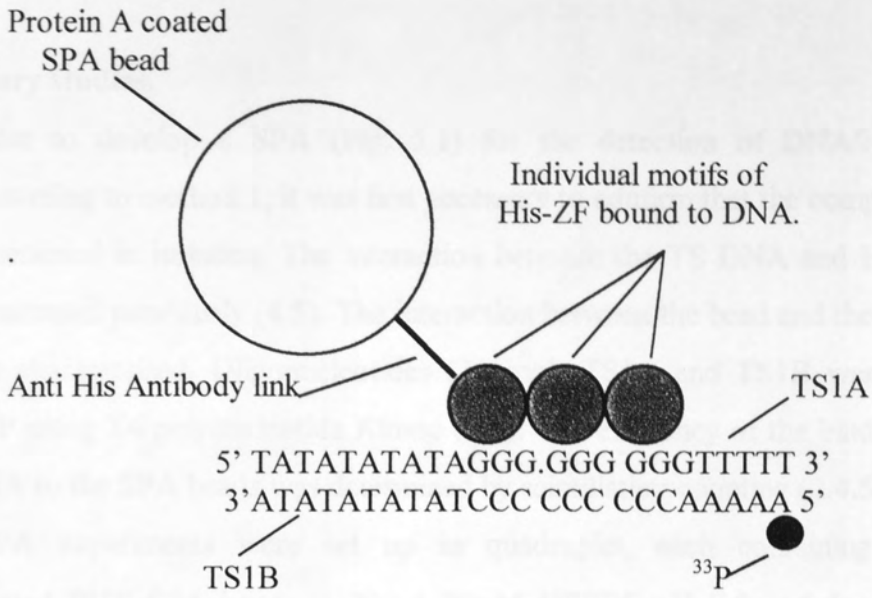



Figure 5.2 SPA method 2. The zinc finger protein is bound by an anti-His antibody, which in turn is bound by protein A on the surface of the SPA bead. Binding of $^{33}$P-labelled DNA to the zinc finger protein will bring the radionucleotide into close enough proximity of the bead to elicit a scintillation event.

5.2 SPA METHOD 1.

**5.2.1 Preliminary studies.**

In order to develop a SPA (Fig. 5.1) for the detection of DNA/His-ZF interactions, according to method 1, it was first necessary to confirm that the component interactions functioned in isolation. The interaction between the TS DNA and His-ZF had been demonstrated previously (4.5). The interaction between the bead and the DNA alone was first characterised. Oligonucleotides (20pmol) TS1A and TS1B were end labelled with $^{32}P$ using T4 polynucleotide Kinase (2.9). The efficiency of the binding of the labelled DNA to the SPA beads was determined by scintillation counting (2.4.5).

Five SPA experiments were set up in quadruplet, each containing 1mg streptavidin coated PVT SPA beads in 200µl 25mM HEPES pH 7.9 and 1pmol of DNA. Binding was allowed to take place at room temperature with gentle agitation for 1hr. The assays were then centrifuged at 6500r.p.m. for 2 min, the supernatant discarded and the beads resuspended in 25mM HEPES pH 7.9. This washing was repeated three times to remove non-specifically bound DNA and free label. The samples were centrifuged again then the microfuge tubes placed in scintillation vials which were counted to determine the amount of bound DNA. The results are illustrated in table 5.1.

| Experiment. | Components | Mean cpm /assay |
|---|---|---|
| Control | Un-incorporated label | 59.1 |
| Positive control | $^{32}P$ TS1A (1pmol/assay, biotinylated) | 997.7 |
| Negative control | $^{32}P$ TS1B (1pmol/assay, non-biotinylated) | 641.5 |
| Annealing experiment1 | TS1A annealed to $^{32}P$ TS1B (1pmol/assay total) | 2573.1 |
| Annealing experiment 2 | $^{32}P$ TS1A annealed to $^{32}P$ TS1B (1pmol/assay total) | 3493.4 |

Table 5.1 Binding of labelled biotinylated and non-biotinylated oligonucleotides to streptavidin coated PVT SPA beads. Assays performed in quadruplet (appendix VI).

The positive control in which the biotinylated $^{32}P$ TS1A was combined with the beads gave surprisingly low counts (997.7cpm/assay). This is probably (taking into consideration difference in the labelling of TS1A and TS1B in Fig. 4.8) due to in-efficient labelling of TS1A, possibly caused by the inhibition of the T4 polynucleotide kinase by the 5' biotin. Unfortunately, since the specific activity of the DNA was not measured the efficiency of labelling cannot be determined. However the sample in which $^{32}P$ TS1B was added alone demonstrates that this oligonucleotide alone does not bind efficiently to the SPA beads (642.5cpm/assay). The counts obtained when $^{32}P$ TS1B was pre-annealed to TS1A (2573.1cpm/assay) were far greater than those obtained with $^{32}P$ TS1B alone. Similarly the counts obtained when $^{32}P$ TS1B was pre-annealed to $^{32}P$ TS1A were, as expected, approximately the sum of those obtained from $^{32}P$ TS1B pre-annealed to TS1A and $^{32}P$ TS1A alone (3493.4cpm/assay).

The results of the annealing experiment show that the biotinylated oligonucleotide interacts with the streptavidin-PVT beads and indicate that the oligonucleotides TS1A and TS1B are annealing efficiently.

The amount of SPA beads required to interact optimally with 1pmol of oligonucleotide in a in 200µl reaction was then determined by varying the amount of beads added to the reaction. TS1A was labelled with $^{32}P$ using T4 polynucleotide kinase (2.9). SPA assays were performed with ratios of bead to oligonucleotide ranging from 0.1mg bead/pmol DNA to 10mg bead /pmol. DNA Controls contained 1mg SPA beads and un-incorporated label. Binding was allowed to take place at room temperature with gentle agitation for 1hr in 25mM HEPES pH7.9. The assays were then centrifuged at 6500rpm for 2 min and resuspended in 25mM HEPES pH7.5. This washing was repeated three times to remove non-specifically bound DNA and free label. Samples were centrifuged a final time then the microfuge tubes placed in scintillation vials and counted. The results are illustrated in figure 5.3.

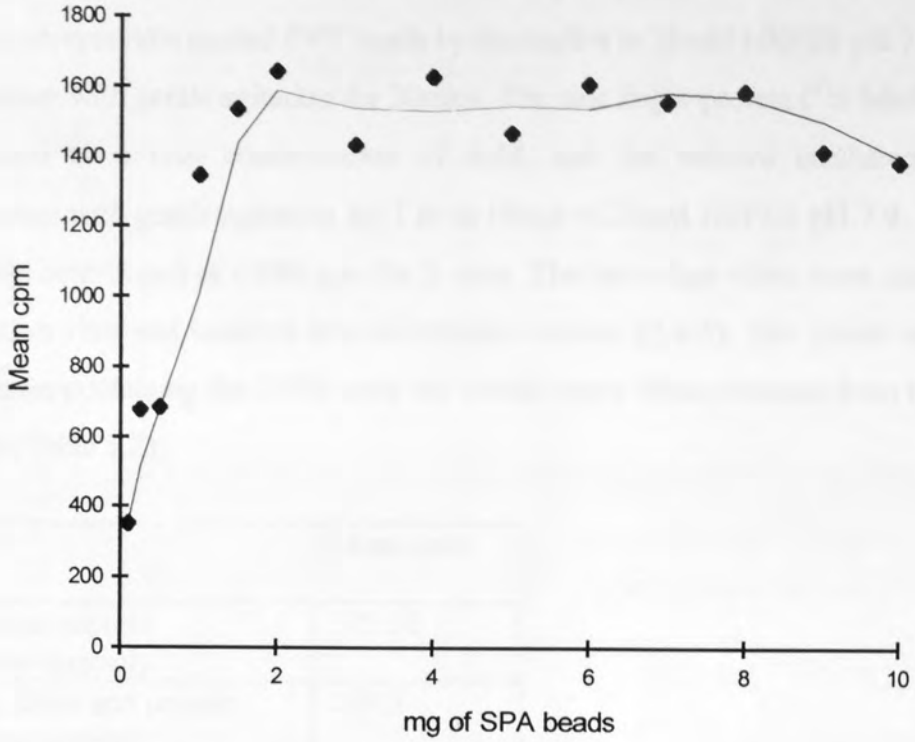



Figure 5.3 Graph to show cpm versus quantity of streptavidin -PVT SPA beads in the presence of 1pmol $^{32}$P-labelled TS1A/TS1B complex. Assays performed in quadruplet, in a total volume of 200µl. Full data shown in appendix VII.

A quantity of 1mg of SPA beads / pmol of DNA was selected for future assays, since this quantity of SPA beads gives virtually optimal counts (Fig. 5.3) and is convenient to handle.

### 5.2.2 SPA using method 1.

An attempt was made to detect the interaction between His-ZF and the TS DNA using SPA method 1 (Fig. 5.1). This should have been possible since interaction between the both the DNA and the protein (2.3.3) and the DNA and the SPA beads (5.2.1) had been demonstrated.

Experiments were prepared in duplicate. Negative controls contained no DNA allowing the level of non-specific interaction between the protein and the beads to be estimated. The DNA was first equilibrated with the beads. The His-ZF protein was then allowed to interact with the complex.

TS DNA (TS1A annealed to TS1B, 2.6.2) to a final concentration of 8nM was bound to streptavidin coated PVT beads by incubation in 25mM HEPES pH 7.9 at room temperature with gentle agitation for 30mins. The zinc finger protein ($^{35}S$ labelled, 4.4.1) was added to a final concentration of 8nM, and the mixture incubated at room temperature with gentle agitation for 1 hr in 100µl of 25mM HEPES pH 7.9. The beads were then centrifuged at 6500r.p.m.for 2 mins. The microfuge tubes were placed inside scintillation vials and counted in a scintillation counter (2.4.5). The counts detected in the samples containing the DNA were not raised above those obtained from the control samples (Table 5.2).

| Assay | Mean cpm |
|---|---|
| Beads and protein (negative control) | 305.65 |
| Beads, DNA and protein (positive control) | 269.3 |

Table 5.2 SPA assay results obtained from pre-coupling beads and DNA. Total cmp/assay = 1658, full data shown in appendix VIII.

Failure of the assay to detect the interaction between His-ZF and the DNA could be due to interference of the binding caused by the bulk of the SPA bead. If this is the case it may be advantageous to pre-couple His-ZF to the DNA, rather than the DNA to the beads.

This was achieved by incubating the DNA and protein together before adding the beads. Assays were performed in duplicate, negative controls contained no DNA. TS DNA (TS1A annealed to TS1B, 2.6.2) to a final concentration of 8nM was combined with the zinc finger protein ($^{35}S$ labelled, 4.4.1) to a final concentration of 8nM, the mixture was incubated at room temperature for 30 mins in 25 mM HEPES pH 7.9. The complex was then bound to streptavidin coated PVT beads by incubation in 25mM HEPES pH 7.9 (total volume of 100µl) at room temperature with gentle agitation for 1 hr. The beads were then pelleted in a microcentrifuge at 6500r.p.m. for 2 mins. The microfuge tubes were placed inside scintillation vials and counted in a scintillation counter (2.4.5). Again the number of counts observed in the samples containing DNA was not raised above those in the negative control samples (Table 5.3).

| Assay | Mean cpm |
|---|---|
| Beads and protein (negative control) | 134.55 |
| Beads, DNA and protein (positive control) | 106.05 |

Table 5.3 SPA assay results obtained from pre-coupling protein and DNA. Total cmp/assay = 1597, full data shown in appendix VIII.

Pre-coupling the DNA and the protein failed to demonstrate the DNA protein interaction. Failure of the assay might still be due to steric hindrance caused by the bead. The low specific activity of the protein may also mean that any interaction which is occurring between the DNA and His-ZF cannot be detected.

## 5.3 SPA METHOD 2.

### 5.3.1 Preliminary studies.

Failure of the assay to detect the interaction between the protein and the DNA using method 1 could be due to the low specific activity of the protein or hindrance from the SPA bead. Linking the protein to the beads increases the distance between the bead and the interacting molecule (DNA) and labelling of DNA to a high specific activity is easily achieved. It should therefore be possible to overcome the problems associated with SPA method 1 (Fig 5.1) by performing the assay using method 2 (Fig. 5.2).

Initially the method of linking His-ZF to the beads was investigated. With the labelled protein it had previously been demonstrated that the protein was capable of binding directly to the yttrium silicate beads (data not shown), but as the binding was thought to be non-specific, use of this method may be unreliable. The use of an antibody which recognised the His tag on the protein would provide a more suitable method for attaching the His-ZF protein to protein A-coated SPA beads.

In order to identify an antibody capable of binding to the protein a dot blot was carried out using a Qiagen Anti-His selection Kit (2.11.4). Within each experiment the zinc finger protein, anti-His antibodies and conjugate were applied to the nitro-cellulose membrane. In all experiments the antibody and conjugate controls produced a positive result, which indicated that the conjugate bound the antibody and that the blot was being developed correctly. However none of the antibodies in the kit apparently recognised the protein.

To confirm that the antibodies did not recognise the His-ZF protein a scintillation proximity assay was carried out. Both yttrium silicate and PVT SPA (beads coated with protein A) were used as illustrated in figure 5.2.

Assays contained 25mM HEPES pH 7.9, 10mM DTT, 1mg/ml BSA, 1.5μg antibody, 1mg SPA beads and 10nM TS DNA. TS DNA was made by annealing (2.6.2) TS1A to TS1B labelled with $^{33}P$ using T4 Kinase (2.9). The reactions were equilibrated at room temperature for 1 hr then the beads pelleted in a microfuge at 6500r.p.m. for 2 mins. The microfuge tubes were placed in a scintillation vial and analysed in a scintillation counter (2.4.5) to give values for the background binding of the DNA. Protein to a final concentration of 52nM (final volume 100μl) was added to each sample and the beads resuspended by gentle vortexing and the sample re-equilibrated at room

temperature with gentle agitation for 1hr. The beads pelleted by brief centrifugation at 6500r.p.m. and the samples re-counted (Table 5.4). The amount of protein bound to the DNA / bead complex was predicted by scintillation counting (2.3.4). Total cpm/assay was determined by placing a 1pmol labelled DNA in scintillation fluid.

| Content of assay | Mean background cpm | Mean cpm in the presence of His-ZF | Difference in cpm due to addition of protein |
|---|---|---|---|
| Yttrium beads, no antibody | 7611.5 | 15818.5 | 8207.0 |
| Yttrium beads, RGS His antibody | 5736.0 | 7381.0 | 1618.0 |
| PVT beads, no antibody | 2559.0 | 2878.0 | 319.0 |
| PVT beads, RGS His antibody | 3687.5 | 3289.0 | -222.5 |
| PVT beads, Tetra His antibody | 5469.0 | 5849.5 | 380.5 |
| PVT beads, Penta His antibody. | 4633.0 | 4476.5 | -156.5 |

Table 5.4 Results obtained from testing the various formats of the SPA assay. Assays were performed in duplicate. Total cpm/assay = 28453, full data is shown in appendix VIII.

None of the antibodies tested (Table 5.4) were capable of inducing binding of the labelled DNA to the SPA beads on addition of His-ZF. This confirms the results obtained from the dot blot, which demonstrates the failure of the antibodies to recognise the His-ZF protein. However the zinc finger protein seems to be capable of inducing the binding of the DNA to the yttrium silicate SPA beads. Binding of the DNA to the yttrium silicate SPA beads / His-ZF complex is far more efficient in the absence of the antibodies than in their presence. This is likely to be due to the binding of the antibody to the SPA bead "blocking" the binding of His-ZF to the beads.

An assay was then performed in the apparently most successful format (Yttrium silicate beads and no antibody) with a range of protein concentrations. This indicated that there is a correlation between increased protein concentration and increased binding of DNA to beads (Fig. 5.4).

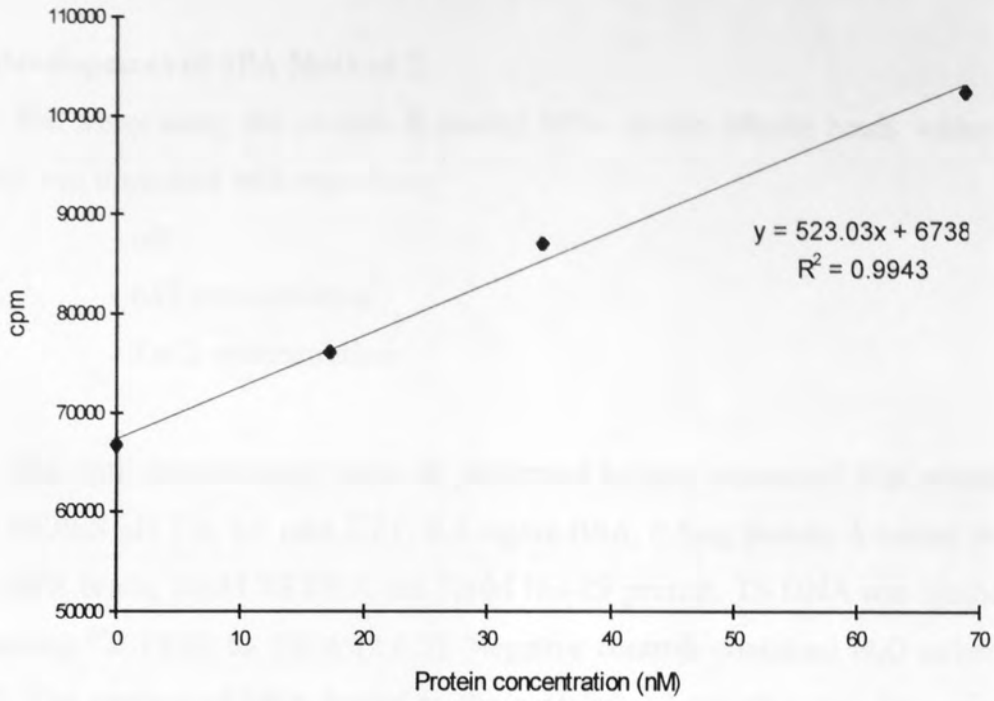



Figure 5.4 Effect on binding of DNA to yttrium silicate SPA beads of increasing concentrations of His-ZF. Concentration of $^{33}P$ labelled DNA 25.6nM/assay. Assays performed singularly, total cpm/assay = 847056.

These preliminary results suggested that the zinc finger protein was capable of binding to the yttrium silicate beads in a non specific manner, inducing binding of the DNA to the beads. Yttrium, like nickel, exists in solution as a divalent cation which may be used for the metal affinity purification of His tagged proteins. It was possible that the His Tag on the protein was capable of binding to yttrium ions immobilised in the SPA beads. This method of interaction of the zinc finger protein with the beads would allow the zinc finger motifs to freely interact with the DNA. Therefore it was likely that this method of attachment of His-ZF to the SPA beads was stable enough for use in the detection of protein-DNA interactions.

### 5.3.2 Development of SPA Method 2.

The assay using the protein A coated SPA yttrium silicate beads without the antibody was optimised with regards to;

- pH
- KCl concentration
- $ZnCl_2$ concentration

The optimisation assays were all performed in total volume of 50µl containing; 25mM HEPES pH 7.9, 0.5 mM DTT, 0.5 mg/ml BSA, 0.5mg protein A coated yttrium silicate SPA beads, 10nM TS DNA and 52nM His-ZF protein. TS DNA was synthesised by annealing $^{33}P$ TS1B to TS1A (2.6.2). Negative controls contained $H_2O$ rather than His-ZF. The amount of DNA bound to the protein/bead complex was determined by counting samples in a scintillation analyser (2.4.5). Total cpm/assay was measured by counting 1pmol of labelled DNA in 0.5ml of scintillation fluid. Data was statistically analysed using Microsoft Excel 5.0, all t-test unless otherwise stated were two tailed.

A range of pH from 6.8 to 8.0 was investigated in order to determine its effect on the binding of the DNA to the SPA beads (Fig. 5.5).

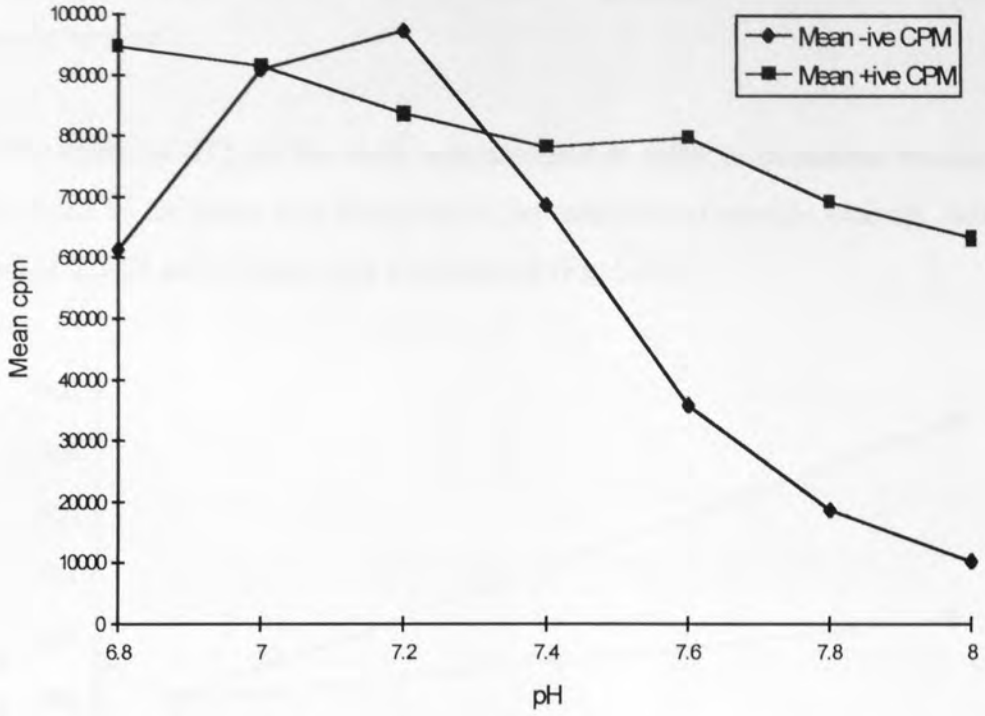



Figure 5.5 Graph showing the effect of pH on the binding of DNA to the SPA beads. Assays performed in quadruplet, +ive contains His-ZF protein,-ive contains no zinc finger protein. Total cpm/assay = 200576. Full data and statistical analysis shown in appendix IX.

The graph (Fig 5.5) indicates that there is a huge increase in non-specific binding of the DNA to the SPA beads between pH 7.0 and 7.4. A students *t*-test was applied to the data in order to determine the significance of the difference between assays containing His-ZF and the negative controls. At pH 6.8, 7.6, 7.8 and 8.0 there is significantly more binding of DNA to SPA beads (P-values 0.0166, 0.0007, $4.8 \times 10^{-6}$ and $1.76 \times 10^{-6}$ respectively) in assays containing His-ZF than in negative controls. At pH 7.0, 7.2 and 7.4 there is no significant difference (P-values 0.926, 0.102 and 0.285 respectively) between samples containing His-ZF and the negative controls. This shows that between pH 6.9 and 7.5 the non-specific binding of the DNA beads is equal to the specific binding.

If addition of the protein causes an alteration of the pH in the assay, pH change rather than the His-ZF protein could be responsible for the binding of the DNA to the SPA beads. For this reason the pH of the assay mixes (at pH 7.9) including and excluding His-ZF was estimated by applying a small quantity to pH paper. A pH of approximately 7.9 in both mixes was confirmed.

Buffer with a pH of 7.9 was retained in subsequent assays in order to avoid high non-specific binding.

The effect of KCl on the assay was assessed in order to determine whether the addition of salt to the assay was beneficial to the detection of specific binding. Addition of the salt at 2.5 25 and 250mM was investigated (Fig 5.6).

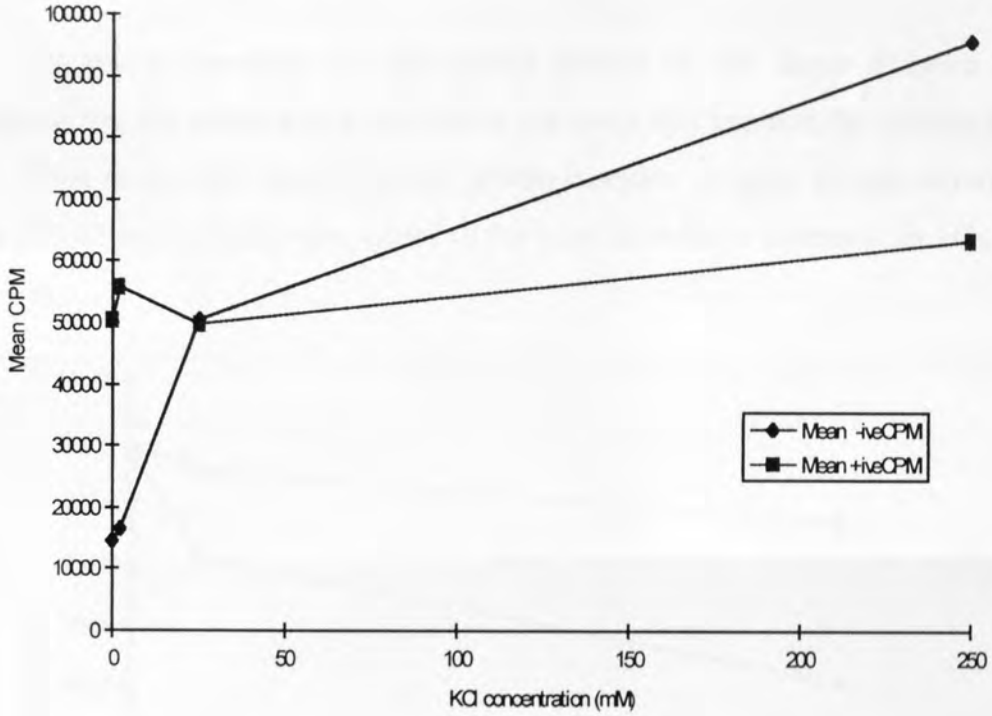



Figure 5.6 Graph illustrating the effect of KCl on the binding of DNA to the SPA beads. Assays performed in quadruplet, +ive contains His-ZF protein,-ive contains no zinc finger protein. Total cpm/assay = 200198. Full data and statistical analysis are shown in appendix X.

The graph (Fig. 5.6) indicates that the addition of KCl causes an increase in the non-specific binding of the DNA to the SPA beads. A students *t*-test was applied to the data in order to determine the significance of the difference between assays containing His-ZF and the negative controls. The binding of DNA to SPA beads in the presence of His-ZF was significantly higher (P-value $4.07 \times 10^{-5}$) than that in the negative control in the absence of KCl and in the presence of 2.5mM KCl (P-value $1.39 \times 10^{-5}$). At 25mM KCl there is no significant difference (P-value 0.642) between the samples containing His-ZF and the negative controls. At 250mM KCl there is significantly (P-value 0.0098) more binding of the DNA to the beads in the negative controls than in the assays

containing His-ZF. This shows that in the absence of KCl and in the presence of 2.5mM KCl the binding of the DNA to the SPA beads is greatly increased by the presence of the His-ZF protein. At higher concentrations of salt there is such a large increase in the non specific binding of the DNA to the SPA beads that the specific binding can no longer be detected.

There seemed to be no advantage associated with the addition of KCl to the assay, it was therefore omitted in subsequent experiments.

As zinc is necessary for the correct folding of zinc finger domains it was considered that the addition of a zinc salt to the assay may improve the specific binding of the DNA to the SPA bead / His-ZF protein complex. A range of concentrations of $ZnCl_2$ (2.5 25 and 250mM) were added to the assay in order to determine its effect (Fig. 5.7).

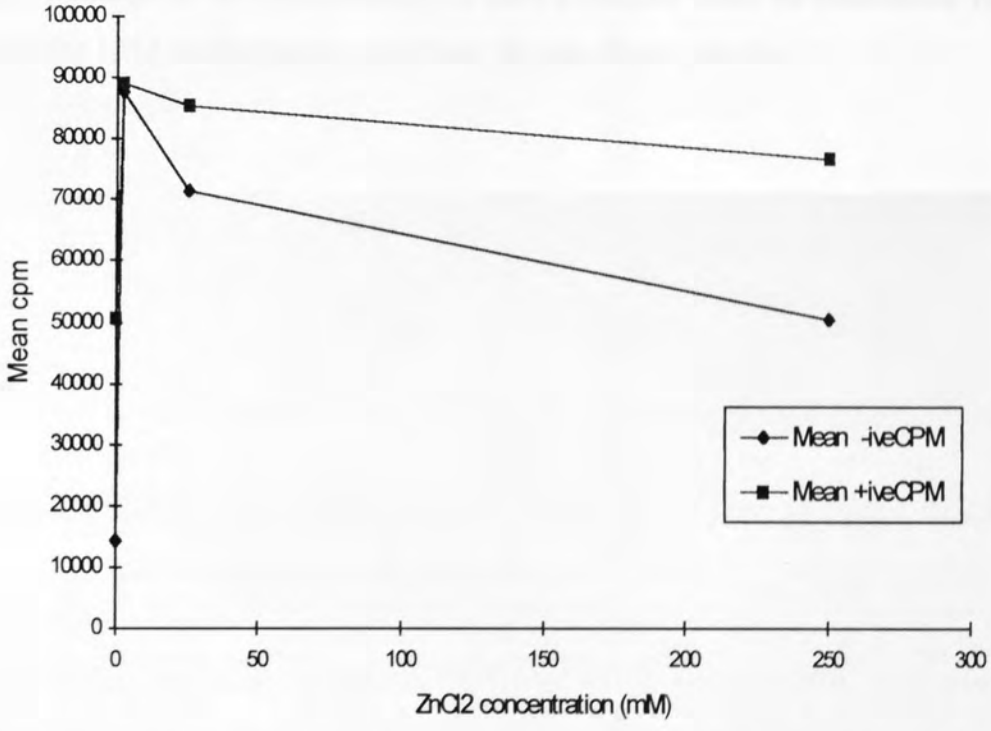



Figure 5.7 Graph demonstrating the effect of $ZnCl_2$ on the binding of DNA to SPA beads. Assays performed in quadruplet, +ive contains His-ZF protein,-ive contains no zinc finger protein. Total cpm/assay = 199848. Full data and statistical analysis shown in appendix XI.

The graph (Fig. 5.7) indicates that $ZnCl_2$ even at very low concentrations causes non specific binding of the DNA to the SPA beads. A students *t*-test was applied to determine the significance of the difference in quantity of DNA bound between assays containing His-ZF and the negative controls. In the absence of $ZnCl_2$ significantly more DNA is bound to the SPA beads in assays containing His-ZF than in the negative controls (P-value $4.07 \times 10^{-5}$). In the presence of 2.5mM $ZnCl_2$ there is no significant difference (P-value 0.823) between the assays. At both 25mM and 250mM the binding of the DNA to the SPA beads is significantly higher (P-values 0.0006 and 0.0003 receptively) in the assays containing His-ZF.

There is a massive increase in non-specific binding of the DNA to the SPA bead / His-ZF protein complex, observed on the addition of even small quantities of $ZnCl_2$ to the assay. This result is particularly concerning as small quantities of various reagents may be present in the protein buffer as a result of the protein purification method (2.11.2). A change in the concentration of such a reagent could be responsible for the binding of the DNA to the beads rather than the zinc finger protein.

## 5.4 INVESTIGATION INTO THE MODE OF BINDING OF DNA TO SPA BEADS.

### 5.4.1 Introduction.

The addition of $ZnCl_2$ to the assay even at low concentrations causes a huge increase in the non-specific binding of the DNA to the SPA beads (Fig 5.7). It was therefore conceivable that a reagent in the protein buffer (other than the His-ZF protein) was causing the DNA to bind non-specifically to the SPA beads.

In order to investigate this possibility the protein was removed from the buffer to determine whether or not the protein buffer alone was capable of causing the DNA to bind to the SPA beads. This confirmed that a factor in the buffer other than the zinc finger protein was responsible for the binding of the DNA to the beads (5.4.2).

Secondly the causative factor of the non-specific binding of DNA to SPA beads was investigated. The protein had been dialysed three times after purification to remove traces of reagents present in the purification buffers. Therefore anything remaining in the protein buffer must have been present at high concentration during purification. The only substances used during protein purification at high concentration were;

- KCl
- Imidazole
- $NiCl_2$

If one of these reagents is capable of at low concentrations mimicking the effect seen on the addition of His-ZF to the assay it would be a strong candidate for the component in the buffer causing the non-specific binding of the DNA to the SPA beads (5.4.3).

Finally the contaminant responsible for causing the non-specific binding of the DNA to the SPA beads was removed. The effect of protein on the binding of DNA to SPA beads was then determined in its absence (5.4.5).

### 5.4.2 Effect of protein buffer on binding of DNA to SPA beads.

In order to determine whether or not the protein was responsible for the binding of the DNA to the SPA beads it was removed from the buffer and the buffer alone used in an assay. Assays and data analysis were performed according to the protocol used for previous assays (5.4.1).

The protein was removed from the buffer using a Microcon 10 (microconcentrator, Amicon). The protein remained in the reservoir whilst the de-proteinated buffer passed through the membrane. To estimate the retention of the protein in the upper reservoir the experiment was initially performed using tritiated protein. The content of the lower reservoir was then monitored by scintillation counting to determine how much protein had passed though the membrane. Tritium labelled His-ZF was removed from a sample of the protein then 10µl of the resulting buffer was placed in scintillation fluid and the counts obtained compared to the counts produced by 10µl of tritiated protein. Mean cpm in sample containing protein = 1465.67, mean cpm in sample obtained from the lower reservoir of the microconcentator = 218.67 (assay performed in duplicate). In this way it was estimated in this way that approximately 85% of the protein was removed from the buffer.

The effect of the de-proteinated buffer on the binding of DNA to protein A coated yttrium silicate SPA beads was then assessed (Table 5.5).

| Assay | Mean cpm | Estimated protein concentration |
|---|---|---|
| $H_2O$ control | 16805.5 | None |
| Protein Buffer | 56126.5 | 10.8nM |
| Protein (His-ZF) | 54321.0 | 54.0nM |

Table 5.5 Effect of protein buffer on non-specific binding. The assay was performed using the same method described in 5.4.1. Protein concentration estimated using known concentrations of His-ZF (4.11.4) and the % removal or protein from the buffer by midcroconcentration (described above). Assays carried out in duplicate. Total cpm = 165131/assay, full data and statistical analysis shown in appendix XII.

There is not a significant difference (P value = 0.587) between the binding of the DNA to the beads in the presence of the protein and in the presence of the protein buffer. This shows that there is a factor in the protein buffer other than His-ZF capable of causing the non-specific binding of the DNA to the SPA beads.

**5.4.3 Identification of the causative agent of the non-specific binding.**

If one of the reagents likely to be present in the buffer (5.4.1) is capable (at low concentration) of mimicking the response caused by adding "protein" to an assay it would be a strong candidate for the component in the buffer causing the non specific binding. Assays and data analysis were performed according to the protocol used for previous assays (5.4.1).

Potassium chloride had already been shown to have little effect on the binding of DNA to the beads at low concentrations (5.3.2).

A range of imidazole concentrations was tested. For each concentration the effect was determined without protein, with protein buffer and with protein (Fig. 5.8).

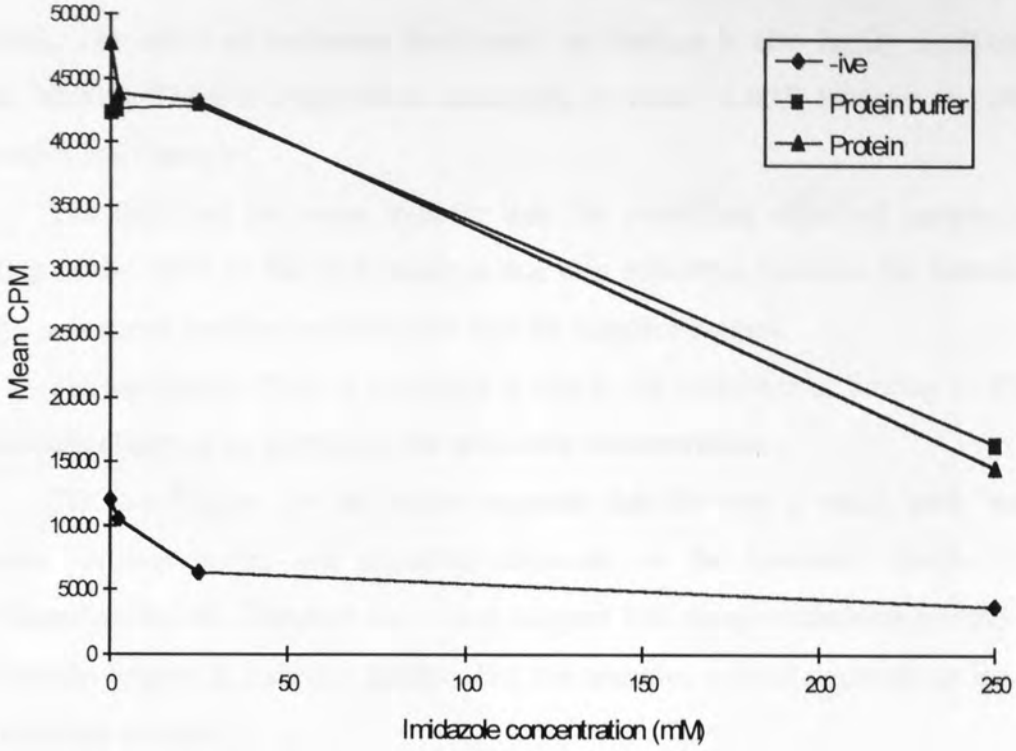



Figure 5.8 Graph illustrating the effect of imidazole on the binding of DNA to SPA beads. Assays performed containing the zinc finger protein, the protein buffer (with the protein removed by microconcentration). The -ive control contains neither protein or protein buffer. Assays carried out in quadruplet, cpm/assay = 171129.5. Full data and statistical analysis shown in appendix XIII.

A students *t*-test was applied to determine whether or not there was a significant difference between the binding of DNA to SPA beads in the presence of the protein buffer and the presence of the protein. In the absence of imidazole there is significantly (P-value 0.0048) more DNA bound to the beads in the presence of His-ZF than there is in its absence. This difference however is not large and was not observed in the previous experiment (Table 5.5), or in subsequent experiments (Figs. 5.9 and 5.10). This result is probably due to experimental error. At 2.5, 25 and 250mM imidazole there is no significant difference between the binding of DNA to SPA beads in assays containing His-ZF and the negative controls. This demonstrates that assays containing the protein and the protein buffer behave in a similar manner.

Analysis of variance revealed that the "sample" (protein, protein buffer and negative control) has a significant effect on the binding of DNA to SPA beads (P-value <0.005). The effect of treatment (imidazole) on binding is also highly significant (P-value <0.005). There is a significant interaction (P-value <0.005) between the effect of treatment and "sample".

The data and the *t*-test indicate that the significant effect of sample on the binding of the DNA to the SPA beads is due to a difference between the behaviour of assays containing protein/protein buffer and the negative control.

The significant effect of treatment is due to the reduction of binding of DNA to SPA beads observed on increasing the imidazole concentration.

The low P-value for interaction suggests that the way in which each "sample" (protein, protein buffer and negative) responds to the treatment (imidazole) are significantly different. The data and *t*-tests suggest that assays containing protein buffer and protein behave in a similar fashion. But the negative control responds to imidazole in a different manner.

The addition of imidazole to the assay does not cause the negative control to behave in a similar manner to the samples containing protein or protein buffer. It is therefore unlikely to be responsible for the non-specific binding of the DNA to the SPA beads.

The effect of $NiCl_2$ on the non-specific binding of DNA to beads was determined by the addition of 0.01, 0.1, 1.0 and 10mM $NiCl_2$ to the assays (Fig 5.9).

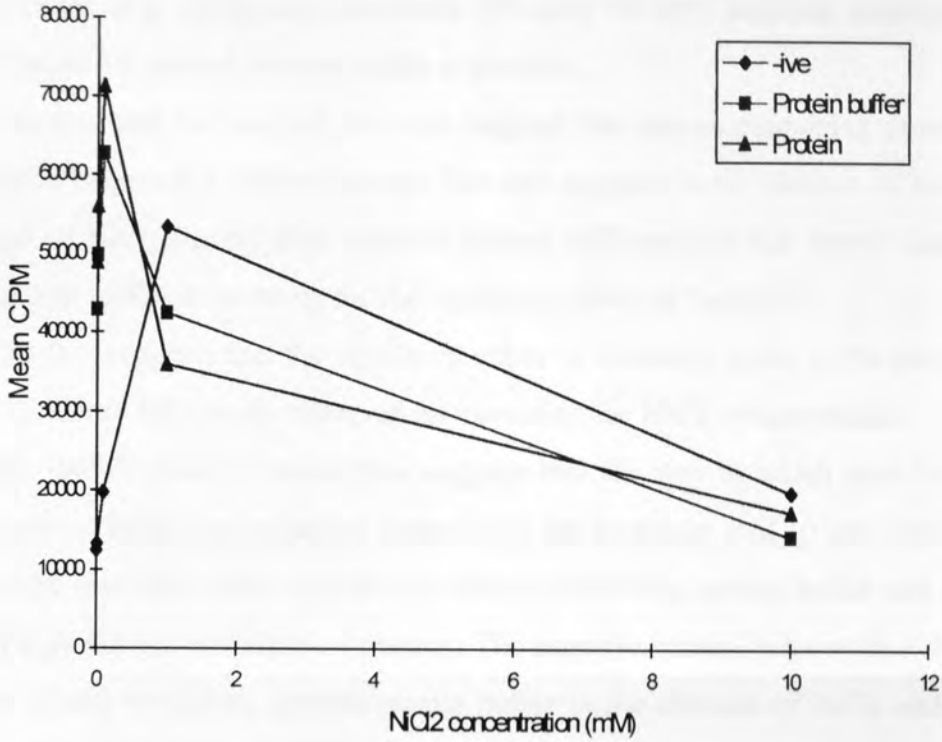



Figure 5.9 Graph showing the effect of $NiCl_2$ on the binding of DNA to SPA beads. Assays performed containing the zinc finger protein, the protein buffer (with the protein removed by microconcentration). The -ive control contains neither protein nor protein buffer. Assays carried out in quadruplet, total cpm/assay = 163324. Full data and statistical analysis shown in appendix XIV.

A students *t*- test was applied to determine whether or not there was a significant difference between the binding of DNA to SPA beads in the presence of the protein buffer and the presence of the protein. In the absence of $NiCl_2$ at 0.01mM and at 1.0mM $NiCl_2$ there is no significant difference (P-values 0.471, 0.254 and 0.078 respectively) between assays containing His-ZF and those containing only buffer. At 0.1mM $NiCl_2$ and 10mM $NiCl_2$ there is a significant difference (P-value 0.038) between the binding of DNA to SPA beads in assays containing His-ZF and the negative controls. But this difference is small and unlikely to represent an effect of His-ZF. The data suggests that

the assays containing protein and protein buffer in are behaving in a similar manner in response to $NiCl_2$.

Analysis of variance revealed that the effect of "sample" (protein, protein buffer and negative control) on the binding of DNA to SPA beads was significant (P-value <0.005). The effect of treatment ($NiCl_2$) on binding is also highly significant (P-value <0.005). There is a significant interaction (P-value <0.005) between treatment and "sample" (negative control, protein buffer or protein).

The data and the analysis by *t*-test suggest that assays containing protein and protein buffer behave in a similar manner. The data suggests in the absence of $NiCl_2$ and at 0.01mM of $NiCl_2$ the negative controls behave differently to the assays containing protein/protein buffer, accounting for the significant effect of "sample".

The data suggests that the significant effect of treatment is due to the increase of binding of DNA to SPA beads observed on increasing the $NiCl_2$ concentration.

The low P-value for interaction suggests that the way in which each "sample" (protein, protein buffer and negative) responds to the treatment ($NiCl_2$) are significantly different. The data and *t*-tests suggest that assays containing protein buffer and protein behave in a similar but not identical manner. The negative control behaves in a different manner to assays containing protein/protein buffer in the absence of $NiCl_2$ and in the presence of 0.01mM $NiCl_2$. At higher concentrations the data seems to suggest that the "samples" (protein/protein buffer and negative) are behaving in a similar manner.

These results seem to indicate that $NiCl_2$ in the protein buffer may be responsible for the non-specific binding of the DNA to the SPA beads, providing there is at least 0.1mM present in the assay. Although these results indicate that $NiCl_2$ may be responsible for the non-specific binding of the DNA to the SPA beads they are not entirely convincing.

In order to further demonstrate that $NiCl_2$ could be the cause of the non-specific binding the effect of a range of EDTA concentrations was assessed (Fig 5.10). EDTA is a strong chelating agent, it will "mop up" any metal ions present in the assay. If the non-specific binding of the DNA to the SPA beads is caused by nickel ions EDTA should dramatically reduce it.

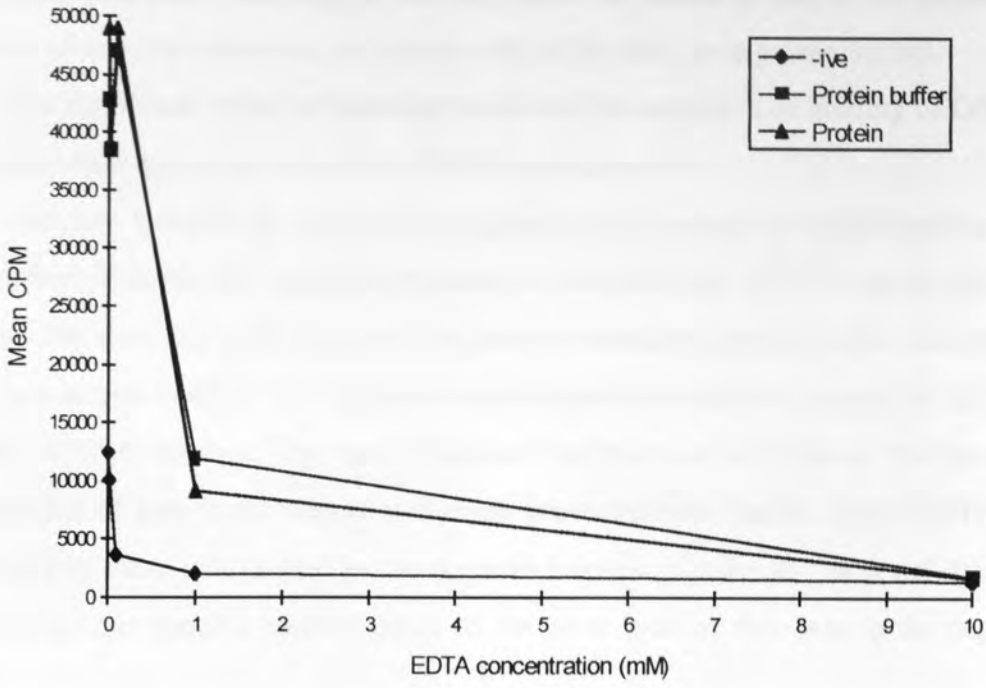



Figure 5.10 Graph illustrating the effect of EDTA on the binding of DNA to SPA beads. Assays performed containing the zinc finger protein, the protein buffer (with the protein removed by microconcentration). The -ive control contains neither protein or protein buffer. Assays carried out in quadruplet, total cpm/assay = 162891.4. Full data shown in appendix XV.

A students *t*-test was applied to determine whether or not there was a significant difference between the binding of DNA to SPA beads in the presence of the protein buffer and the presence of the protein. In the absence of EDTA and at 0.01mM, 0.1mM, and 10mM there is no significant difference (P-values 0.485, 0.968, 0.318 and 0.403 respectively) between the amount of DNA bound to the beads in the presence of His-ZF and that bound in assays containing protein buffer. At 1.0mM EDTA there is significantly (P-value 0.0197) more DNA bound to the beads in the absence of His-ZF than there is in its presence. This difference is however small and likely to be an experimental artefact. This data suggests that assays containing protein and protein buffer behave in essentially the same manner.

Analysis of variance revealed significant effect of "sample" (protein, protein buffer and negative control) on the binding of DNA to SPA beads (P-value <0.005). The effect of treatment (EDTA) on binding is also highly significant (P-value <0.005). There is a significant interaction (P-value <0.005) between the effects of sample and treatment.

The data and *t*-test suggest that the effect of sample is due to the differential behaviour of samples containing protein/protein buffer and the negative control.

The significant effect of treatment is due to the reduction of binding of DNA to SPA beads observed on increasing the EDTA concentration.

The low P-value for interaction suggests that the way in which each sample (protein, protein buffer and negative) responds to the treatment (EDTA) are significantly different. The data and *t*-test suggest that assays containing protein buffer and protein behave in a similar fashion. The negative control however seems to respond to EDTA in a slightly different manner. The results indicate that this may to be due to the increased concentration of ions in the assays containing protein/protein buffer. More EDTA will be required in these assays than in the negative control in order to "mop up" the ions and bring the non-specific binding down to the same level as that seen in the negative control.

The effect of EDTA on the assay supports the evidence that $NiCl_2$ present in the buffer as a result of the method of purification is responsible for the non-specific binding of the DNA to the SPA beads. In order to prove this it is first necessary to show that an appropriate concentration of $NiCl_2$ is present in the protein buffer. As 0.1mM EDTA is required to quench the non-specific binding it would be expected that at least this concentration of $NiCl_2$ would be present in the assay.

### 5.4.4 Estimation of $NiCl_2$ concentration in protein.

In order to estimate the concentration of the $NiCl_2$ present in the protein a range of solutions containing known concentrations of $NiCl_2$ were prepared containing 25mM HEPES pH 7.9, 10mM DTT and 1mg/ml BSA. A brown discoloration the absorbance of which can be measured at 325nm was present in these solutions and its intensity is related to the $NiCl_2$ concentration (Fig. 5.11).

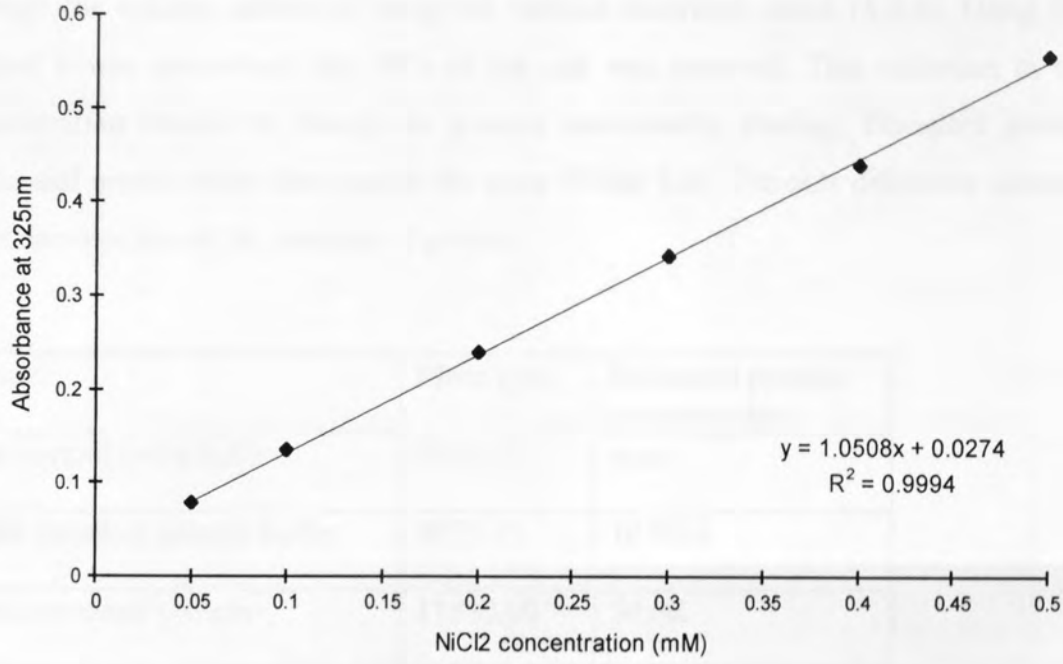



Figure 5.11 Graph showing the relationship between $NiCl_2$ concentration and absorbance at 325nm, used to predict $NiCl_2$ concentrations in the protein.

The absorbance of SPA mixes containing 10mM DTT and the zinc finger protein were also measured and the $NiCl_2$ concentration was predicted from the value given from known concentrations (Fig. 5.11). At the same His-ZF concentration employed in the SPA (54nM) the absorbance at 325nm was 0.413. The concentration of $NiCl_2$ in the assays would have been approximately 0.37mM. The concentration of $NiCl_2$ in the His-ZF protein stock solution was therefore estimated to be18.5mM.

### 5.4.5 Effect of de-salting the protein used in SPA.

Although the EDTA counteracts the effect of the salt in the assay it is also likely to remove the zinc ion from the protein. This will denature the protein preventing specific binding. To investigate whether or not it was possible to detect binding of DNA to His-ZF using this assay, both protein buffer and protein were de-salted by centrifugation through G10 sepharose (Sigma). In order to determine how effective the de-salting was, 500mM $NiCl_2$ was de-salted and the concentration of salt passing through the column calculated using the method described above (5.5.3). Using this control it was determined that 88% of the salt was removed. This reduction in salt concentration should be enough to prevent non-specific binding. De-salted protein buffer and protein were then used in the assay (Table 5.6). The only difference between these samples should be presence of protein.

| Assay. | Mean cpm | Estimated protein concentration |
|---|---|---|
| -ive control (with $H_2O$) | 9042.25 | none |
| With de-salted protein buffer | 9875.25 | 10.8nM |
| With de-salted protein | 11850.00 | 54nM |

Table 5.6 Effect of de-salting protein and protein buffer on the binding of DNA to SPA beads. Protein concentration estimated using known concentrations of His-ZF (4.11.4) and the % removal or protein from the buffer by microconcentration (5.4.2). Assay performed in quadruplet, total cmp/assay = 103912. Full data and statistical analysis shown in appendix XVI.

A students t-test was applied to the data to determine whether or not there was a significant difference between the different treatments. There is no significant difference between the binding of the DNA to the beads in the negative control and the assay containing protein buffer. There is significantly (P-value 0.0322, one tail test) more binding of the DNA to the SPA beads however in the presence of His-ZF than there is in the assay containing buffer. This might indicate that the protein is inducing DNA binding to the SPA beads. These results also suggest that there is nothing else in the buffer causing non-specific binding. A range of protein concentration was then tested using

negative controls containing the amount of $NiCl_2$ calculated to be present in the assays containing protein (Fig. 5.12).

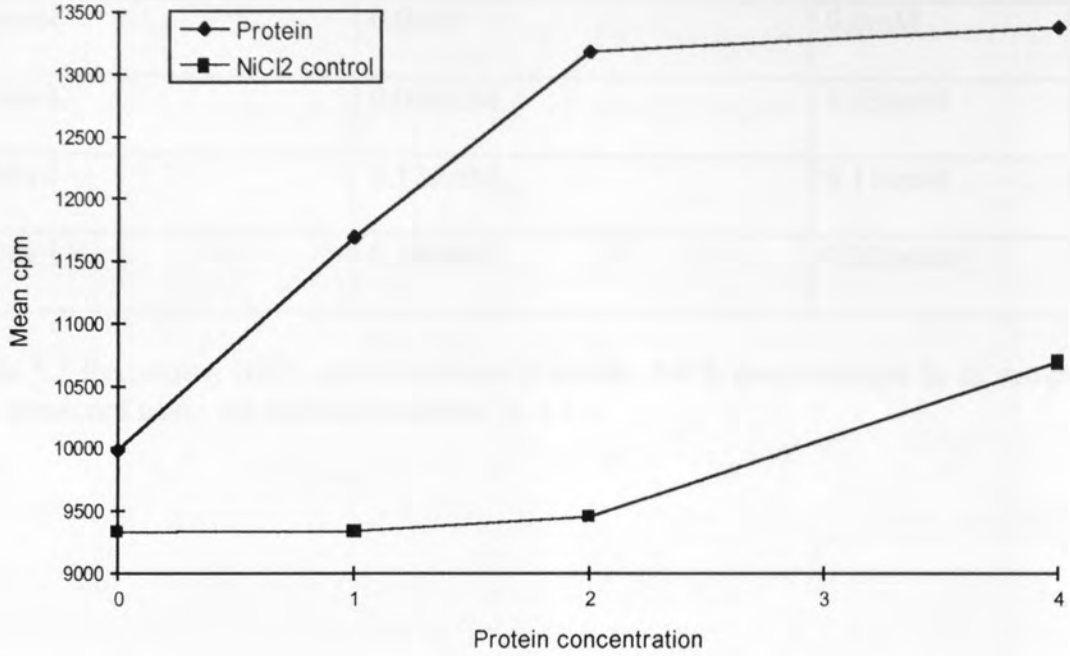



Figure 5.12 Graph demonstrating the effect of increasing concentrations of de-salted protein on the binding of DNA to SPA beads. Effect of increasing concentrations of protein on the binding of DNA to SPA beads. Assays performed in quadruplet, total cpm/assay = 103227. Full data and statistical analysis shown in appendix XVII.

A students t-test was applied to the data to determine the significance of the difference between the assays containing His-ZF and those containing an appropriate quantity of $NiCl_2$. In the absence of the protein there is no significant difference between the assays containing protein and the negative controls. At 1pmol, 2pmol and 4pmol of protein there is significantly (P-values = 0.0244, 0.0051 and 0.0225 respectively) more binding of the DNA to the beads in the presence of His-ZF than there is in its absence.

In order to confirm the validity of this assay the $NiCl_2$ concentration in all samples was measured using the method described in 5.5.3 (Table 5.7).

| | Estimated $NiCl_2$ concentration | |
|---|---|---|
| Concentration of protein (in "protein" assays) | With protein (His-ZF) | Without protein |
| 0pmol | 0.0mM | 0.0mM |
| 1pmol | 0.060mM | 0.055mM |
| 2pmol | 0.120mM | 0.111mM |
| 4pmol | 0.240mM | 0.222mM |

Table 5.7 Remaining $NiCl_2$ concentrations in assays. $NiCl_2$ concentration in all samples was measured using the method described in 5.5.3.

## 5.5 DISCUSSION.

### 5.5.1 SPA method 1 (5.2).

Initially an experiment was carried out which demonstrated that the biotinylated TS DNA was capable of binding to the streptavidin coated PVT SPA beads (Table 5.1). The DNA seemed to be efficiently binding to the beads in its double stranded form and there did not seem to be too much non-specific binding of either unincorporated label or non-biotinylated DNA.

Optimisation of the amount of SPA bead required to bind 1pmol of DNA revealed that 1-1.5mg was sufficient for virtually optimal DNA binding (Fig. 5.3). At high concentrations of SPA beads the counts obtained fluctuated greatly. This is largely due to the difficulties experienced in washing the beads at the higher concentrations. The beads, when in large quantities are easily discarded during measuring.

Assays to detect the binding of the His-ZF protein to the TS DNA, were initially attempted by binding non-labelled TS DNA to the beads and then adding the labelled protein (Table 5.2). The counts obtained from assays containing TS DNA were not raised above those containing no DNA indicating that the protein was not binding to the DNA. It was possible that the large size of the SPA beads was hindering the binding of the protein.

The assay was then attempted by pre-coupling the TS DNA and the zinc finger protein before adding the antibody and SPA beads (Table 5.3). Again the counts obtained in those assays containing TS DNA were not raised above those not containing TS DNA, indicating that the protein was not binding to the DNA.

Since it had already been shown that the protein binds to the DNA by gel retardation assay (4.5.2) it is unlikely that the failure of the assay is due to the His-ZF protein not binding to the TS DNA. It is possible that binding is not detected because the protein has too low a specific activity. This is a major problem as it would be difficult to synthesise suitable quantities of protein with a much higher specific activity. Secondly steric hindrance caused by the size of the bead could prevent the protein binding the DNA. Again this would be difficult to resolve, a possible solution would be to move the target binding DNA site further away from the bead. This would either require testing DNA of different lengths, or due to the possibility of the protein binding

to regions of the DNA sequence that it should not bind to, synthetic linkers may have to be considered. For these reasons it was impractical to pursue the assay using this format.

**5.5.2 SPA method 2 (5.3).**

Failure of the assay to detect the interaction between the protein and the DNA using method 1 could be due to the low specific activity of the protein or hindrance from the SPA bead. Coating the protein onto the beads increases the distance between the bead and the interacting molecules (DNA and zinc finger protein) and labelling of DNA to a high specific activity is easily achieved. It should therefore be possible to overcome the problems associated with SPA method 1 (Fig 5.1) by performing the assay using method 2 (Fig. 5.2).

The only assay format in which His-ZF seemed to cause the DNA to bind to the SPA beads was the one that relied on direct binding of His-ZF to protein A coated yttrium silicate beads (Table 5.4). All other assay formats were unsuccessful. The failure of these formats probably results from the failure of the antibodies available to bind to His-ZF. The zinc finger protein was thought to bind to the SPA beads though co-ordination of the yttrium in the beads by the His tag. If this was the case the binding of the protein to the beads should be both strong and specific. Also it would not interfere with the DNA binding domain. It therefore seemed a suitable method for linking the protein to the beads in order to detect zinc finger / DNA interactions.

Using this assay format it was demonstrated that there was a strong relationship (Fig. 5.4) between increasing His-ZF concentration and increased binding of the DNA to the SPA beads. The assay was then optimised with regards to pH, KCl concentration and $ZnCl_2$ concentration.

At pH 6.8, 7.6. 7.8 and 8.0 there is significantly more binding of DNA to the SPA beads in assays containing His-ZF than there is in its absence. At pH 7.0, 7.2 and 7.4 there was no significant difference between the binding of DNA to SPA beads in the positive and the negative controls (Fig. 5.5).

It was concerning that over a small pH range (7.0-7.4) the DNA bound non-specifically to the SPA beads, as addition of the protein to the assay may simply alter the pH causing non-specific binding. In order to ensure that the pH in the assay remained stable on addition of His-ZF, assay mixes at pH 7.9 both with and without His-ZF were applied to pH paper. This confirmed that an alteration in the pH was not responsible for

causing the DNA to bind to the SPA beads. The reason why DNA binds non-specifically to the beads at pH 7.0-7.4 is unknown and all subsequent assays were performed at pH 7.9 to avoid non-specific binding.

In the absence of salt and at 2.5mM KCl the binding of DNA to the SPA beads was significantly higher in the assays containing His-ZF than in the negative controls. At higher concentrations of KCl (25 and 250mM) there is a large increase in the non-specific binding of the DNA to the SPA beads (Fig. 5.6). These results indicate that KCl at higher concentrations causes the DNA to non-specifically bind to the SPA beads. The binding of the DNA to the beads does not increase (in response to KCl) as much in the assays containing His-ZF as it does in the negative controls. This is possibly due to the presence of the protein "blocking" the beads making less bead available for non-specific binding. As there was no particular advantage to the addition of KCl to the assay it was omitted in subsequent assays.

In the absence of $ZnCl_2$ the binding of the DNA to the SPA beads was significantly higher in the presence of His-ZF than it was in its absence. At 2.5mM $ZnCl_2$ there is no significant difference between the binding of DNA to SPA beads in the assays containing His-ZF and the negative controls. At higher concentrations of $ZnCl_2$ (25 and 250mM) the binding of the DNA to the SPA beads is significantly higher in the assays containing protein than in the negative controls. The results show that even at low concentrations of $ZnCl_2$ there is a massive increase in the non-specific binding of DNA to SPA beads (Fig. 5.7). At the higher concentrations (25 and 250mM) the binding of DNA to the SPA beads is greater in the samples containing His-ZF.

The effect of relatively low concentrations of $ZnCl_2$ on the non-specific binding of DNA to SPA beads is particularly concerning. The buffer that the protein is in is likely to contain small quantities of various reagents as a result of the method of protein purification. Binding of the DNA to the beads could be attributed to non-specific binding induced by contaminants rather than specific binding of the DNA to His-ZF.

**5.5.3 Investigation into the mode of binding of DNA to SPA beads (5.4).**

Removal of the protein from the buffer and subsequent use of the buffer in the assay demonstrated that there was no significant difference between the effect of His-ZF and its buffer alone on the binding of DNA to SPA beads (Table 5.5). This strongly suggested that something other than the His-ZF protein was causing the DNA to bind non-specifically to the SPA beads.

Due to the preparation and subsequent dialysis of the protein only KCl, imidazole and $NiCl_2$ are possible contaminants in the assay at concentrations which may effect non-specific binding. If one of these reagents was capable, at low concentration, of mimicking the effect of adding His-ZF to the assay, it would be a good candidate for the "rogue" substance.

The effect of KCl on the assay has been previously determined (5.3), and at low concentrations the effect is not significant.

The effect of imidazole (Fig. 5.8) was determined by the addition of 0.01, 0.1, 1.0 and 10mM imidazole to the assay. In all three samples imidazole causes a decrease in the binding of the DNA to SPA beads implying that it is not responsible for the non-specific binding of the DNA to the SPA beads. In retrospect this was because the imidazole "mops up" the $NiCl_2$ which was causing the DNA to non-specifically bind to the SPA beads.

The effect on $NiCl_2$ on the binding of the DNA to the SPA beads was assessed by the addition of of 0.01, 0.1, 1.0 and 10mM of the salt to the assay (5.9). As a divalant cation ($ZnCl_2$, 5.2.2) had already been shown to cause the non-specific binding of DNA to the SPA beads therefore $NiCl_2$ seemed the most likely candidate for the "rogue" substance in the protein buffer. Small quantities of $NiCl_2$ were capable of causing the DNA to non-specifically bind to the SPA beads. The results implied that providing the concentration of $NiCl_2$ in assays containing the protein or protein buffer is above approximately 0.1mM the salt could be causing the non-specific binding.

In order to provide further evidence that the non-specific binding is caused by nickel cations in the buffer, a range of concentrations of EDTA (Fig. 5.10) was added to the assays. As EDTA "mops up" ions in the buffer it should inhibit the non-specific binding of the DNA to the SPA beads. EDTA (at 1.0 and 10mM) as expected causes a huge decrease in the binding of DNA to the SPA beads. This supports the theory that

that $NiCl_2$ was causing the non-specific binding. The results also indicate that the concentration of $NiCl_2$ in the assay must be greater than 0.1mM.

In order to confirm that $NiCl_2$ was indeed responsible for the non-specific binding of the DNA to the SPA beads it was necessary to infer how much was present in the assay. By comparison of the assay mix with similar mixes containing known concentrations of $NiCl_2$ (Fig. 5.11) it was possible to predict that in the assay containing protein there was 0.37mM $NiCl_2$. This is above the concentration predicted by previous experimentation (5.4.2) to be capable of causing the DNA to non-specifically bind to the SPA beads. Due to the dialysis of the protein after purification it is unlikely that there should be this much $NiCl_2$ present in the buffer. A possible reason why it has been retained is that the His tag has "held" onto it.

Finally the effect of removing the $NiCl_2$ from the buffer was investigated, in order to determine whether or not specific binding of the DNA to the protein could be detected in its absence. It was possible to remove 88% of the salt form the protein and the protein buffer (5.4.5). De-salted samples were then used in the assays. No significant difference was found between the amount of DNA bound to the beads in the assays containing protein buffer and the negative control. The DNA bound to the beads in assays containing His-ZF was a significantly greater than that bound in the assay containing protein buffer (Table 5.6).

A range of protein concentrations were then tested and the negative controls used which reflected the amount of $NiCl_2$ present in the "positives" (Fig. 5.12 and Table 5.7). A strong correlation was observed between the amount of protein in the assay and the amount of DNA bound to the beads. There was no significant difference between the binding of the DNA to the beads in the positive and negative samples in the absence of the protein. In the presence of His-ZF (at all concentrations) there was significantly more binding of the DNA to the SPA beads than there was in the negative controls.

It does seem that the assay is detecting a specific interaction between the DNA and the protein. It is problematic however that the assay relies on non-specific binding of the protein to the yttrium silicate beads. The non-specific binding of DNA to the SPA beads in the presence of small quantities of salt is also unacceptable. It would therefore be more practical to use PVT beads for this assay, unfortunately with no antibody that recognises His-ZF there is no method by which the protein can be attached to PVT SPA beads. In order to develop an assay which is reliable it is necessary either to continue

with SPA method 1 (this is impractical for the reasons outlined above, 5.5.1), or to change the protein such that it can be attached to PVT SPA beads. A frequently used expression/purification system is glutathione S-transferase (GST) fusion. Fusion of the zinc finger gene to GST would provide a convenient method by which the protein could be attached to protein A coated PVT SPA beads (via an anti-GST antibody).

**5.5.4 Summary.**

- Biotinylated double stranded target sequence DNA has been shown to successfully interact with streptavidin coated PVT SPA beads (5.2.1).

- SPA method 1 (Fig. 5.1) was not successful, an interaction between the DNA and the protein was not detected. This is possibly because the protein was not of high enough specific activity. But steric hindrance of the SPA bead due to the short length of the DNA may also be important (5.2.2).

- SPA method 2 (Fig 5.2) initially seemed successful (5.3.1) but the binding of the DNA to the SPA beads was later shown to be caused by $NiCl_2$ in the buffer rather than His-ZF.

- On removal of the $NiCl_2$ from the protein it seemed possible to detect the interactions between the DNA and His-ZF. But due to the non-specific binding caused by small quantities of $NiCl_2$ and the unknown nature of the interaction between His-ZF and the beads the assay format was considered unreliable (5.3.5).

- Neither of the assay formats tested here are practical, the zinc finger domain of His-ZF will therefore be fused to GST in order to allow it to be attached to protein A coated PVT beads, via an anti GST antibody.

# 6. FUSION OF THE ZINC FINGER GENE TO A GST GENE.

## 6.1 INTRODUCTION.

Since the His-ZF protein could not be linked to PVT SPA beads, the assay format shown in figure. 5.2 was ineffective for detecting interactions between the zinc finger protein and DNA. To develop a reliable assay format, the zinc finger domain of His-ZF might be fused to a Glutathione S-transferase (GST gene). Anti-GST antibodies that might be used to attach the zinc finger protein to protein A-coated PVT beads are readily available.

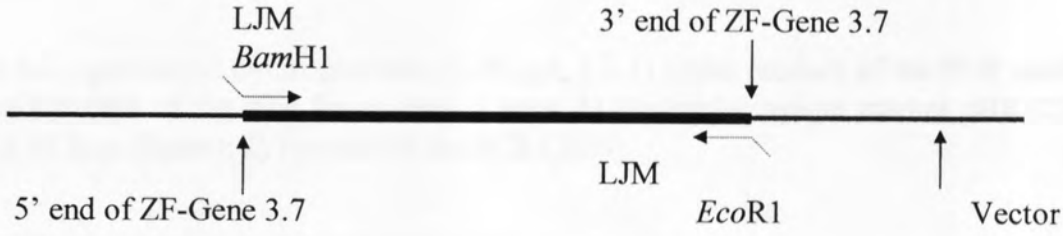
Expression and purification systems for GST-fused proteins are also readily available. A GST fusion vector (pGEX-2TK, Amersham Pharmacia Biotech) was selected. This particular vector also incorporates a kinase site, allowing the protein to be radiolabelled during purification, and a thrombin cleavage site allowing removal of the GST subunit from the protein of interest. The pGEX-2TK vector contains a lac promoter which controls the expression of the GST fused genes. This allows expression of the gene to be induced with IPTG.

As they were no longer required the His tag and the ATG "start" codon were removed from the 5' end of the gene. To allow the zinc finger gene to be sub-cloned into the pGEX-2TK vector, the ends of the gene were altered in order to incorporate appropriate restriction sites (5' *Bam* H1 and 3' *Eco* R1).

## 6.2 PRODUCTION OF THE ZINC FINGER GST-FUSED GENE.

### 6.2.1 Alteration of the zinc finger gene.

For the cloning of the zinc finger gene into pGEX-2TK such that it is in frame with the GST gene, a *Bam*H1 site was required at the 5' end of the gene. An *Eco*R1 site at the 3' end of the gene would allow the fragment to be cloned directionally. These sites were incorporated into two oligonucleotide primers, which were otherwise complementary to the zinc finger gene and suitable for amplification of the gene by PCR (Fig. 6.1).




LJM *Bam*H1

5' TT**G GA TCC** ACC GGC GAA AAA CCG TAC 3'

LJM *Eco*R1

5' AA **GAA TTC** TTA ATG CGT GCG CTG ATG 3'

Figure 6.1 Method and primers used for alteration of the ends of the zinc finger gene. Restriction sites to be added to the end of the gene are shown in bold within the primers.

The gene was amplified by PCR (2.8.1) using primers LJM *Bam*H1 and LJM *Eco*R1 in a 50µl reaction. ZF-Gene 3.7 (the zinc finger gene cloned into pT7-7), pre-restricted with *Sau*3A (2.9) was used as a template. Pre-restriction of the template

prevents it being transformed into the host bacteria at a later stage, as this would cause the original pT7-7 construct to be isolated in preference to the altered construct. The product was analysed by gel electrophoresis (Fig. 6.2). The remaining 40μl of the product was placed in a blunt end reaction (2.9.2) and then ligated (2.9.3) into *Sma* 1 cut phosphatased pUC 19 (appendix III). The resulting constructs were transformed (2.7.1) into competent DH5α cells (Table 6.1).

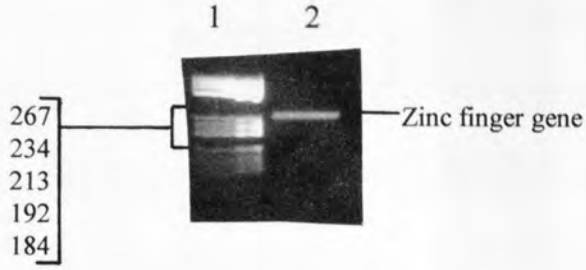



Figure 6.2 Agarose gel electrophoresis (2 2% gel, 2.5.1) of the product of the PCR used to alter the ends of the zinc finger gene. Lanes; 1) Molecular weight marker pBR322 *Hae* III (0.3μg; Sigma); 2) Product of the PCR (20%).

| Experiment | White colonies | Blue colonies |
|---|---|---|
| Fragment ligation. | 48 | 19 |
| Self ligation | 3 | 28 |

Table 6.1 Results of transformation of altered zinc finger gene into competent DH5α cells.

Possible recombinant colonies were screened by PCR (2.9.1) using pUC19 forward and reverse primers (appendix V). The products of the PCR were examined by gel electrophoresis (2.5.1). Twelve colonies containing the gene were identified. The plasmid DNA from three of these clones was prepared using a Wizard Mini-Prep kit (Promega, 2.8). The DNA was then sequenced using a Thermo-sequenase $^{33}P$ terminator kit (Amersham Pharmacia Biotech, 2.9). None of the clones sequenced were complete: all had sequences missing from one or both ends. One clone (ZFG, Fig. 6.3)

was only missing part of the 3' *Eco* R1 site, but was orientated into pUC-19 such that there was an available *Bam* H1 site at either end of the gene (Fig. 6.4).

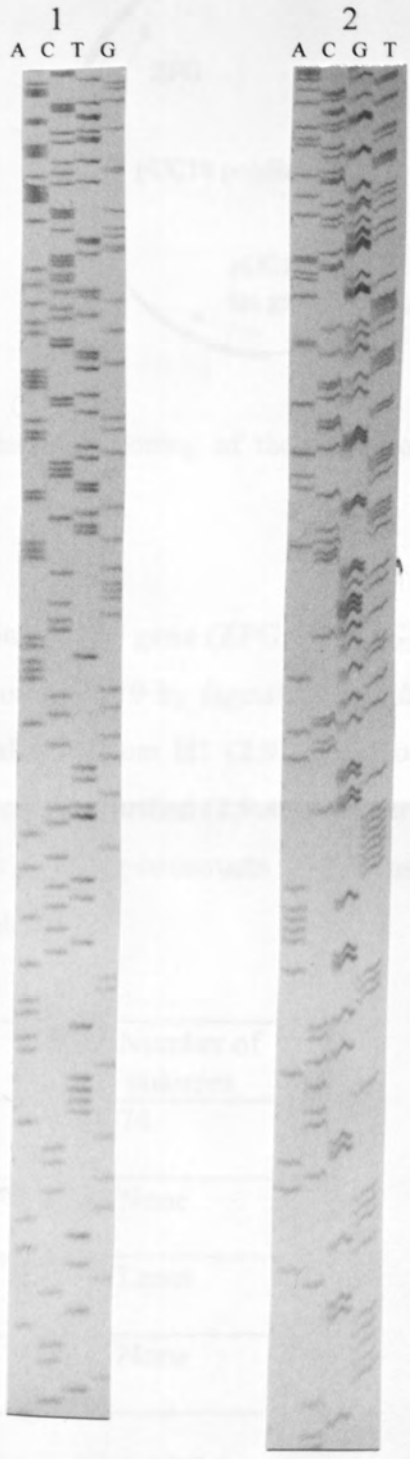



Figure 6.3 Sequences of pUC19 clones containing the zinc finger gene (ZFG) with gene termini altered by site directed mutagenesis. 1) Sequence obtained pUC19 forward primer 2) Sequence pUC19 reverse primer.

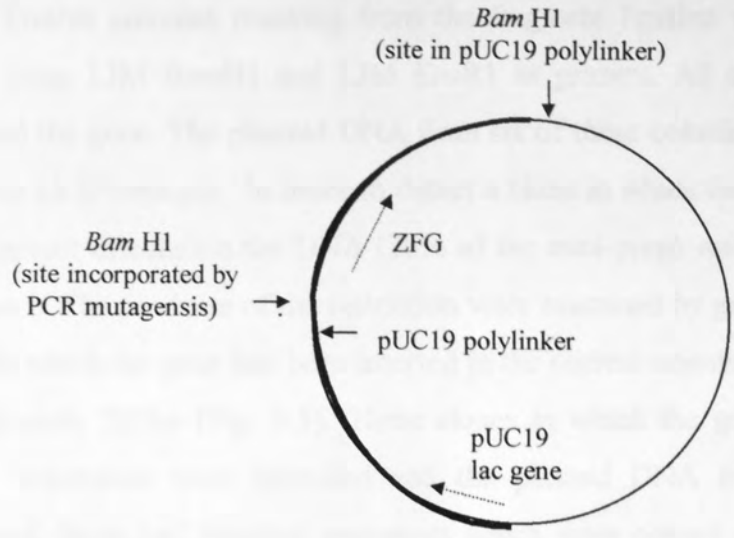



Figure 6.4 Construct obtained by cloning of the PCR constructed altered zinc finger gene.

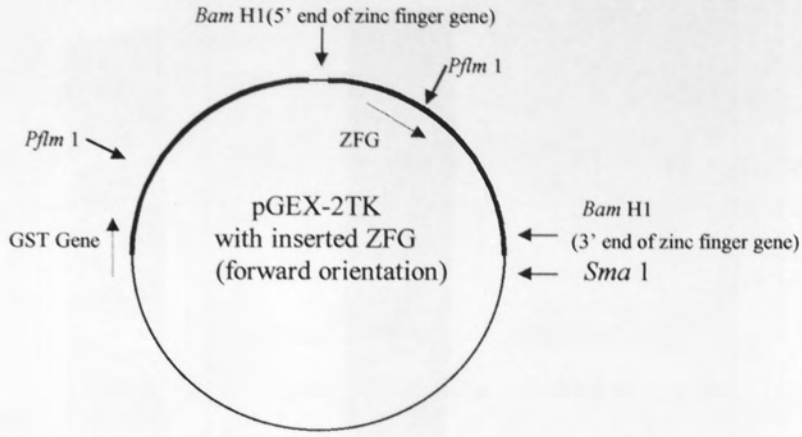
### 6.2.2 Sub-cloning of the zinc finger gene (ZFG) into pGEX-2TK.

ZFG was excised from pUC19 by digestion with *Bam* H1. The vector (pGEX-2TK) was similarly digested with *Bam* H1 (2.9) and phosphatased using CIAP (2.9). Both the vector and gene were gel purified (2.9.4). The purified gene was then ligated in to the vector (2.9.3) and the resulting constructs were transformed into MAX Efficiency DH5α competent cells (Table 6.2).

| Experiment | Number of colonies |
|---|---|
| Fragment ligation | 74 |
| Self ligation | None |
| Control transformation | Lawn |
| Antibiotic resistance. | None |

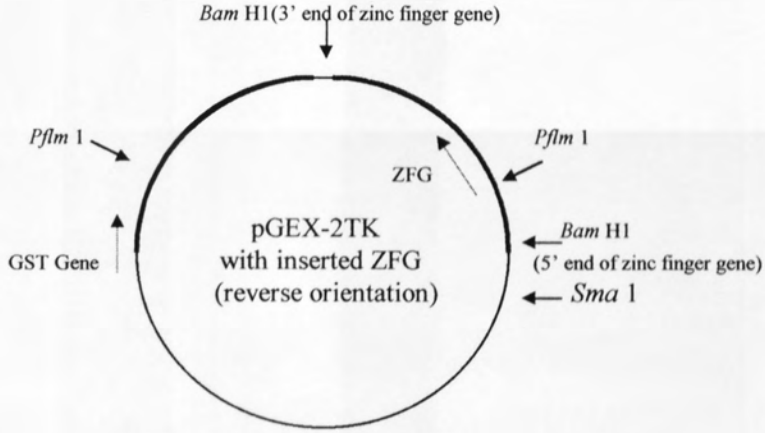
Table 6.2 Results of transformation of ZFG (sub-cloned into pGEX-2TK) into MAX Efficiency DH5α competent cells (Gibco BRL).

Twelve colonies resulting from the fragment ligation were screened by PCR (2.9.1) using LJM *Bam*H1 and LJM *Eco*R1 as primers. All of the colonies selected contained the gene. The plasmid DNA from six of these colonies was prepared using a mini-prep kit (Promega). In order to detect a clone in which the insert had been cloned in the correct orientation the DNA (20% of the mini-prep) was restricted with *Pflm* 1 and *Sma* 1. The products of the restriction were examined by gel electrophoresis Those clones in which the gene had been inserted in the correct orientation yield a fragment of approximately 220bp (Fig. 6.5). Three clones in which the gene was inserted in the correct orientation were identified and the plasmid DNA from two of these was sequenced. Both had identical sequences which were correct and confirmed that the gene had been sub-cloned in the correct orientation (Fig. 6.6).




On restriction with *Pflm* 1 and *Sma* 1 sizes of fragments obtained approximately;

2288bp
2756bp
200bp




On restriction with *Pflm* 1 and *Sma* 1 sizes of fragments obtained approximately;

2288bp
2881bp
100bp

Figure 6.5 Diagrams illustrating pGEX-2TK containing ZFG both in the correct orientation and incorrect orientation.

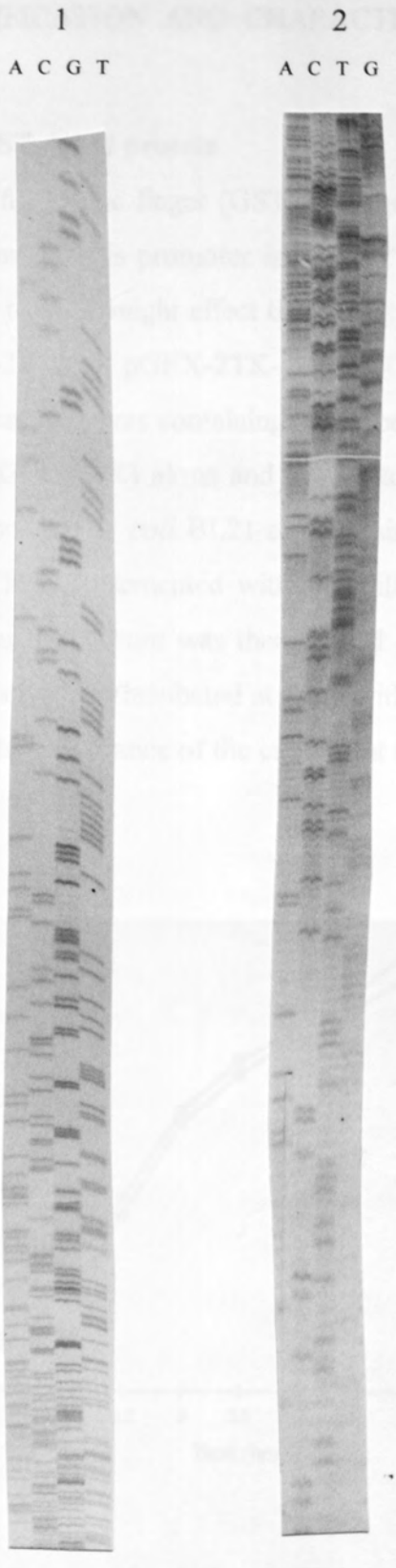



Figure 6.6 Sequence of GST zinc finger fusion gene (GST-ZFG). 1) Sequence obtained from primer LJM *Bam*H1. 2) sequence obtained from primer LJM *Eco*R1.

## 6.3 EXPRESSION, PURIFICATION AND CHARACTERISATION OF THE GST FUSION PROTEIN.

### 6.3.1 Expression of the GST-fused protein.

GST and the GST-fused zinc finger (GST-ZFG) genes are under the control of the lac promoter. The control of this promoter is slightly "leaky". Since the zinc finger gene product is likely to be toxic, it might effect the viability of the cells. The viability of cultures containing pGEX-2TK and pGEX-2TK-GST-ZFG was assessed by measuring the growth of cultures of transformants containing the respective genes.

Both the vector (pGEX-2TK) alone and the vector containing the fused gene were transformed into competent *E. coli* BL21 cells. A single transformed colony was used to inoculate 30ml of LB supplemented with ampicillin and the culture incubated overnight at 37°C. The resulting culture was then diluted 1/1000 into fresh LB (30ml) containing ampicillin and the culture incubated at 37°C with shaking. The growth of the cultures was measured by the absorbance of the cultures at 600nm (Fig. 6.7).

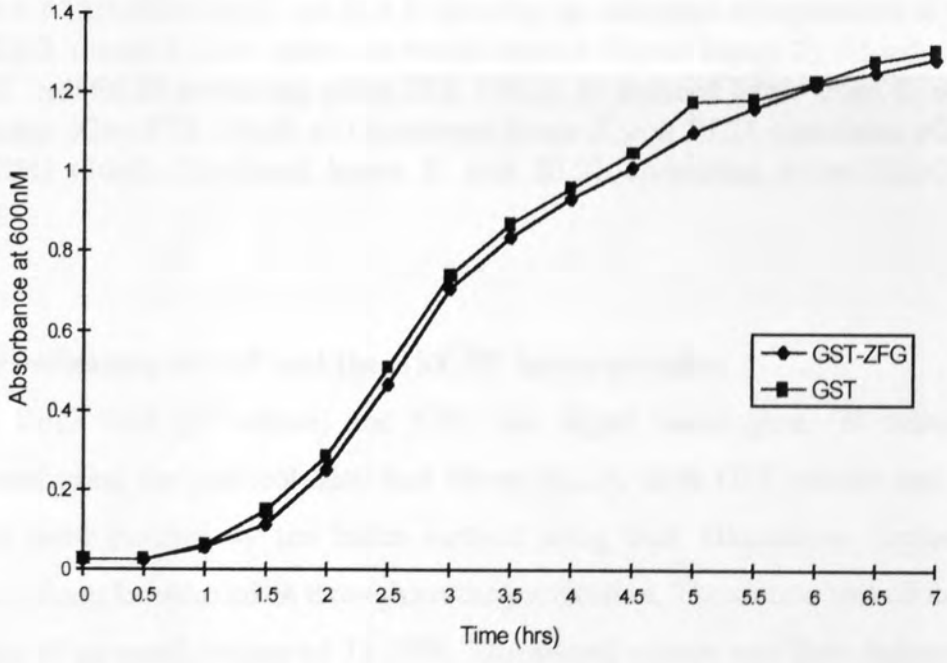



Figure 6.7 Graph showing the growth of *E. coli* BL21 transformed with pGEX-2TK alone and containing the GST-fused gene.

Comparison of the growth of the cultures indicates that viability is not adversely affected by the presence of the zinc finger gene.

An overnight culture was prepared from fresh transformants of both pGEX-2TK alone and containing the fused gene (as described above). Proteins were expressed in 2l cultures (2.11.1). Samples of the culture were taken before induction and 3hrs after induction. These samples were examined by SDS page gel electrophoresis (Fig. 6.8).

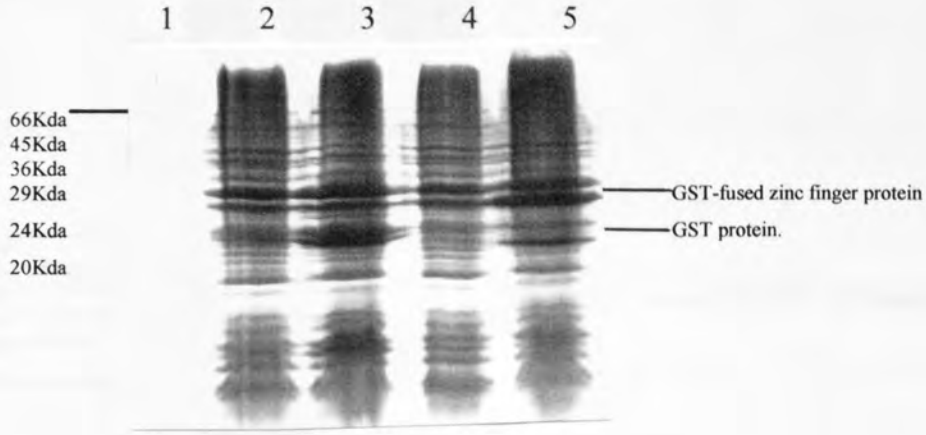



Figure 6.8 15% SDS-PAGE gel (2.5.3) showing the induction of expression of GST and GST-ZFG. Lanes: 1) Low molecular weight marker (Sigma 10μg); 2) Un-induced lysate from *E. coli* BL21 containing pGex-2TK (30μl); 3) Induced lysate from *E. coli* BL21 containing pGex-2TK (30μl); 4) Un-induced lysate *E. coli* BL21 containing pGex-2TK-GST-ZFG (30μl); 5)Induced lysate *E. coli* BL21 containing pGex-2TK-GST-ZFG (30μl);

### 6.3.2 Purification of GST and the GST-ZF fusion proteins.

Both GST (2l culture) and GST zinc finger fusion gene (6l culture) were expressed using the protocol described above (6.3.1). Both GST protein and GST-ZF protein were purified by the batch method using bulk Glutathione Sepharose 4B (2.11.3). Samples were taken throughout the purification. The elutate was diluted by the addition of an equal volume of 1x PBS. The diluted elutate was then dialysed (2.5.3) against the buffers described in 2.11.3. Samples were examined by SDS PAGE gel electrophoresis (Fig. 6.9)

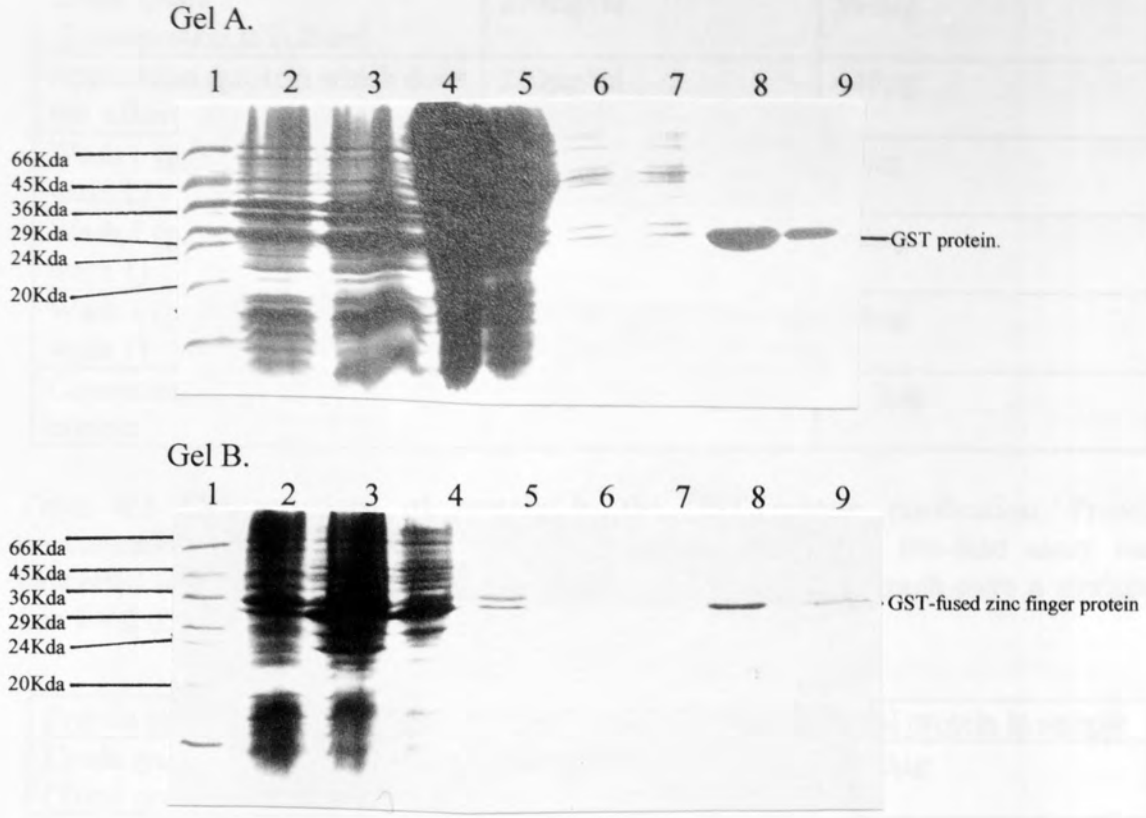



Figure 6.9 SDS-PAGE (15%) gels showing the purification of GST and GST-ZF. Gel A Purification of GST. Gel B Purification of GST-ZF. Samples on both gels are loaded in the same order. Lanes: 1) Low molecular weight marker (Sigma 10μg); 2) Un-induced lysate; 3) Induced lysate; 4) Wash 1(30μl); 5) Wash 2 (30μl); 6) Wash 3 (30μl); 7) Wash 4 (30μl); 8) Elutate (30μl). 9) Dialysed product (10μl), 3.9μg GST protein and 415μg GST-fused zinc finger gene.

Protein concentrations for both the GST protein (Table 6.3) and the GST-ZF protein (Table 6.4) were estimated using the Bio-Rad protein assay. Calibration was carried out using known quantities of BSA.

| Protein sample (GST) | Protein concentration | Total protein in sample |
|---|---|---|
| Crude lysate (Total protein in culture) | 270μg/ml | 540μg |
| Application (protein which does not adhere to the sepharose | 220μg/ml | 440μg |
| Wash 1 (protein removed in wash 1) | 50μg/ml | 50μg |
| Wash 1 (protein removed in wash 1) | 9.2μg/ml | 9.2μg |
| Wash 1 (protein removed in wash 1) | 2.8μg/ml | 2.8μg |
| Concentration of dialysed protein | 390μg/ml | 11.7μg |

Table 6.3 Concentrations of protein in the GST protein purification. Protein concentrations were estimated by Bio-Rad protein assay. The Bio-Rad assay was calibrated using known concentrations of BSA. The calibration graph gave a straight-line $y = 0.0324 x + 0.02$ and $R^2 = 0.9854$.

| Protein sample (GST-ZF) | Protein concentration | Total protein in sample |
|---|---|---|
| Crude lysate (Total protein in culture) | 271μg/ml | 1900μg |
| Application (protein which does not adhere to the sepharose | 223μg/ml | 1560μg |
| Wash 1 (protein removed in wash 1) | 71μg/ml | 213μg |
| Wash 1 (protein removed in wash 1) | 7.6μg/ml | 23μg |
| Wash 1 (protein removed in wash 1) | 2.0μg/ml | 6.1μg |
| Concentration of dialysed protein | 41.5μg/ml | 31.1μg |

Table 6.4 Concentrations of protein in the GST-ZF protein purification. Protein concentrations were estimated by Bio-Rad protein assay. The Bio-Rad assay was calibrated using known concentrations of BSA. The calibration graph gave a straight-line $y = 0.0324 x + 0.02$ and $R^2 = 0.9854$.

### 6.3.3 Characterisation of GST-ZF.

It was necessary to confirm that the zinc finger GST fusion product was capable of binding to both its target DNA sequence and to the Anti-GST antibody.

Initially the activity of the GST (both in the GST and GST fusion protein) was demonstrated using the CDNB assay (Amersham Pharmacia Biotech). Both proteins were shown to have activity (Fig. 6.11).

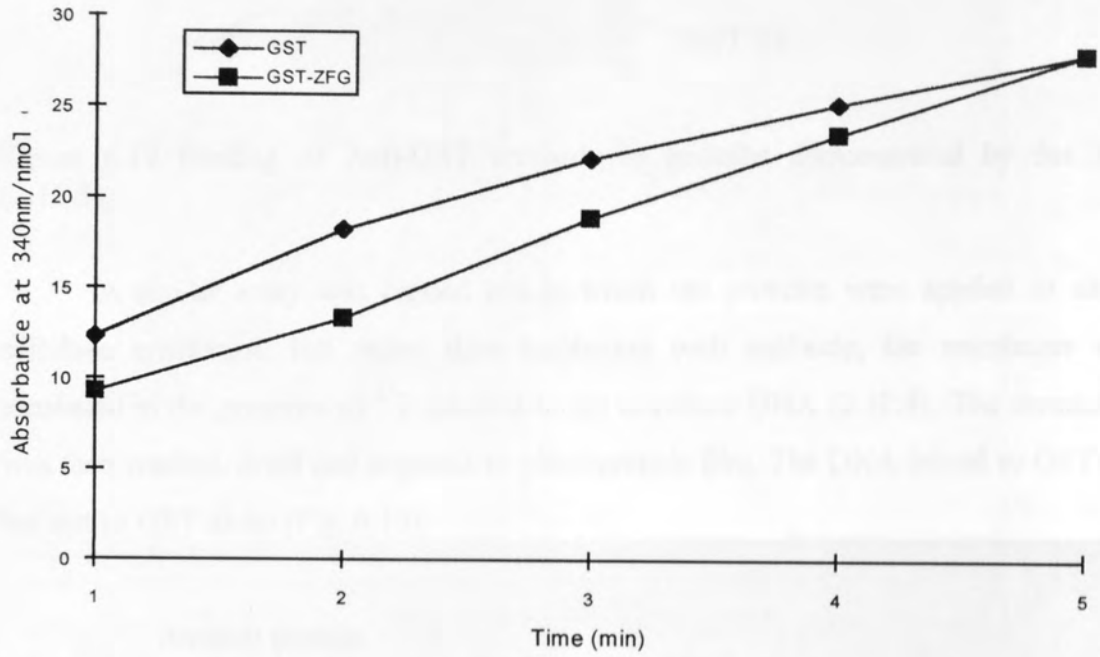



Figure 6.11 Graph showing activity of the GST using the CDNB assay.

Binding of the Anti-GST (Amersham Pharmacia Biotech) antibody to the protein was demonstrated by dot blot (2.11.5). Both GST and GST-ZF were applied to nitro-cellulose membrane. As a control, a small quantity of Anti-GST and the conjugate was also applied. The results demonstrated that the antibody bound to both proteins (Fig. 6.12).

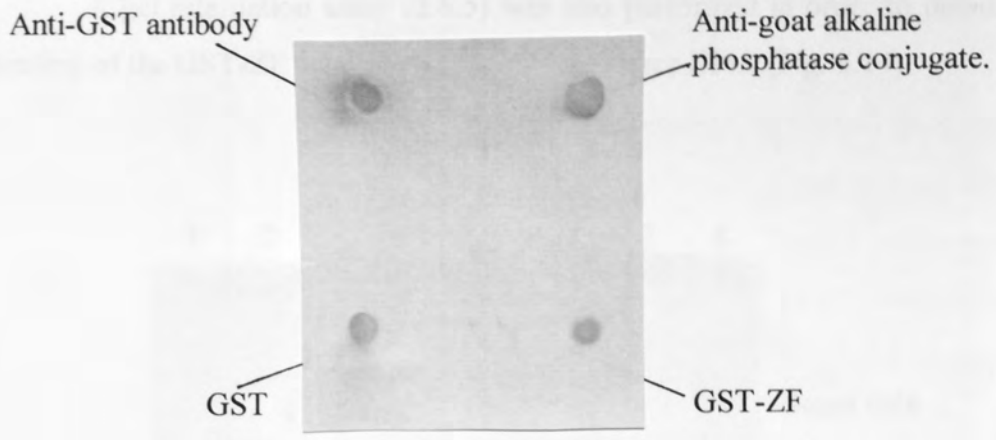



Figure 6.12 Binding of Anti-GST antibody to proteins demonstrated by dot blot (2.11.5).

A similar assay was carried out in which the proteins were applied to nitro-cellulose membrane, but rather than incubation with antibody, the membrane was incubated in the presence of $^{33}P$ labelled target sequence DNA (2.12.5). The membrane was then washed, dried and exposed to photographic film. The DNA bound to GST-ZF but not to GST alone (Fig. 6.13).

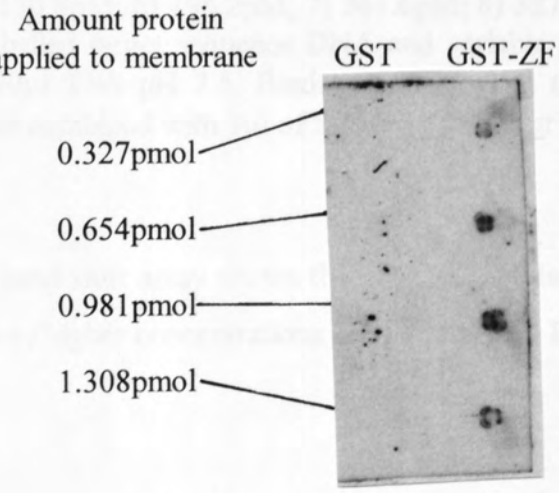



Figure 6.13 Dot blot demonstrating the binding of the DNA to GST-ZF (2.11.6).

A gel retardation assay (2.6.5) was also performed in order to demonstrate the binding of the GST-ZF protein to the target sequence DNA (Fig. 6.14).

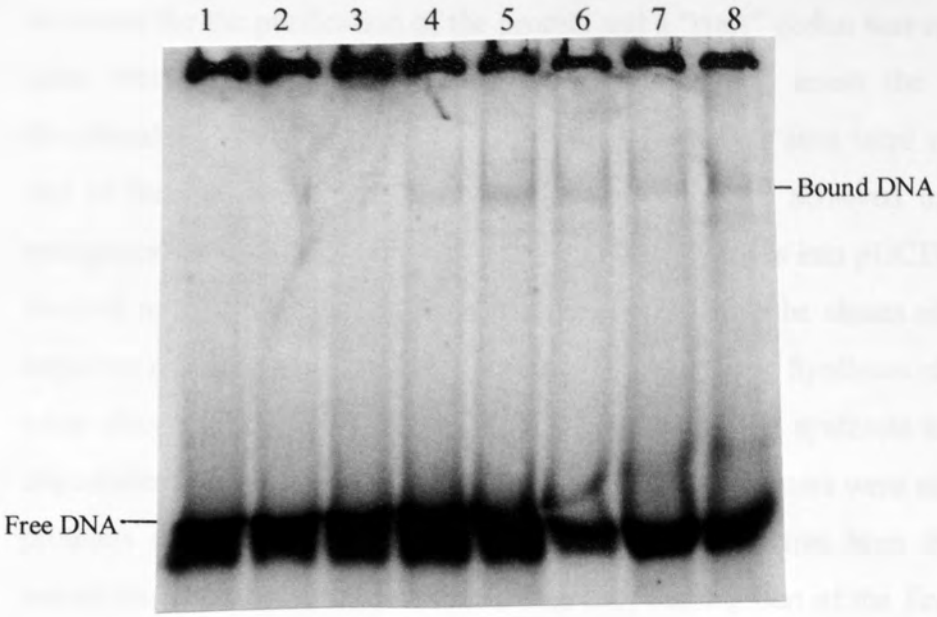



Figure 6.14 Band shift assay demonstrating the binding of GST-ZF to the target sequence DNA. Protein concentrations in assays 1) 0.0pM; 2) 3.27pM; 3) 32.7pM; 4) 65.4pM; 5) 130.8pM; 6) 196.2pM; 7) 261.6pM; 8) 327pM. Binding reactions contained 10nM $^{33}$P-labelled target sequence DNA and variable protein concentrations in a total volume of 20µl TBS pH 7.5. Binding was allowed to take place at 4°C for 30min. Samples were combined with 5µl of 50% glycerol prior to loading on a 9% gel (2.5.5).

The band shift assay shows that GST-ZF binds to the target sequence DNA, in the presence of higher concentrations of protein, more DNA is bound.

## 6.4 DISCUSSION.

### 6.4.1 Fusion of the ZF gene to a GST gene.

In order to clone the zinc finger gene into the vector pGEX-2TK it was first necessary to remove the His tag and ATG "start codon". The His tag was no longer necessary for the purification of the protein and a "start" codon was not required as the gene, would be fused to the GST gene. In order to insert the zinc finger gene directionally behind the GST gene, additional restriction sites were required on either end of the ZF gene. The alterations to the gene were achieved using site directed mutagenesis, followed by blunt end cloning of the products into pUC19 (6.2.1). The site directed mutagenesis was not entirely successful and all the clones obtained had some sequence missing from either the 5' or 3' end of the gene. Synthesis of oligonucleotides takes place in the 3'-5' direction, as result of incomplete synthesis the 5' ends of the oligonucleotides may be deleted. As the oligonucletide primers were not gel purified the products of incomplete oligonucleotide synthesis may have been the source of the mutations. One of the clones however was only missing part of the *Eco*R1 site from the 3' end of the gene (after the stop codon). Due to the orientation of this insert in pUC19 it was possible to excise the gene using *Bam*H1.

This made the task of sub-cloning slightly more difficult as a clone in which the gene had been inserted in the correct orientation must be selected. This did not, however prove to be problematic and several clones containing the zinc finger gene in the correct orientation were identified, two of which were sequenced, confirming that the sequence was correct (6.2.2).

### 6.4.1 Expression, Purification and Characterisation of GST-ZF.

The product of the GST fused zinc finger gene was successfully expressed (6.3.1) and purified (6.3.1)

The protein product of the gene (GST-ZF) was then characterised. Firstly the activity of the GST was tested using the CDNB assay (Amersham Pharmacia Biotech). This demonstrated that the GST portion of the protein was functional therefore folded correctly (Fig. 6.10). A dot blot was then used to confirm that the anti-GST antibody bound to the purified GST and GST-ZF. The antibody successfully recognised both proteins (6.11).

In order to confirm that the protein bound to its target DNA sequence a dot blot using labelled TS DNA as a probe was carried out. This demonstrated that the DNA bound GST-ZF and that it did not bind to GST (Fig. 6.12). Finally a band shift assay was carried out, this demonstrated that the protein bound to the DNA and that in the presence of increasing concentrations of protein more DNA was bound (Fig. 6.13).

GST-ZF is capable of binding both to the anti GST antibody and to its target DNA sequence. It should therefore be possible to detect the interaction between the zinc finger protein and its target DNA sequence using protein A coated PVT SPA beads.

**6.4.3 Summary.**

- The zinc finger gene was altered in order to incorporate restriction sites which allowed it to be sub-cloned in to the vector pGex-2TK, as a GST fusion gene.
- The protein product of the fusion gene was successfully expressed and purified.
- GST-ZF was shown to bind to both the anti-GST antibody and the TS DNA.

# 7 DEVELOPMENT OF AN ASSAY FOR THE DETECTION OF DNA / GST-ZF INTERACTIONS.

## 7.1 INTRODUCTION.

An assay was required in order to detect the binding of DNA to the zinc finger protein. A scintillation proximity assay format in which the DNA had been linked to the beads and the binding of labelled protein monitored, failed to detect the interaction between the DNA and the protein (5.2). The second assay format in which the zinc finger protein was bound to yttrium silicate SPA beads was impractical to use because of the problems associated with non-specific binding to this bead type (5.3). In view of these results, a method by which the zinc finger protein could be attached to PVT SPA beads was required. PVT SPA beads are available coated in protein A. A variety of antibodies can therefore be bound by these beads. The zinc finger protein had been fused to a GST domain. Antibodies that recognise GST are readily available. Such antibodies may therefore be used to attach the zinc finger protein specifically to protein A-coated PVT SPA beads.

The assay utilises SPA technology (Amersham Pharmacia Biotech) in order to detect the binding of labelled DNA to the zinc finger protein, which is fused to GST (Fig. 7.1). The zinc finger protein (GST-ZF) is attached to protein A coated SPA beads using an anti-GST antibody. In the event of binding of the DNA to the protein, the label is brought into close enough proximity of bead for the scintillant to emit a signal. Binding of the DNA to GST-ZF can therefore be detected using a scintillation counter.

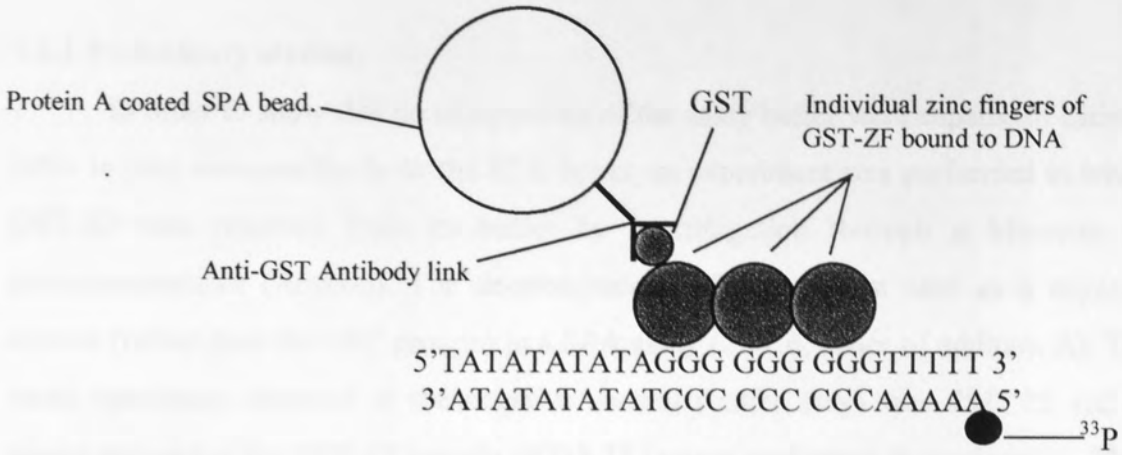



Figure 7.1 Assay for the detection of DNA protein interactions using GST-ZF.

The gene encoding the zinc finger protein had been successfully cloned into a GST fusion vector (pGEX-2TK, 6.2) and the GST fused product (GST-ZF) had been expressed (6.3.1) and purified (6.3.2). The constitutive interactions for the SPA format illustrated in figure 7.1 have been show to function in isolation. It has been demonstrated by gel retardation assay (Fig. 6.14) and dot blot (Fig. 6.13) that GST-ZF binds to the double stranded target sequence (TS) DNA 5'GGG GGG GGG 3'. Binding of the anti-GST antibody to GST-ZF has been shown by dot blot (Fig. 6.12). It should, therefore, be possible to develop a SPA assay to detect the interaction between DNA and the GST-ZF protein.

## 7.2 SPA ASSAY DEVELOPMENT USING GST-ZF.

### 7.2.1 Preliminary studies.

In order to show that no components of the assay buffer were capable of causing DNA to bind non-specifically to the SPA beads, an experiment was performed in which GST-ZF was removed from its buffer by centrifugation through a Micocon 30 microconcentrator (Amicon). The de-proteinated buffer was then used as a negative control (rather than the GST protein) in a SPA assay (2.11.6, order of addition A). The mean cpm/assay observed in the negative control (buffer only) was 7547.75 and in assays containing the GST-ZF protein 44218.75 (assays performed in quadruplet). Thus more DNA bound to the SPA beads in the presence of the GST-ZF protein than did in its absence. This indicates strongly that the GST-fused zinc finger protein is responsible for the binding of the DNA to the SPA beads.

The concentration of protein used in the assays was then varied (Table 7.1). A student's *t*-test was applied to determine the significance of the difference of DNA binding between assays containing GST and GST-ZF protein (Table 7.1).

| Protein concentration | Mean GST-ZF cpm | Mean GST cpm | Specific binding | Signal to noise ratio | P-values |
|---|---|---|---|---|---|
| 8.2nM | 34369.25 | 7886.75 | 26482.5 | 4.3:1 | 0.00618 |
| 16.4nM | 39076 | 4626 | 34450 | 8.4:1 | 0.00215 |
| 32.8nM | 49431.5 | 4428.25 | 45103.25 | 11.2:1 | 0.00016 |
| 65.3nM | 48102.5 | 4862.5 | 43240 | 9.9:1 | 0.00143 |
| 98.4nM | 33711.75 | 13232.5 | 20479.25 | 2.5:1 | 0.0358 |

Table 7.1 Effect of variation of protein concentration on the binding of DNA to the SPA beads. Specific binding = difference between GST and GST-ZF mean cpm. P values = significance of the difference between assays containing GST and GST-ZF. Total cpm / assay = 165169. Assays performed in quadruplet. Full data and statistical analysis shown in appendix XVIII.

A range of protein concentrations was added to the assay. At a concentration of 32.8nM (Fig. 7.1) the difference between the assays containing GST and those containing GST-ZF is the greatest, also the signal to noise ratio is highest. The difference between the experimental and negative controls is significant at all protein concentrations, at 10.8nM the difference is highly significant (P value <0.005). The significant P-values indicate that it is extremely unlikely that the increased binding of the DNA to the SPA beads seen in assays containing GST-ZF is due to chance.

This experiment provides further evidence that GST-ZF alone is responsible for the binding of the DNA to the SPA beads. The GST protein, which had been expressed and purified in the same way as GST-ZF, was used in the negative controls in all subsequent assays. The GST should not only contain the same reagents that are in the buffer but will also contain any contaminating proteins that may be in GST-ZF. Using GST in the negative control also demonstrates that the DNA does not bind to GST, but rather binds to the zinc finger motifs in the fusion protein.

### 7.2.2 Optimisation of the assay buffer.

The buffer in which the SPA was carried out in was lacking several ingredients which may improve the performance of the assay. The buffer was optimised with regards to;

DTT

BSA

$ZnCl_2$

DTT is a reducing agent. It is present in many protein buffers since it prevents the formation of di-sulphide bridges between cysteine residues. It is particularly important that the cysteine residues in GST-ZF do not form di-sulphide bridges that would inhibit the co-ordination of the zinc ion and therefore prevent the protein from folding correctly. If the protein did not fold correctly, it would not bind the DNA.

The effect of addition of 0.5mM and 5mM DTT to the assay was determined (Table 7.2). Assays were carried out using order of addition A (2.11.6). A student's *t*-test was applied to determine the significance between assays containing DTT with no DTT present (Table 7.2).

| [DTT] | Mean GST-ZF cpm | Mean GST cpm | Specific binding | Signal to noise ratio | P-values GST-ZF | P-values GST |
|---|---|---|---|---|---|---|
| 0.0mM | 43878.75 | 5871.25 | 37907.5 | 7.5:1 | | |
| 0.5mM | 37597.25 | 5910.25 | 31687.0 | 6.3:1 | 0.0304 | 0.9347 |
| 5.0mM | 37318.25 | 5306.5 | 32011.75 | 7:1 | 0.1105 | 0.2485 |

Table 7.2 The effect of DTT on the binding of DNA to SPA beads. Total cpm / assay = 164078.5. Specific binding = difference between GST and GST-ZF means. P value show the significance of the effect of DTT compared to assays containing no DTT. Assays were performed in quadruplet. Full data and statistical analysis are shown in appendix XIX.

The binding of DNA to the SPA beads seems to be largely unaffected by DTT, the signal to noise ratio and the difference between GST and GST-ZF assays remains similar in all cases. The effect of the DTT on the GST controls is not significant (P values > 0.05). There is a significant difference between the GST-ZF assays without DTT and containing 0.5mM DTT (P value <0.05). The results indicate that this is due to a reduction of specific binding of the DNA to the protein at 0.5mM DTT. At 5.0mM DTT there is no significant difference in the binding of DNA to protein compared with that observed in the absence of the DTT. It was concluded that there is no particular advantage associated with adding DTT to the assays.

BSA is frequently used (in blots and ELISA) to reduce non-specific binding. It might therefore have reduced the non-specific binding of the DNA to the SPA beads in this assay.

The effect of BSA on the binding of DNA to the SPA beads was determined by the addition of 5.0mg/ml BSA to the assay (Table 7.3). A student's *t*-test was applied to determine the significance of this effect.

| | Mean cpm at 0.0gm/ml BSA | Mean cpm at 5.0gm/ml BSA | P-values demonstrating the significance of the addition of BSA |
|---|---|---|---|
| GST | 7771.75 | 3104.75 | 0.158 |
| GST-ZF | 53304.75 | 41831 | 0.347 |
| Difference between GST and GST-ZF | 45533 | 38726.25 | |
| Signal :noise ratio | 6.8:1 | 13.5:1 | |

Table 7.3 Effect of BSA on the binding of DNA to SPA beads. Total cpm / assay = 127307.5. Assays were performed in quadruplet. Full data and statistical analysis are shown in appendix XX.

BSA had the effect of lowering the binding of the DNA in both GST and GST-ZF assays. The difference in counts between the positive and the negative assays was greatest in the absence of BSA, though the signal to noise ratio was improved by its presence. The P- values indicate that whilst BSA seems to have a positive effect on the assay this result is not significant. BSA was not therefore added to subsequent assays.

As zinc is necessary for the correct folding of zinc finger domains it was considered that the addition of a zinc salt to the assay may improve the specific binding of the DNA to the SPA bead / GST-ZF complex.

The effect of a range of $ZnCl_2$ concentrations (0.0mM, 0.1mM, 1.0mM and 10mM) was tested. A student's *t*-test was applied to determine if the differences between GST and GST-ZF assays were significant (Table 7.4, Fig. 7.2).

| [$ZnCl_2$] | mean cpm GST | mean cpm GST-ZF | Specific binding (cpm) | Signal: noise ratio | P-values |
|---|---|---|---|---|---|
| 0.0mM | 4543.75 | 39135 | 34591.25 | 8.6:1 | 0.003 |
| 0.1mM | 13293.5 | 31916.5 | 18623 | 2.4:1 | 0.001 |
| 1.0mM | 56421.25 | 103314.25 | 46893 | 1.8:1 | 0.006 |
| 10mM | 102040.25 | 125431.75 | 23391.5 | 1.2:1 | 0.040 |

Table 7.4 Effect of $ZnCl_2$ on the binding of DNA to SPA beads. Specific binding = difference between GST and GST-ZF means P-values demonstrate the significance the effect of ZF protein at various $ZnCl_2$ concentrations. Total cpm / assay = 112862. Assays are performed in quadruplet. Full data and statistical analysis are shown in appendix XXI

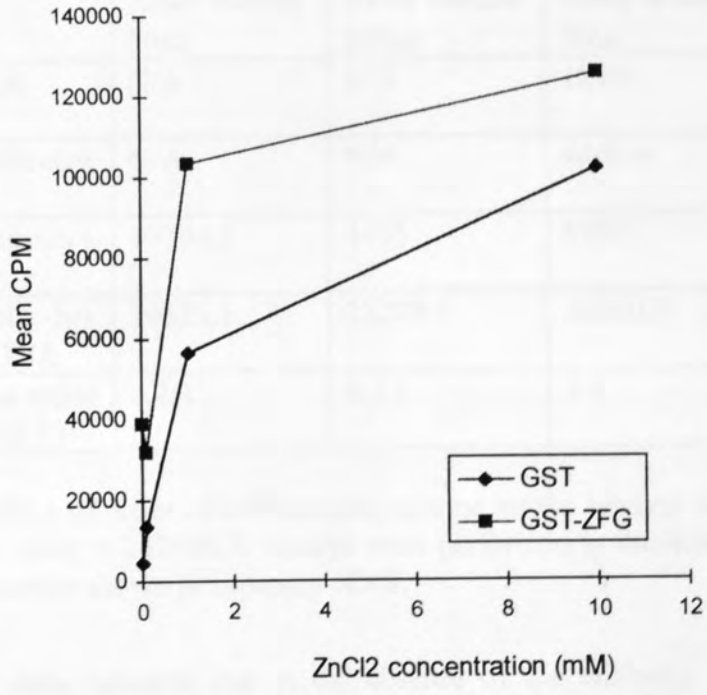



Figure 7.2. Graph showing the effect of $ZnCl_2$ on the binding of DNA to SPA beads in assays containing GST and GST-ZF. Full data shown in appendix XXI.

The addition of $ZnCl_2$ to the assay seems to raise both the specific and the non-specific binding of the DNA to the beads (Fig. 7.2). The signal to noise ratio is greatest (Table 7.4) in the absence of salt. There is a significant difference between GST and GST-ZF assay at all concentrations but their seems to be no particular advantage associated with the addition of salt.

### 7.2.3 Optimisation of assay format.

The order of addition of reagent to the assay may effect the specific and the non-specific binding of the DNA to the SPA beads. The effect of the order of addition of reagents to the assay and the assay volume were determined alongside an experiment testing the omission of several reagents from the assay. Two orders of addition were used order A (in which the protein and DNA are pre-coupled) and order B (in which the protein and bead are pre-coupled). Results are shown in Table 7.5.

| Reagent present in assay | Mean cpm, order of addition A | | Mean cpm, order of addition B | |
|---|---|---|---|---|
| | Assay volume 50µl | Assay volume 100µl | Assay volume 50µl | Assay volume 100µl |
| Beads+DNA | N/A | N/A | 12769 | 21116.7 |
| Beads+Antibody+ DNA | N/A | N/A | 4468.4 | 2841.7 |
| Beads+Antibody+ GST+DNA | 10794.3 | 4695 | 5907 | 3431.3 |
| Beads+Antibody+ GST-ZF+DNA | 34629.1 | 22279.7 | 30530.7 | 24441.3 |
| Signal:noise ratios (GST-ZF:GST) | 3.2:1 | 4.7:1 | 5:1 | 7:1 |

Table 7.5 Effect of order of addition and volume on the binding of DNA to SPA beads. Total cpm / assay = 242691.7. Assays were performed in duplicate. Full data appendix statistical analysis shown in appendix XXII.

The data indicates that in the absence of the antibody, the DNA binds non-specifically to the SPA beads. For this reason, the order of addition of reagents to the assays may be important in preventing non-specific binding. Order of addition B in which the beads are coated with antibody and protein prior to the addition of the DNA, is likely

to help prevent non specific binding. Order addition B is also more practical and was therefore adopted in subsequent assays. There may also be some advantage found in adding more antibody to the assays. Similarly an increase in assay volume appears to improve the signal to noise ratio. None of these results were statistically significant, nevertheless order of addition B was used in subsequent assays in order to prevent the non-specific binding of the DNA to the beads.

In view of the non-specific binding of the DNA to the SPA beads exhibited in the absence of antibody, the effect of antibody concentration in the assay was investigated. The effect of 0.0mg/ml, 1.0mg/ml, 1.5mg/ml, 2.0mg/ml and 3.0mg/ml antibody on the binding of DNA to SPA beads was tested (Table 7.6).

| Antibody concentration | DNA only | GST | GST-ZF |
|---|---|---|---|
| 0.0mg/ assay | 12637 | 10976.25 | 10097.2 |
| 1.0mg/ assay | 3988.8 | 4323.1 | 25713.1 |
| 1.5mg/ assay | 3076 | 2913.65 | 23952.75 |
| 2.0mg/ assay | 2850.6 | 3272.1 | 28137.8 |
| 3.0mg/ assay | 3015.95 | 3707.2 | 21870.65 |

Table 7.6 Effect of the amount of antibody on the non-specific binding of DNA to SPA beads. Total cpm / assay = 242691.7. Assays were performed in duplicate. Full data and statistical analysis are shown in appendix XXIII.

The data does not seem to show any particular advantage to the addition of more antibody than has previously been used in this assay (Table 7.8). The amount of antibody used in subsequent assays was therefore retained at 1.5mg/assay.

## 7.3 DEMONSTRATION OF THE SPECIFICITY OF THE BINDING OF GST-ZF TO THE TARGET SEQUENCE DNA.

Binding of the zinc finger proteins to the DNA is typically specific with regards to the DNA sequence. The GST-ZF protein is based on a zinc finger domain which binds specifically to the DNA sequence 5'GGG3' (3.1). GST-ZF should therefore bind specifically to TS DNA (Fig. 7.1). Alteration of bases in the target DNA binding site may alter the affinity of the binding of the protein to the DNA. The preference of the protein for various DNA binding sites may therefore be measured using a competition assay. The binding of the protein to labelled TS DNA in a competition assay was challenged with the presence of non-labelled competitor DNA. A drop in the scintillation counts obtained allowed the detection of binding of the protein to the non-labelled competitor in preference to the labelled TS DNA.

Assays were carried out using order of addition B (2.11.6). Experiments were set up in which the binding of the $^{33}$P-labelled TS DNA to the GST-ZF protein / SPA bead complex was challenged by the addition of non-labelled double stranded DNA. The binding was challenged with TS DNA, C1, C2, C3 and C4 (Fig 7.3). Data was statistically analysed using Microsoft Excel 5.0. Results of the analysis are shown in tables 7.7 and 7.8.

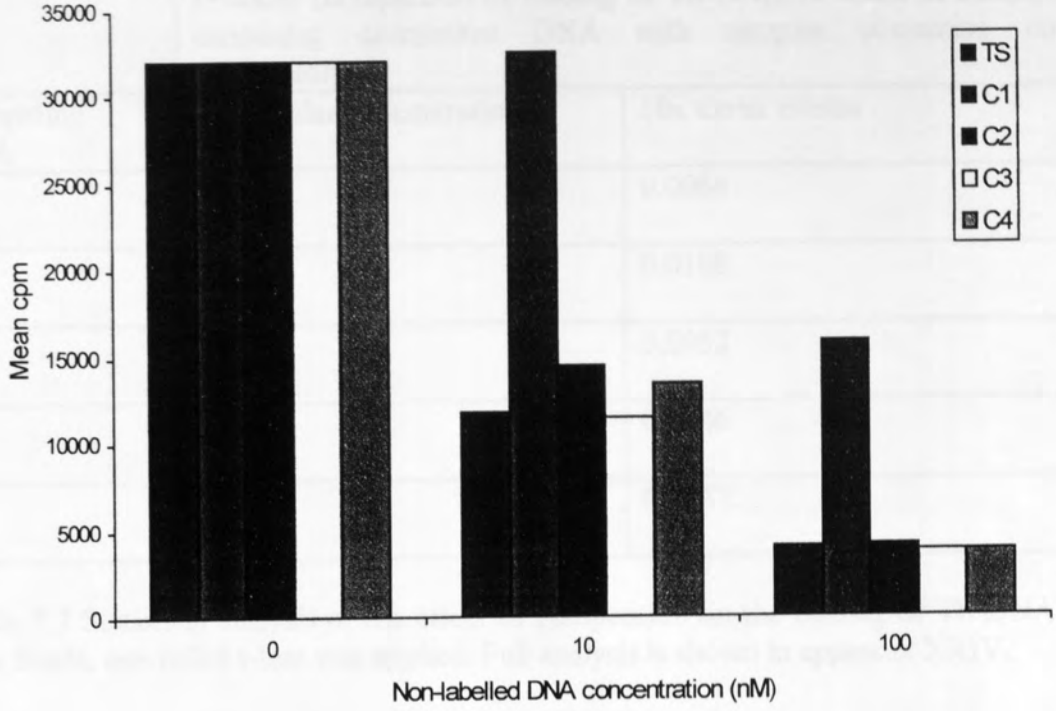



Figure 7.3 Graph to show the effect on the binding of TS DNA to SPA beads of addition on non-labelled competitor DNA. Background binding (binding in GST assays) is subtracted. Assays were performed in quadruplet. Full data is shown in appendix XXIV.

| | |
|---|---|
| TS | 5′TATATATATA GGG GGG GGG TTTTT 3′<br>3′ATATATATAT CCC CCC CCC AAAAA 5′ |
| C1 | 5′TATATATATA TTT TTT TTT TTTTT 3′<br>3′ATATATATAT AAA AAA AAA AAAAA 5′ |
| C2 | 5′TATATATATA GGG TTT GGG TTTTT 3′<br>3′ATATATATAT CCC AAA CCC AAAAA 5′ |
| C3 | 5′TATATATATA TTT GGG GGG TTTTT 3′<br>3′ATATATATAT AAA CCC CCC AAAAA 5′ |
| C4 | 5′TATATATATA GGG GGG TTT TTTTT 3′<br>3′ATATATATAT CCC CCC AAA AAAAA 5′ |

| | P-values (comparison of binding of TS DNA to beads in samples containing competitor DNA with samples containing no competitor) | |
|---|---|---|
| Competing DNA | Equi-molar concentration | 10x molar excess |
| TS | 0.0304 | 0.0064 |
| C1 | 0.3962 | 0.0188 |
| C2 | 0.0136 | 0.0052 |
| C3 | 0.0095 | 0.0046 |
| C4 | 0.0108 | 0.0047 |

Table 7.7 Statistical analysis of the effect of competition on the binding of TS DNA to SPA beads, one tailed t-test was applied. Full analysis is shown in appendix XXIV.

As expected, when non-labelled TS DNA was added to the assay at equi-molar concentration there was a significant (P-value< 0.05) decrease in the binding of the labelled DNA to the SPA beads. The amount of labelled DNA bound to the SPA bead/protein complex was reduced by approximately 50%. This demonstrates that the protein binds the labelled TS DNA and the non-labelled TS sequence DNA with equal affinity. When a 10x molar excess of non-labelled TS DNA was added, the binding of the labelled TS DNA was reduced (P-value< 0.05) to a level similar to that seen in the negative controls.

At equi-molar concentration the binding of TS DNA to the SPA bead/protein complex is unaffected by the presence of C1 DNA. At a 10x molar excess of C1 DNA there is an approximately 50% reduction (P-value< 0.05, Table 7.7) in the binding of the TS DNA to the SPA beads. This show that the alteration of the binding site from 5' GGG GGG GGG 3' to 5' TTT TTT TTT 3' significantly impairs the binding of the DNA to the SPA beads. This would be expected as the binding sites for all three zinc finger motifs in the protein have been altered.

At equi-molar concentration C2, C3 and C4 DNA significantly (P-values< 0.05, Table 7.10) reduce the binding of the TS DNA to the SPA bead/protein complex. Approximately 50% of the TS DNA seem to have been competed off the protein by these competitors. This implies that the protein binds as efficiently to these competitors as it does to the TS DNA.

| Non-labelled competing DNA | P-Value | |
|---|---|---|
| C1 | 0.0005 | Preference for TS over C1 |
| C2 | 0.2222 | Equal reference for both TS and C2 |
| C3 | 0.0422 | Slight preference for C3 over TS |
| C4 | 0.0592 | Equal reference for both TS and C2 |

Table 7.8 comparison of the binding of various oligonucleotides to the SPA bead/ protein complex at equi-molar concentrations, t-tests performed between the cpm observed in samples containing non-labelled TS competitor and those containing C1, C2, C3 and C4. Statistical analysis shown in appendix XXIV.

Two tailed t-tests (Table 7.11) were performed between amount of TS bound to the protein in the presence of non-labelled 1x TS DNA and the amount bound in the presence of non-labelled TS competitor DNA. The data suggests that the protein will preferentially bind TS DNA rather than C1 DNA (P-value< 0.005). There is no significant preference for the TS DNA over C2. The protein seems to bind significantly better (P-value< 0.05) to C3 than to the TS. There is no significant preference for the TS DNA over C4. As competitor oligonucleotides were not labelled they were also not spun through a C10 column (Sigma) in order to remove unincorporated label. The oligonucleotides used to make C2, C3 and C4 were gel purified in order to remove the products of incomplete synthesis. There may therefore be traces of urea in these samples. This may interfere with the assay and result in non-specific binding of the DNA to the SPA beads, accounting for the unexpected binding of the zinc finger protein to "inappropriate" sites.

## 7.4 CALCULATION OF A DISSOCIATION CONSTANT FOR THE INTERACTION BETWEEN GST-ZF AND TS DNA.

### 7.4.1 Identification of the range of concentrations over which assays should be carried out in order to calculate a dissociation constant.

The zinc finger protein on which the zinc finger domains of GST-ZF were based has a dissociation constant with regards to its target DNA sequence of 2nM (Desjarlais and Berg, 1993; Shi and Berg, 1995a). The synthetic zinc finger protein was therefore expected to have a $K_d$ within the low nM range. In order to predict the concentration range required for an experiment to calculate the $K_d$, several preliminary experiments were carried out.

The assay was performed using the same conditions as those employed in previous experiments (2.11.6, order of addition B). Protein (16.4nM) was allowed to interact with the DNA over a wide range of DNA concentrations from 0.1-100nM (Fig. 7.4).

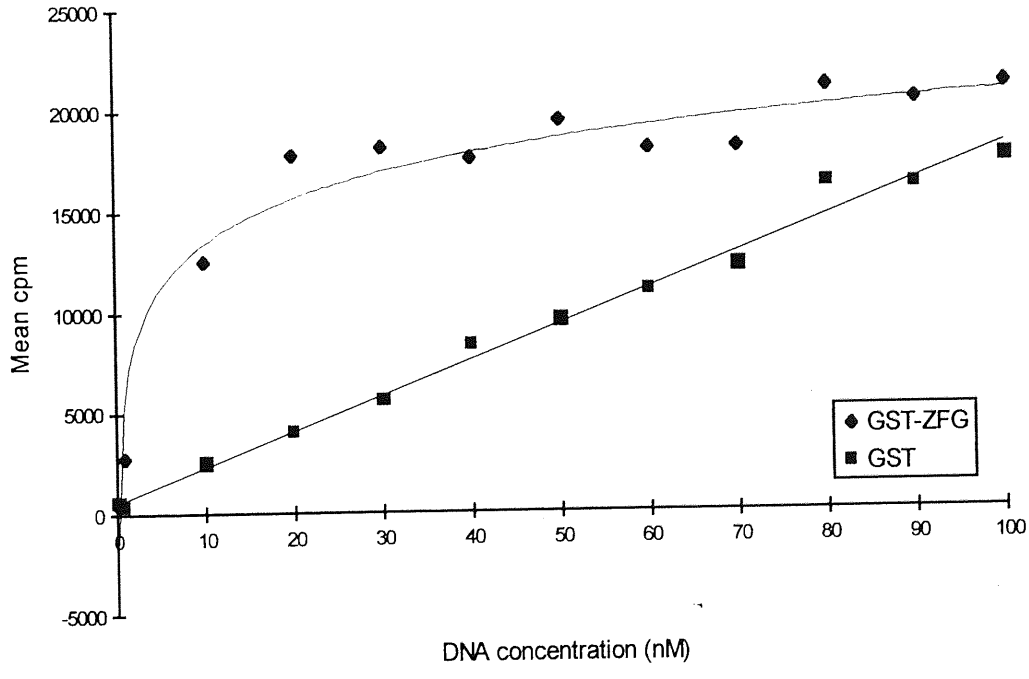



Figure 7.4 Graph showing the effect of a wide range of DNA concentrations on the SPA. Assays performed in quadruplet, full data shown in appendix XXV.

The graph shows that the assay is essentially saturated at about 20-30nM DNA. The increase in binding of DNA to beads after this point is likely to be due to the increase in non-specific binding. This indicates that in order to calculate a dissociation constant for the DNA/GST-ZF interaction the range of concentrations required is from approximately 1-20nM.

To further identify the range of concentrations over which an experiment to calculate the dissociation constant should be performed, assays were carried out over lower ranges of both DNA and protein concentrations. Initially assays were performed over a range of protein concentrations, with a constant DNA concentration of 10nM (Fig. 7.5).

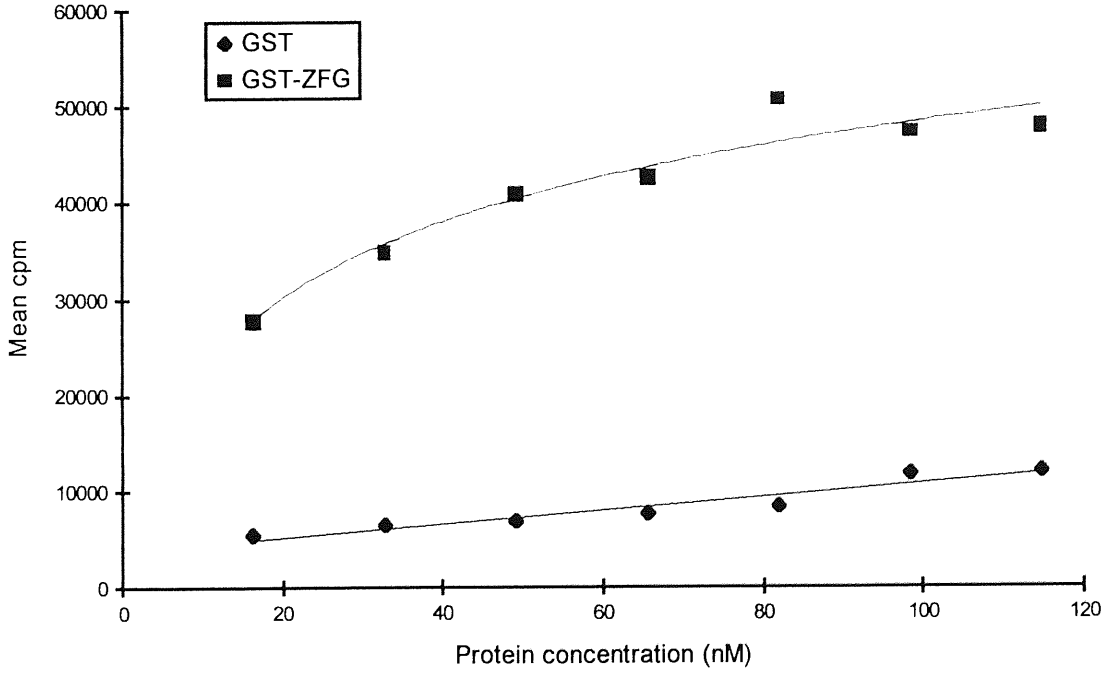



Figure 7.5 Graph showing the effect of a range of protein concentrations on the binding of DNA to SPA beads. Assays were performed in quadruplet. Full data is shown in appendix XXVI.

The effect of varying the DNA concentration over a low range was demonstrated using a constant protein concentration of 16.4nM (Fig. 7.6).

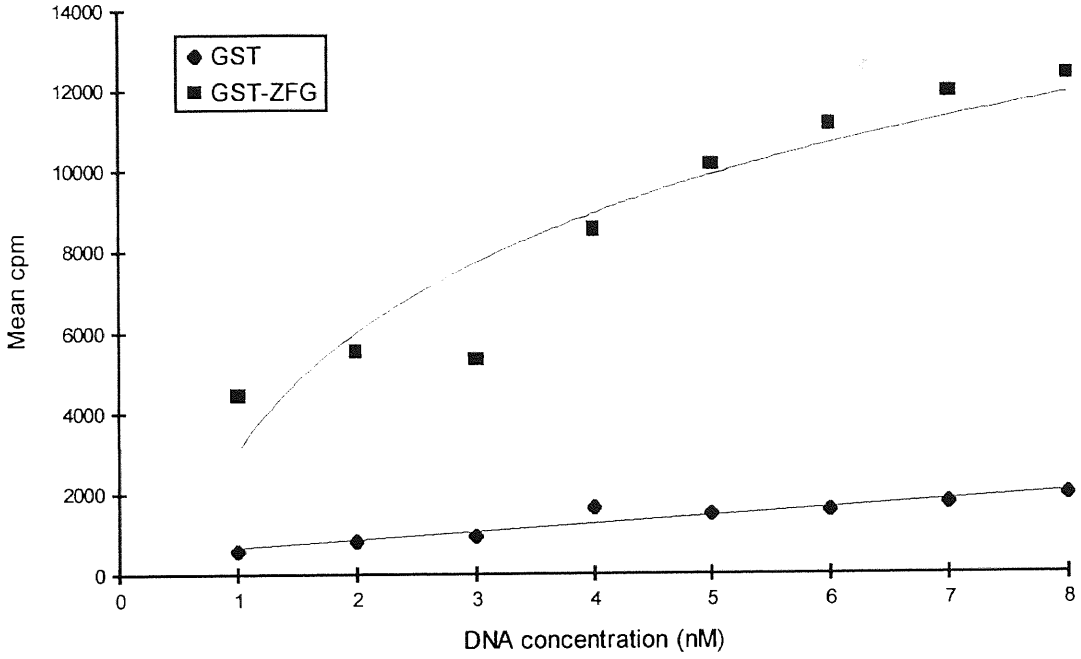



Figure 7.6 Graph showing the effect of a low range of DNA concentrations. Assays were performed in quadruplet. Full data is shown in appendix XXVII.

### 7.4.2 Calculation of dissociation constant.

The previous experiments (7.4.1) revealed that a range of assays with concentrations of DNA varying from approximately 1-20nm would be necessary in order to gather the data required to calculate the dissociation constant between the protein and its DNA target sequence. A range of assays were performed at 16.4nM protein with DNA concentrations varying from 1-16nM (Fig. 7.7).

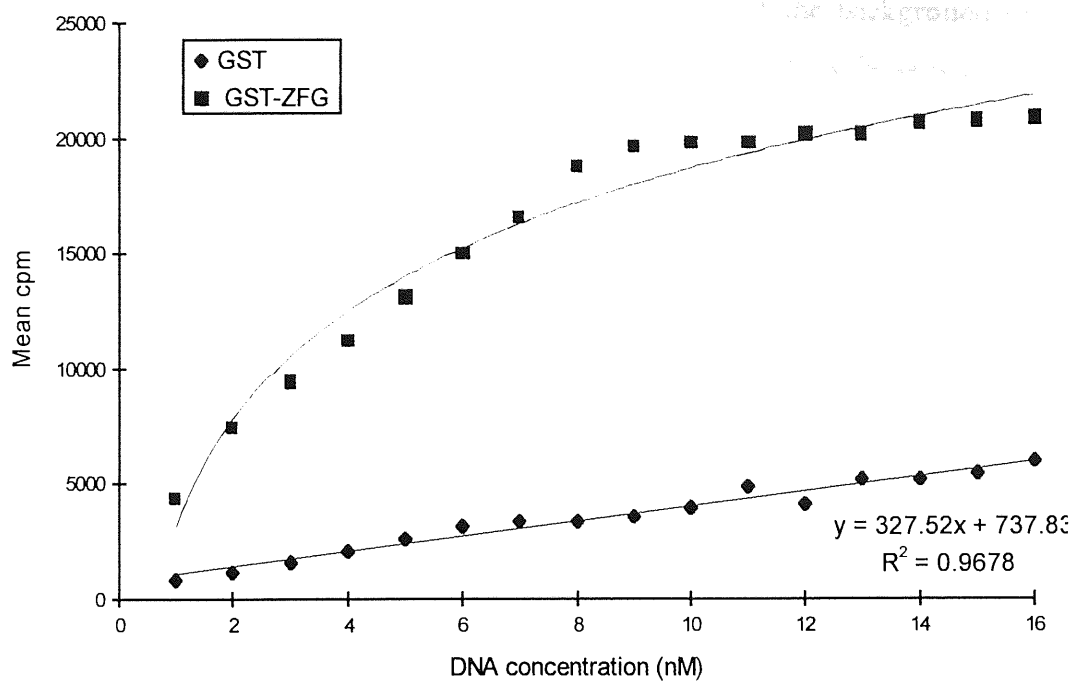



Figure 7.7 Graph showing the SPA data used in the calculation of the dissociation constant. Assays were performed in quadruplet. Full data is shown in appendix XXVIII.

The data from this experiment was plotted using KaleidaGraph (Synergy software), and the dissociation constant calculated by fitting the following equation to the data.

$B_{max}$ = The maximum amount of ligand which may bind
n = Hill coefficient

$$\text{Bound ligand} = \frac{B_{max}\ \text{Free ligand}^n}{\text{Free ligand}^n + Kd^n}$$

Where; m0 = Free ligand
m1 = $B_{max}$
m2 = n (Hill coefficient)
m3 = Kd

The dissociation constant was calculated both without the background (assays containing only GST) subtracted from the specifically bound DNA (assays containing GST-ZF) (Fig. 7.8), and with the background subtracted (Fig. 7.9). The values - background binding of the DNA to the beads was calculated by plotting a graph of the GST assays and fitting a straight line to it. The equation for this straight line was then subtracted from the values obtained from the assays containing GST-ZF.

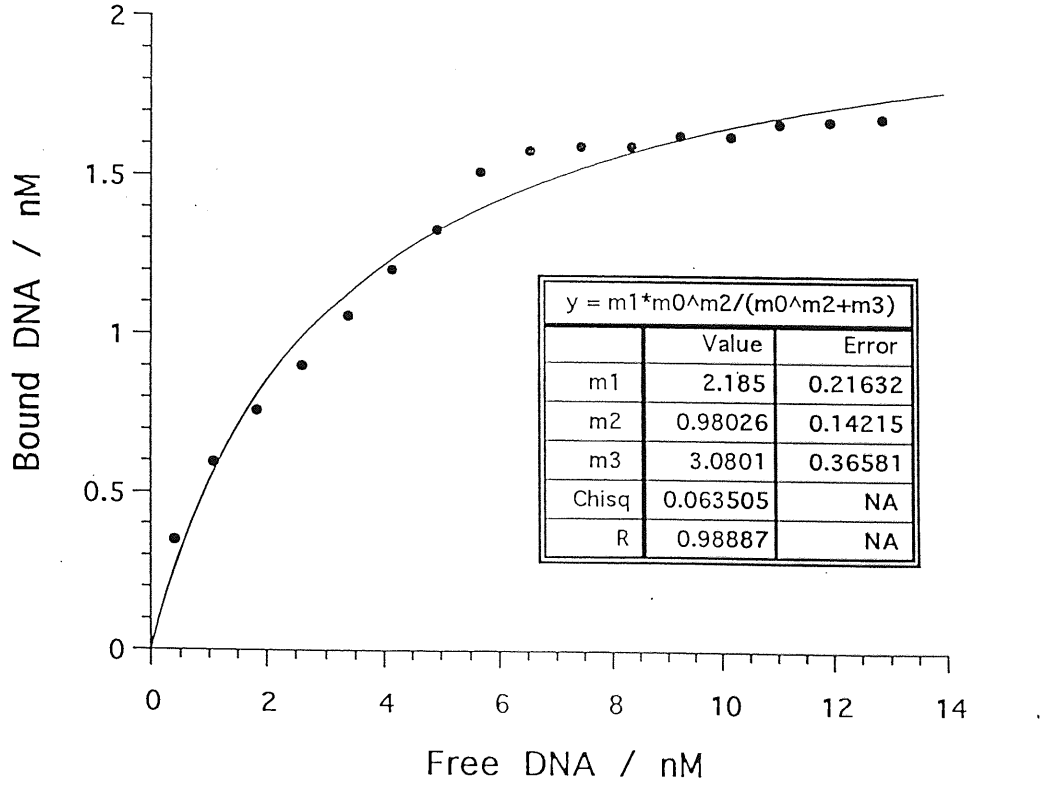



Figure 7.8 Calculation of this dissociation constant is without background subtracted, calculation of values plotted on graph shown in appendix XXVIII.

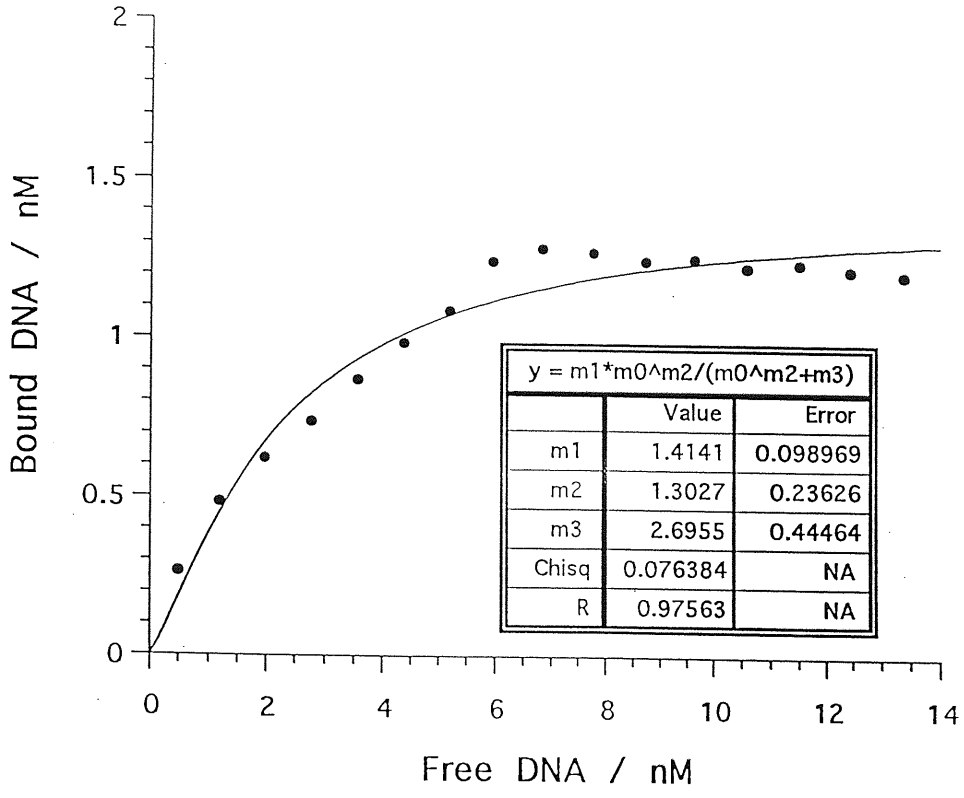



Figure 7.9 Calculation of this Dissociation constant is with background subtracted, calculation of values plotted on graph shown in appendix XXVIII.

7.5 DISCUSSION.

### 7.5.1 Development of the SPA to detect GST-ZF / DNA interactions.

The binding of radio labelled DNA to the protein A-coated PVT SPA beads was shown to be induced in the presence of both anti-GST antibody and GST-ZF. A correlation was shown between increased concentrations of GST-ZF protein and increased binding of the DNA to the beads (7.2.1).

Two different negative controls were used in order to confirm that the binding of the DNA to the SPA beads in this assay was due only to the presence of the zinc finger protein. By using the buffer from GST-ZF (7.2.1) as the negative control in an assay it was confirmed that no other small molecules were responsible for the binding of the DNA to the SPA beads. Secondly the use of GST (synthesised and purified in the same way as GST-ZF) as a negative control showed that no other large molecule (such as a contaminating *E. coli* protein), or GST itself were causing the DNA to bind to the SPA beads.

The assay buffer was optimised with regards to DTT, BSA and $ZnCl_2$ (7.2.2). The addition of none of these reagents seemed to confer any particular advantage, either in lowering the non-specific binding of the DNA to the beads or increasing the specific binding.

The order of addition of reagent to the assay was tested (7.2.3) revealing that the addition of antibody and protein to the SPA mix prior to the addition of the DNA may lower the non-specific binding of the DNA to the beads. This experiment also indicated that there may be some advantage to increasing the amount of antibody added to the assay. Further investigation however did not seem to indicate that this may be advantageous.

### 7.5.2 Specificity of binding of GST-ZF.

The specificity of the binding of GST-ZF to the TS DNA was demonstrated by challenging the labelled DNA with non-labelled competitor, at both equi-molar concentration and at a 10x molar excess (7.3).

As expected there is a preference for the TS DNA over C1 DNA (which has all three zinc finger binding sites altered). It would also be expected that the zinc finger protein would bind C2, C3 and C4 (all of which have only one zinc finger binding site

altered) with a lower affinity that it does TS DNA. C2 with the central binding site altered would be expected should show the least affinity (out of C2, C3, and C4) for the protein (Choo, 1998). In this experiment however it seems to have the same affinity for the protein as the TS DNA. Likewise C4 seems to bind to the protein with a similar affinity to the TS DNA. C3 however seems to bind the protein with a slightly higher affinity than the TS DNA.

These results suggest that the binding of GST-ZF is sequence specific with regards to the alteration of the binding site from 5' GGG GGG GGG 3' to 5' TTT TTT TTT 3'. It was not expected that the zinc finger protein would bind quite so well to the DNA sequences with one zinc finger binding site altered. This may be a true reflection of the binding of these sequences to the GST-ZF protein. It is possible however that these results are due to contaminating urea in C2 C3 and C4 interfering with the assay.

### 7.5.3 Calculation of dissociation constant.

The range of concentrations over which an experiment to calculate the Kd of the TS DNA/ GST-ZF interaction should be carried out was predicted from the results of several preliminary experiments (Fig. 7.4, 7.5 and 7.6).

The specific binding of the DNA to the SPA beads observed in the presence of GST-ZF increases in response to increased DNA concentration, at high DNA concentrations binding reaches a plateau (Fig 7.4). The response of the non specific binding of the DNA to the SPA beads is however linear. As the concentration of the DNA in the assay increases the level of non-specific binding represents disproportionately more of the specific binding (Fig 7.4). Non-specific binding (in the GST assays) seems to be constantly approximately 10% of the total counts. In the presence of GST-ZF the concentration of DNA available to non-specifically bind will be lower than that in the GST counterpart (as a certain percentage of the DNA is already specifically bound). The amount of non-specific binding is therefore probably not as great in the assays containing GST-ZF as it is in the assays containing GST.

The dissociation constant was calculated both with the background (binding in the presence of only GST) subtracted (7.9) and without the background subtracted (7.8). Without the background subtracted the fit of the equation to the data (R= 0.989) was higher than the fit with the background subtracted (R= 0.975). The Kd estimation from the data without background subtracted was slightly more conservative (3.07nM) than

the value from the data with the background binding subtracted (2.69nM). In both cases the Hill coefficient is predicted to be approximately 1 implying that as expected the interaction between GST-ZF and the DNA is 1:1. The Vmax (amount of GST-ZF binding sites available) in both cases is far lower than the amount of protein added to the assay. This is probably due to the fact that not all of the protein present is functional and, or not all of the functional protein placed in the assay is bound to the SPA beads. Low binding of the protein to the SPA beads could be due to lack of protein A on the beads, lack of functional antibody or poor interaction between the protein and the antibody.

For the reasons outlined above it is probably justifiable to not subtract the background (GST assay values) from the specific binding values in the calculation of the Kd. Particularly as the calculation of the Kd without subtracting the background provides the more conservative estimate.

### 7.5.4 Summary.

- The binding of the DNA to the SPA beads in the presence on GST-ZF was shown to be due to the presence of the protein not another factor in the buffer, and a correlation between increased protein concentration and increased binding of the DNA to the SPA beads was demonstrated.

- The assay was optimised with regards to DTT, BSA and $ZnCl_2$. The addition of these reagents to the assay seemed to have no particular benefit.

- Order of addition of reagents to the assay was investigated and this revealed that in the absence of the antibody high non-specific binding of the DNA to the SPA beads was observed. This suggested that it was beneficial to pre-couple the antibody and the protein to the beads before the addition of DNA.

- The binding of GST-ZF to its target sequence is specific at least with respect to the alteration of the entire binding site. The alteration of individual binding sites however seems to have less effect on the binding of GST-ZF to the DNA

- It was possible to calculate a Kd of 3.08nM with respect to the binding of the DNA to the TS DNA. The fit of the equation to the data was satisfactory (R= 0.989), and the Hill coefficient of 0.989 suggests that there is a 1:1 interaction between the protein and the DNA.

# 8. GENERAL DISCUSSION.

## 8.1 SUMMARY AND DISCUSSION OF WORK COMPLETED.

### 8.1.1 Summary of the work completed.

A zinc finger gene has been designed, and the protein product synthesised by the controlled expression of its gene in *E. coli*. The zinc finger protein contains three zinc finger motifs each is based on a zinc finger consensus sequence (Krizek *et al.*, 1991). The DNA contacting residues used within each domain are those found in the first zinc finger of the transcription factor Sp1 (Shi and Berg, 1995a). The zinc finger motifs are preceded by a consensus linker sequence, which is necessary for the binding of the protein in the correct orientation to DNA. Also the linker sequence allows the three zinc finger motifs to be correctly "spaced" in order to bind to their 3 base pair subsites (Choo and Klug, 1993). Use of a consensus zinc finger framework with predetermined contacting residues should give predictable DNA binding preferences (Shi and Berg, 1995a). The synthetic zinc finger protein designed here is designed to bind to the dsDNA sequence 5' GGG GGG GGG 3'. In the presence of zinc the consensus zinc finger domains fold into the ββα structure (Kim and Berg, 1996) this formation (described in 1.2.2) is typical of zinc finger proteins.

The gene encoding the zinc finger protein uses, where possible, the optimal codon bias for maximal expression in *E. coli* (Wada, *et al.*, 1992). Initially unique restriction enzyme sites were incorporated at each end of the gene allowing the gene to be inserted into the expression vector pT7-7 and a His tag was incorporated for the purification of the protein product. Flanking the central zinc finger domain there are two *Age* 1 sites, which enable the removal or replacement of the central motif. Within the central zinc finger either side of the area containing the contacting residues there are two unique sites (5' *Bsi* W1 and 3' *Hind* III). These sites allow the cassette mutagenesis of the region responsible for the specificity of the binding of the protein to DNA.

The gene was synthesised using three different methods, all of which relied on the use of oligonucleotides. Difficulties in isolating a copy of the gene, which did not contain single or multiple base deletions, were experienced. The reason why such a high frequency of copies of the gene which contained deletions were isolated is not clear, but

it is likely that some quality unique to the oligonucleotides used, possibly in combination with the sequence of the gene, caused the mutations.

The synthetic zinc finger protein described here was initially attached to a His tag, to allow it to be purified by metal affinity chromatography. Initially an assay was attempted in which biotinylated DNA was attached to streptavidin-coated PVT SPA beads, in order to detect the binding of labelled DNA. This failed to detect interaction between the protein and the DNA, probably due to the low specific activity of the protein and steric hindrance from the SPA bead. To overcome these problems it was intended to attach the protein to protein A-coated PVT SPA beads via an anti-His antibody, in order to detect the binding of labelled DNA. Since none of the antibodies available were capable of recognising the zinc finger protein, the protein was therefore linked directly to yttrium silicate SPA beads. This assay format seemed to function but was unfortunately prone to huge non-specific binding of the DNA to the SPA beads in the presence of small quantities of divalent cations. This format was therefore considered impractical because there was a high chance of producing false positives from contamination with divalent cations.

The zinc finger gene was then fused to a Glutathione S-Transferase (GST) gene in order that the protein product could be purified by Glutathione affinity chromatography. The GST-fused protein was linked to protein A-coated PVT SPA beads via an anti-GST antibody, in order to detect the binding of labelled DNA to the protein. A SPA assay using the GST-fused zinc finger gene proved far more successful. This format was also sensitive to salt concentrations although the non-specific binding was not increased to the same levels as the specific binding. The reason why in both assay formats non-specific binding of the DNA to the SPA beads is induced by salt remains unsolved.

The binding of the zinc finger protein to the target DNA was demonstrated using both "traditional" methods (dot blot and gel retardation assay) and the SPA technology. The zinc finger protein was shown to bind with some degree of specificity to its DNA target sequence. The predicted mode of binding of the protein to the DNA is discussed below.

### 8.1.2 Discussion of the observed interaction between the synthetic zinc finger protein and DNA.

The GST fused zinc finger gene was found to bind to its target DNA sequence with a dissociation constant of approximately 3nM. The binding of the zinc finger protein to the DNA is likely to stabilised by six hydrogen bonds made to the to the phosphate backbone and a minimum of nine base contacts (Fig 8.1).

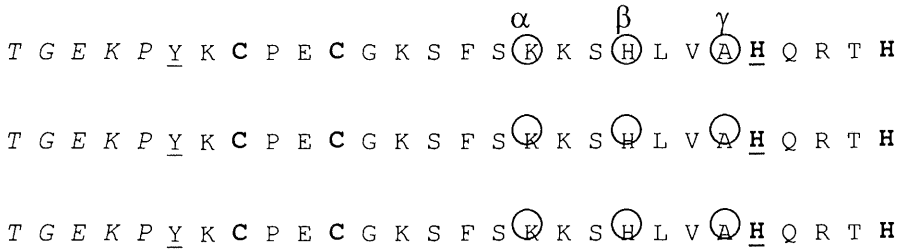



Figure 8.1 Amino acid sequence of the synthetic zinc finger peptide. Linker sequence is shown in Italics. Zinc co-ordinating residues shown in bold. Contacting residues α, β and γ are indicated by the letter above the amino acid sequence. Predicted base contacting residues are encircled (Pavlatich and Pabo, 1991; Kim and Berg, 1995; Kim and Berg, 1996; Yokono *et al.*, 1998). Predicted phosphate backbone contacting residues are underlined (Kim and Berg, 1996).

The binding of the zinc finger protein to the DNA seems to be sequence specific. The protein shows a strong preference for its target DNA sequence over a competitor in which all three binding sites have been altered. The zinc finger protein also unexpectedly bound to a competitor sequence used in a gel retardation assay (4.5.2) and possibly binds to competitor sequences used in SPA in which one of the three subsites was altered (7.3).

There is no x-ray crystal structure available that demonstrates how the contacting residues of the first zinc finger of Sp1 recognise their three base subsite. It has been suggested that only the lysine at position α contributes to the specificity of the binding of the Sp1 finger 1 domain to the DNA (Yokono *et al.*, 1998). X-ray crystal structures confirm that Lysine residues in zinc finger motifs are capable of making either one or two base specific contacts to G (Kim and Berg, 1996). This may be the case but comparison to the x-ray crystal structure of other zinc finger domains suggests that the histidine at position β will also make a base specific contact. Histidine residues in zinc

finger motifs have been shown to interact with both A (Kim and Berg, 1996) and G (Pavlatich and Pabo, 1991). The alanine residue at position γ is unlikely to make a base contact to the DNA (Yokono *et al.*, 1998). The serine residue, however which lies immediately to the amino side of residue β is capable of contacting the DNA. But as serine can act as a hydrogen bond acceptor or a hydrogen bond donor this contact is unlikely to contribute to the specificity of the binding between the protein and the DNA (Kim and Berg, 1995). Based on these findings the way in which I predict the zinc finger protein interacts with the target sequence DNA is shown in figure 8.2.

Although zinc finger 1 of the Sp1 protein binds to the consensus sequence 5'GGG 3' (Desjarlais and Berg, 1993) the structure and data available (discussed above) on the behaviour of the contacting residues it utilises suggest that it could bind to 5' N A/G G 3'. Possible DNA targets to which the synthetic zinc finger protein could bind are 5' N A/G G N A/G G N A/G G 3'. This sequence is quite degenerate and probably, in part, accounts for the binding of the protein to sequences that it was not initially expected to bind.

Despite the fact that the potential binding site of this protein may be "degenerate" it still binds with a higher specificity to its target sequence than to unrelated DNA sequences (7.3). The binding of the zinc finger protein to DNA sequences it would not be expected to bind to occurred in competition experiments were the comparison of binding was not entirely fair. Where the protein bound the competitor oligonucleotide in the gel shift experiment the concentration of competitor was much higher than the concentration of the target DNA (4.5.2). The binding of the zinc finger protein to competitor DNA in the SPA assay may be due to the different preparation of the competitor and target DNA (7.3). Further experimentation is now required in order to confirm that the binding of the zinc finger protein to the DNA is affected by the alteration of the zinc finger subsites within the target DNA.

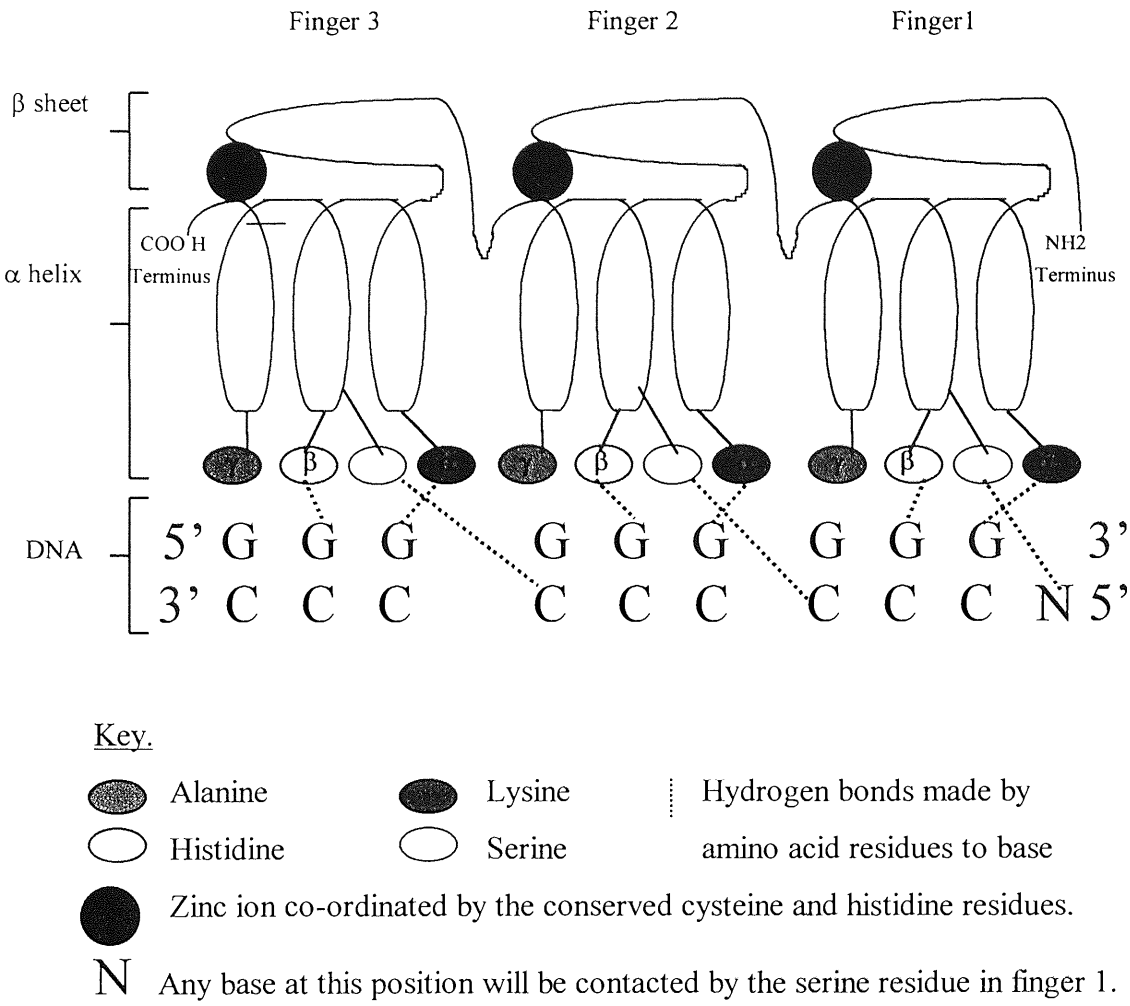



Figure 8.2 Diagram showing the predicted amino acid base contacts made by the synthetic zinc finger protein (Pavletich and Pabo, 1991; Kim and Berg, 1995; Kim and Berg, 1996; Yokono *et al.*, 1998).

In conclusion I believe that the zinc finger protein described here is capable of binding DNA with a high affinity and specificity. This zinc finger protein in conjunction with the scintillation proximity assay will provide a suitable system for future work.

## 8.2 FUTURE WORK.

The zinc finger protein and assay developed here may be used for the elucidation of a code of interaction between zinc fingers and DNA, and for the selection of zinc finger proteins with novel ligand binding specificities.

In the elucidation of a code of interactions between zinc finger proteins and DNA the residues responsible for the specificity of the binding of the protein to the DNA will be altered by controlled randomisation. This may be achieved by cassette mutagenesis of the central zinc finger in the protein. The combinatorial libraries produced can then be analysed using the high throughput SPA developed here.

The SPA developed here is extremely versatile and may be used not only for the detection of zinc finger DNA interaction. A variety of ligands may be tested using this assay including RNA and proteins.

# 9.0 REFERENCES

Baleja, J. D., Marmorstein, R., Harrison. S. C., Wagner, G. (1992) Solution structure of the DNA-binding domain $Cd_2$-GAL4 from *Saccharomyces cerevisiae. Nature*, **356**, 450-453.

Barlow, P. N., Luisi, Bb., Milner, A., Elliott, M., Everett, R., (1994) Solution structure of the C3HC4 domain by $^1$H-nuclear magnetic resonance spectroscopy: a new class of zinc-finger. *J. Mol. Biol.* **237**, 201-211.

Becker, K. G., Nagle, J. W., Canning, R. D., Biddison, W. E., Ozato, K., Drew, P. D. (1995) Rapid isolation and characterization of 118 novel $C_2H_2$-type zinc finger cDNAs expressed in human brain. *Human Molecular Genetics*, **4**, 685-691.

Berg, J. M. (1987) Proposed structure for three zinc binding domains from transcription factor IIIA end related proteins. *Proc. Natl. Acad. Sci. USA*, **85**, 99-102.

Berg, J. M. (1990) Zinc finger domains: hypotheses and current knowledge. *Annu. Rev. Biophys. Biophys. Chem*, **19**, 405-421.

Berg, J. M. (1993) Zinc-finger proteins. *Current Opinion in Structural Biology*, **3**, 11-16.

Caddick, M. X., Greenland, A. J., Jepson, I., Krause, K., Qu, N., Riddell, K. V., Salter, M. G., Schuch, W., Sonewald, U., Tomsett, A. B. (1998) An ethanol inducible gene switch for plant used to manipulate carbon metabolism. *Nature Biotechnology*, **16**, 177-180.

Call, K. M., Glaser, T., Ito, C. Y., Buckler, A. J., Pelletier, J., Haber, D. A., Rose, E. A., Kral, A., Yeger, H., Lewis, W. H., Jones, C., Housman, D. E. (1990) Isolation and characterization of a zinc finger polypeptide gene on the Human Chromosome 11 Wilms' Tumor Locus. *Cell*, **60**, 509-

Cheng, C. & Young, E. T. (1995) A single amino acid substitution in zinc finger 2 of Adr1p changes its binding specificity at two positions in UAS1. *J. Mol. Biol*, **251**, 1-8.

Chongwoo, A. K., Berg, J. M. A 2.2 A resolution crystal structure of a designed zinc finger protein bound to DNA. *Nature structural biology*, **3**, 940-945.

Choo, Y. & Klug, A. (1993) A role in DNA binding for the linker sequences of the first three zinc fingers of TFIIIA. *Nucleic Acids Research*, **21**, 3341- 3346.

Choo, Y. & Klug, A. (1994a) Towards a code for the interactions of zinc fingers with DNA: Selection of randomized fingers displayed on phage. *Proc. Natl. Acad. Sci. USA*, **91**, 11163-11167.

Choo, Y. & Klug, A. (1994b) Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc. Natl. Acad. Sci. USA*, **91**, 11168-11172.

Choo, Y. & Klug, A. (1997) Physical basis of a protein-DNA recognition code. *Current Opinion in Structural Biology*, 7, 117-125.

Choo, Y. (1998) End effects in DNA recognition by zinc finger arrays. *Nucleic Acids Research.*, **26,** 554-557.

Choo, Y., Castellanos, A., Garcia-Hernandez, B., Sanchez-Garcia, I., Klug, A. (1997) Promoter-specific Activation of a Gene Expression Directed by Bacteriophage-selected Zinc fingers. *J. Mol. Biol.*, **273,** 525-532.

Choo, Y., Sanchez-Garcia, I. & Klug, A. (1994) In *vivo* repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature*, **372**, 642-645.

Choo, Y., Schwabe. W. R. (1998) All wrapped up. *Nature structural biology*, **5**, 253-255.

Choo, Y., Schwabe. W. R. (1998) Recognition of DNA methylation by zinc fingers. *Nature structural biology*, **5**, 264-265.

Christy, B. A., Lau, L. F., Nathans, D. (1988) A gene activated in the mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences. *Proc. Natl. Acad. Sci. USA*, **85**, 7857-7861.

Crick, F. H. C. (1958) On protein synthesis *Symp.Soc. Exp. Biol.*, **12**,138-163.

Darnell, J. E., *et al.*, (1982) Variety in the level of gene control in eucaryotic cells. *Nature*, **297**, 365-371.

Derman, E., Krauter, K., Walling, L., Weinberger, C., Ray, M., Darnell, J E., (1981) Transcriptional control in the production of liver specific mRNAs. *Cell* **23**, 731-739.

Desjarlais, J. R. & Berg, J. M. (1992a) Redesigning the DNA binding specificity of a zinc finger protein: A data base-Guided approach. *Proteins structure function and gentetics*, **12**,101-104.

Desjarlais, J. R. & Berg, J. M. (1992b) Towards rules relating zinc finger protein sequences and DNA binding site preferences. *Proc. Natl. Acad. Sci. USA*, **89**, 7345-7349.

Desjarlais, J. R. & Berg, J. M. (1994) Length encoded multiplex binding site determination: Application to zinc finger proteins. *Proc. Natl. Acad. Sci. USA*, **91**, 11099-11103.

Desjarlais, J. R. & Berg, J. M. (1993) Use of a zinc finger consensus sequence framework and specificity rules to design specific DNA-binding proteins. *Proc. Natl. Acad. Sci. USA*, **90**, 2256-2260.

Dutnall, R. N., Neuhaus, D. & Rhodes, D. (1996) The solution structure of the first zinc finger domain of SW15: A novel structural extension to a common fold. *Structure*, **4**, 599-611.

Elrod-Erickson, M., Benson, T. E., Pabo, C. O. (1998) High resolution structures of variant Zif268 protein-DNA complexes: implications for understanding zinc finger-DNA recognition. *Structure*, **6**, 451-464.

Elrod-Erickson, M., Rould, M. A., Nekludova, L., Pabo, C. O. (1996) Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interaction. *Structure*, **4**, 1171-1180.

Engels, J. W., Uhlmann. E., (1989) Gene Synthesis. *Angew. Chem. Int.*, **28**, 716-735.

Fairall, L., Schwabe, J. W. R., Chapman, L., Finch, J. T. & Rhodes, D. (1993) The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger /DNA recognition. *Nature*, **366**, 483-485.

Freedman, L. P., Luisi, B. F. (1993) On the mechanism of DNA binding by nuclear hormone receptors: A structural and functional perspective. *J. Cell Biochem*, **51**, 140-150.

Freemont, P. S., Hanson, I. M., Trowsdale, J. (1991) A novel Cysteine-rich sequence motif. *Cell*, **64**, 483-484.

Freyd, G., Kim, S. K., Horvitz, H. R. (1990) Novel cysteine-rich motif and homeodomain in the product of the *Caenorhabditis elegans* cell line lineage *lin-II*. *Nature*, **344**, 876-879.

Gidoni, D., Dynan, S. W., Tijian, R. (1984) Multiple specific contacts between a mammalian transcription factor and its cognate promoters. *Nature*, **312**, 409-412.

Gogos, A. J., Jin, J., Wan, H., Kokkinidis, M., Kafatos, F. C. (1996) Recognition of diverse sequences by class I zinc fingers: Asymmetries and indirect effects on specificity in the interaction between CF2II and A+T rich sequence elements. *Proc. Natl. Acad. Sci. USA,* **93**, 2159-2164.

Greisman, H. A., Pabo, C. O. (1997) General strategy for selecting High-Affinity Zinc finger proteins for diverse DNA target sites. *Science*, **275**, 657-661.

Hommel, U., Zurini, M., Luyten, M., (1994) Solution structure of a cysteine rich domain of rat protein kinase C. *Nature Struct. Biol.* **1**, 383-387.

Hoovers, J. M. N., Mannens, M., John, R., Bliek, J., Van-Hexningen, V., Porteus, D. J., Leschot, N. J., Westerveld, A., Little, P. F. R., (1992) High resolution localization of 69 potential human zinc finger protein genes: A number are clustered. *Genomics*, **12**, 254-263.

Isalan, M., Choo, Y., Klug, A. (1997) Synergy between adjacent zinc fingers in sequence-specific DNA recognition. *Proc. Natl. Acad. Sci. USA*, **94**, 5617-5621.

Jacobs, G. H. (1992) Determination of the base recognition positions of zinc fingers from sequence analysis. *EMBO Journal*, **11**, 4507-4517.

Jamieson, A. C., Wang, H., Kim, S. (1996) A zinc finger directory for high-affinity DNA recognition. *Proc. Natl. Acad. Sci. USA,* **93**, 12834-12839.

Johnston, M., (1987) A model fungal regulatory mechanism: the GAL genes of *saccharomyces cerevisiae. Microbiol. Rev.* **51**, 458-476.

Judd, A. K., Sanchez, A., Bucher, D. J., Huffman, J. H., Bailey, K., Sidwell, R. W. (1997) In Vivo Anti-Influenza virus activity of a zinc finger peptide. *Antimicrobial Agents and Chemotherapy*, **41**, 687-692.

Karlsson, O., Thor, S., Norberg, T., Ohlsson, H., Edlund, T., (1990)Insulin gene enhancer-binding protein *Isc*-1 is a member of a novel class of proteins containing both a homeodomain and a His-Cys domain. *Nature*, **344**, 879-882.

Kel, O. V., Romaschenko, A. G., Kel, A. E., Wingender, E., Kolchanov, N. A. (1995) A compilation of composite regulatory elements affecting gene transcription in vertebrates. *Nucl. Acids. Res.* **23**, 4097-4103.

Kim, C. Y., Berg, J. M.. (1995) Serine at position 2 in the DNA recognition Helix of a $Cys_2His_2$ Zinc finger motif is not in general responsible for base recognition. *J. Mol. Biol.*, **252**, 1-5.

Kim, C. Y., Berg, J. M.. (1996) A 2.2Å resolution crystal structure of a designed zinc finger protein bound to DNA. *Nature Structural Biology.*, **3**, 940-945.

Kim, J., Kim, J., Cepek, K. L., Sharp, P. A., Pabo, C. O. (1997) Design of TATA box-binding protein/zinc finger fusion's for targeted regulation of gene expression. *Proc. Natl. Acad. Sci. USA*, **94**, 3616-3620.

Kim, J., Pabo, C. O. (1997) Transcriptional repression by zinc finger peptides. *The Journal of Biological Chemistry*, **272**, 29795-29800.

Kim, J., Pabo, C. O. (1998) Getting a handhold on DNA: Design of poly-zinc finger proteins with femtomolar dissociation constants. *Proc. Natl. Acad. Sci. USA*, **95**, 2812-2817.

Kim, Y., Cha, J., Chandrasegaran, S. (1996) Hybrid restriction enzymes: Zinc finger fusion's to *Fok* 1 cleavage domain. *Proc. Natl. Acad. Sci. USA*, **93**, 1156-1160.

Kinzler, K. W., Ruppert, J. M., Bigner, S. H., Vogelstein, B. (1988) The GL1 gene is a member of the *Kruppel* family of zinc finger proteins. *Nature*, **332**, 371-

Klevit, R. E., Herriott, J. R., Horvath, S. J. (1990) Solution structure of a single zinc finger domain of Yeast ADR1. *Proteins: structure function and genetics*, **7**, 215-226.

Klug, A. & Rhodes, D. (1987) Zinc Fingers: A novel protein motif for nucleic acid recognition. *Trends in Biochemical Sci*, **12**, 464-499.

Klug, A. & Schwabe, J. W. R. (1995) Zinc Fingers. *FASEB Journal*, **9**, 597-604.

Kraulis, P. J., Raine, A. R. C., Gadhavi, P. L., Laue, E. D. (1992) Structure of the zinc-containing domain of GAL4. *Nature* **356**, 448-450.

Krizek, B. A., Amann, B. T., Kilfoil, V. J., Merkle, D. L. & Berg, J. M. (1991) A consensus zinc finger peptide: Design, high-affinity metal binding, a pH-dependent structure, and a His to Cys sequence variant. *J. Am. Chem. Soc*, **113**, 4518-4523.

Kumar, V., Green, S., Stack, G., Berry, M., Jin, J., Chambon, P. (1987) Functional domains of the Estrogen receptor. *Cell*, **51**, 941-951.

Landschultz, W. H., Johnson, P. F., McKnight, S. L. (1988) The Leucine Zipper: A hypothetical structure common to a new class of DNA binding protein. *Science*, **240**, 1759-1764.

Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., Wright, P. E. (1989) Three dimentional structure of a single Zinc Finger DNA-Binding Domain. *Science*, **245**, 635-637.

Lusi, B. F., Xu. W. X., Otwinowski, Z., Freedman, L. P., Yamamoto, K. R., Sigler, P. B. (1993) Crytalogrhaphic analysis of the interaction of the Glutocorticoid receptor with DNA. *Nature*, **352**, 497-505.

Michael, S. P., Kiddoh, V. I., Schmide, M. H., Amann, B. T. & Berg, J. M. (1992) Metal binding properties of a minimalist Cys2His2 zinc finger peptide. *Proc. Natl. Acad. Sci. USA,* **89**, 4796-4800.

Miller, J. H., Reznikoff, W. S., eda. TheOperon. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1978.

Miller, J., McLachlan, A. D. & Klug, A. (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus*. *The EMBO Journal*, **4**, 1609-1614.

Nagaoka, M., Sugiura, Y. (1996) Distinct phosphate backbone effect revealed by some mutant peptides of Zinc finger protein Sp1: Effect of Protein-Induced bending on DNA recognition. *Biochemistry*, **35**, 8761-8768.

Nakaseko, Y., Neuhaus, D., Klug, A., Rhodes, D. (1992) Adjacent Zinc-finger motifs in multiple Zinc finger peptides from SWI5 form structurally independent, flexibly linked domains. *J. Mol. Biol.* **228,** 619-636.

Nardelli, J., Gibson, T. J., Vesque, C. & Charnay, P. (1991) Base sequence discrimination by zinc-finger DNA-binding domains. *Nature*, **349**, 175-178.

Nekludova, L., Pabo, C. O. DNA Conformations in protein-DNA complexes. (1994) *Proc. Natl. Acaf. Sci. USA*. **91**. 6948-6952.

Neuhaus, D., Nakaseko, Y., Swhabe J. W. R., Klug, A. (1992) Solution structure of two Zinc-finger Domains from SW15 Obtained using two dimensional $^{1}H$ Nuclear Magnetic Resonance Spectroscopy. *J. Mol. Biol.* **228,** 637-651.

Nichols, J. & Nimer, S. D. (1992) Transcription factors, translocations, and leukemia. *Blood*, **80**, 2953-2963.

Nolte, R. T., Conlin, R. M., Harrison, S. C., Brown, R. S. (1998) Differing roles for zinc fingers in DNA recognition: Structure of a six-finger transcription factor IIIA complex. *Proc. Natl. Acad. Sci. USA*, **95**, 2938-2943.

Ominchinski, J. G., Clore, G. M., Schaad, O., Felsenfeld, G., Trainor, C., Appella, E., Stahl, S. J., Gronenborn, A. M. (1993) NMR structure of a specific DNA binding complex of Zn-containing DNA binding domain of GAGA-1. *Science*, **261**, 438-446.

Ominchinski, J. G., Pedone, P. V., Felsenfeld, G., Gronenborn, A. M., Clore, G. M. (1997) The solution structure of a specific GAGA factor-DNA complex reveals a modular binding mode. *Nature strucural biology*, **4**, 122-132.

Ono, Y., Fuji, T., Igarashi, K., Kuno, T., Tanaka, C., Kikkawa, U., Nishizuka, Y. (1989) Phorbol ester binding to protein kinase C requires a cysteine-rich zinc finger-like sequence. *Proc. Natl. Acad. Sci. USA*, **89**, 4868-4871.

Pabo, C. O., Aggarwal, A. K., Jordan, S. R., Beamer, L. J., Odeysekare, U. R., Harrison, S. C. (1990) Conserved residues Make similar contacts in two repressor-Operator complexes. *Science*, **247**, 1210-1213.

Pabo, C. O., Sauer, R. T. (1992) Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem*, **61**, 1053-1095.

Pan, T., Coleman, J. E. (1989) GAL4 transcription factor is not a "zinc finger" but forms a $Zn(II)_2Cys_6$ binuclear cluster. *Proc. Natl. Acad. Sci. USA*, **87**, 2077-2081.

Pavletich, N. P. & Pabo, C. O. (1991) Zinc finger DNA recognition: Crystal structure of a ZIF268 DNA complex at 2.1Å. *Science*, **252**, 809-817.

Pavletich, N. P. & Pabo, C. O. (1993) Crystal structures of a five-finger GLI-DNA complex: New perspective on zinc fingers. *Science*, **261**, 1701-1707.

Petrez-Alvarado, G., Miles, C., Michelson, J. W., Louis, H. A., Winge, D. R., Beckerly, M. C., Summers, M. F., (1994) Structure of the C-terminal LIM domain from the cysteine rich protein CRP. *Nature Struct. Biol.* **1**, 388-398.

Pomerantz, J. L., Sharp, P. A., Pabo, C. O. (1995) Structure-based design of Transcription factors. *Science*, **267**, 93-96.

Prodromou, C., Pearl, L. H., (1992) Recursive PCR: a novel technique for total gene synthesis. *Protein Engineering,* **5**, 827-829.

Ptashne, M. N. (1986) A genetic Switch. Palo Alto, CA: Blacwell.

Quain, X., Gozani, S. N., Yonn, H., Jeon, C., Agarweal, K., Weiss, M. (1993) Novel zinc finger motif in the basal transcriptional machinery: Three dimensional NMR studies of the nucleic acid binding domain of the transcriptional elongation factor TFIIS. *Biochemistry* **32**, 9944-9959.

Raumann, B. E., Brown, B. M., Saur, R. T. (1994) Major groove DNA recognition by β-sheets; the ribbon-helix-helix family of gene regulatory proteins. *Curr Opin Strut Biol*, **4**, 36-43.

Rebar, E. J., Greisman, H. A. and Pabo, C. O. (1996) Phage display methods for selecting proteins with novel DNA binding specificity's. *Methods in Enzymology*, **267**, 129-149.

Rebar, E. J., Pabo, C. O. (1994) Zinc finger Phage: Affinity selection of fingers with New DNA-Binding Specificity's. *Science*, **263**, 671-673.

Rhodes, D. & Klug, A. (February 1993) Zinc Fingers - They play a key part in regulating the activity of genes in many species, from yeast to humans. Fewer than 10 years ago no one new they existed. *Scientific American*, 56-65.

Sadler, I., Crawford, A. W., Michelson, J. W., Beckerle, M. C., (1992) Biochemical evidence suggesting that the LIM domain may be involved in protein-protein interactions. *J. Cell Biol.*, **199**, 1573-1587.

Sauer, R., Yocum, R. R., Doolittle, R. F., Lewis, M., Pabo, C. O. (1992) Homology among DNA-binding proteins suggests use a conserved super secondary structure. *Nature*, **298**, 447-451.

Schmiedeskamp, M. & Klevit, R. E. (1994) Zinc finger diversity. *Current Opinion in Structural biology*, **4**, 28-35.

Schwabe, J. W. R. & Klug, A. (1994) Zinc mining for protein domains. *Structural Biology*, **1**, 345-349.

Schwabe, J. W. R. (1997) The role of water in protein-DNA interactions. *Current Opinion in Structural Biology*, 7, 125-134.

Schwabe, J. W. R., Chapman, L., Finch, J. T. & Rhodes, D. (1993) The crystal structure of the Estrogen receptor DNA binding domain bound to DNA: How receptors discriminate between their response elements. *Cell*, **75**, 567-579.

Seeman, N. C., Rosenberg, J. M., Rich, A. (1976) Sequence specific recognition of double helical nucleic acids by proteins. *Proc. Natl. Acad. Sci. USA.*, **73**, 804-808.

Shi, Y. & Berg, J. M. (1995a) A direct comparison of the properties of natural and designed zinc-finger proteins. *Chemistry & Biology*, **2**, 83-89.

Shi, Y. & Berg, J. M. (1995b) Specific DNA-RNA hybrid binding by Zinc finger proteins. *Science*, **268**, 282-284.

Stemmer, W. P. C. (1994a) DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution. *Pro. Natl. Acad. Sci. USA*, **91**, 10747-10751.

Stemmer, W. P. C. (1994b) Rapid evolution of a protein *In vitro* by DNA shuffling. *Nature*, **370**, 389-391.

Summers, M. F., South, T. L., Kim, B., Hare, D. R. (1990) High resolution structure or an HIV fingerlike domain via a new NMR-based distance geometry approach. *Biochemistry* **29**, 329-340.

Suzuki, M., Brenner, S. E., Gerstein, M., Yagi, N. (1995) DNA recognition code for transcription factors. *Protein Engineering*, **8**, 319-328.

Tabor, S. (1990) Expression using the T7 RNA polymerase/promoter system, In: F.A. Ausubel, R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith, K. Struhl, (eds.), *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York, pp.16.2.1-16.2.11

Takatsuji, H,. (1996) A single amino acid difference determine the specificity for the target sequence of two zinc finger proteins in plant. *Biochemical and Biophysical research communications* **224**, 219-223.

Van Dyke, M. W., Sirito, M. & Sawadogo, M. (1992) Single step purification of bacterialy expressed polypeptides containing an oligo-histidine domain. *Gene*, **111**, 99-104.

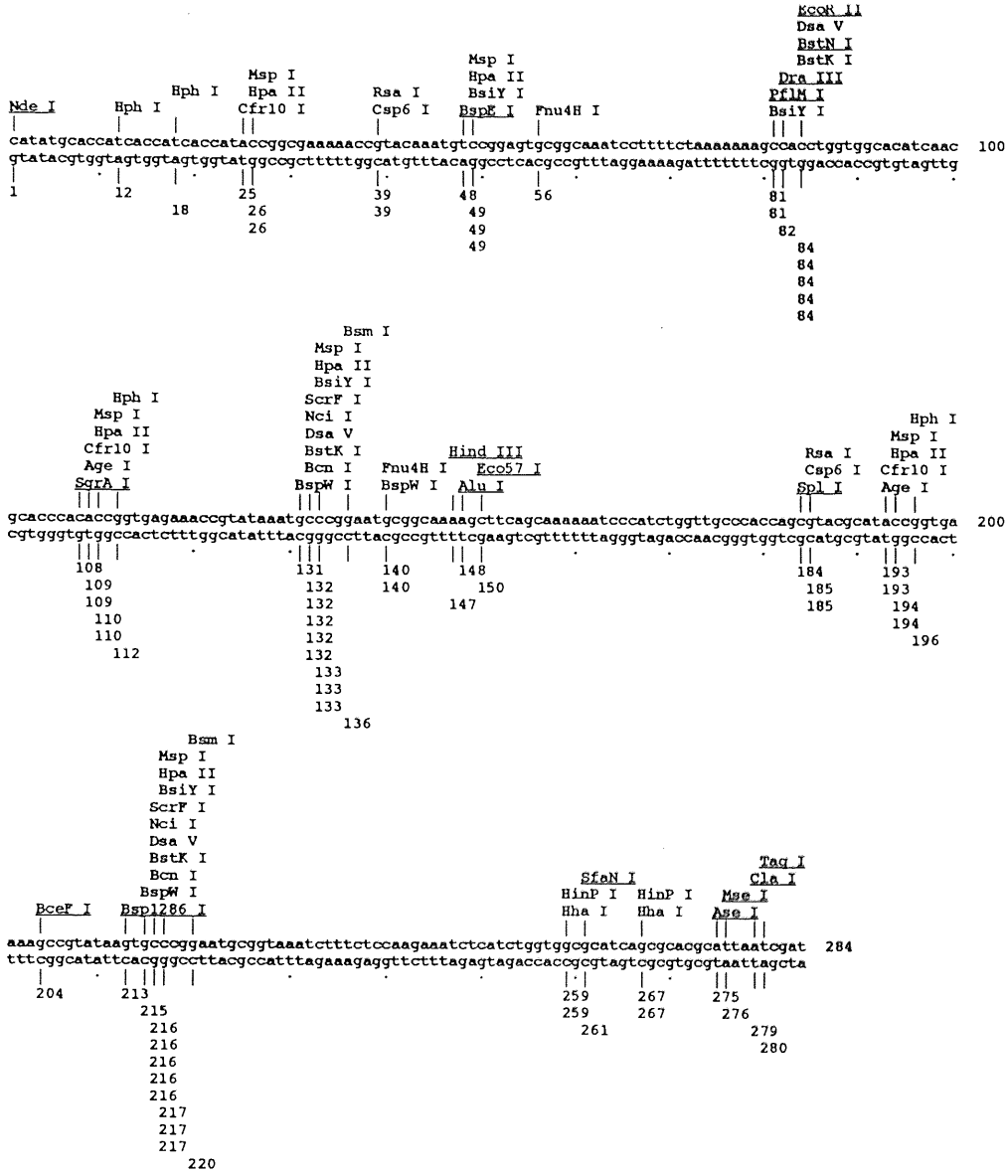
Wada, K., Wada. Y., Ishibashi, F., Gojobori. T., Ikemi, T., (1992) Codon usage tabulated from the GenBankgenetic sequence data. *Nucl. Acids. Res.* **20**, 2111-2118.

Yokono, M., Ssegusa, N., Matsushita, K., Sugiura, Y. (1998) Unique DNA Binding Mode of the N-Terminal Zinc Finger of Transcription Factor Sp1. *Biochemistry*, **37**, 6824-6832.

# APPENDICES.

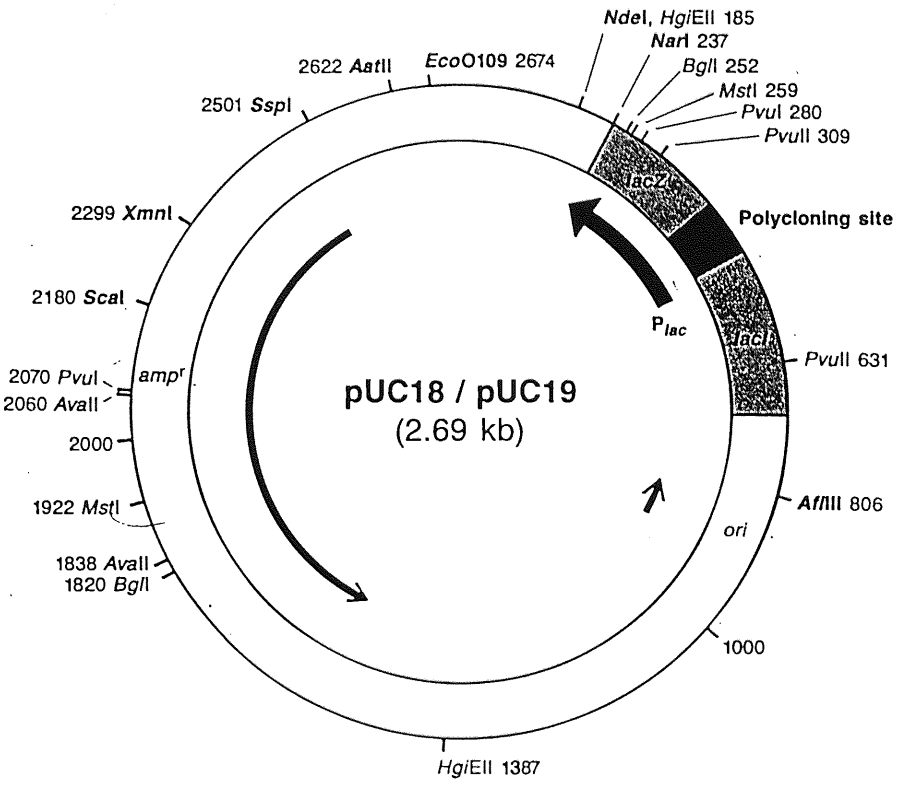
## APPENDIX I

Full restriction enzyme map of the zinc finger gene. Sequence analysis performed on DNA Strider™.



## APPENDIX II

Genetic map of pUC19 showing major restriction sites and polycloning site. Taken from Molecular cloning a Laboratory manual. Second Edition, Sambrook, Fritsch and Maniatis.




**Polycloning Sites**

pUC18

| | 1 | 2 | 3 | 4 | 5 | 6 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Thr | Met | Ile | Thr | Asn | Ser | Ser | Ser | Val | Pro | Gly | Asp | Pro | Leu | Glu | Ser | Thr | Cys | Arg | His | Ala | Ser | Leu | Ala | Leu | Ala |
| ATG | ACC | ATG | ATT | ACG | AAT | TCG | AGC | TCG | GTA | CCC | GGG | GAT | CCT | CTA | GAG | TCG | ACC | TGC | AGG | CAT | GCA | AGC | TTG | GCA | CTG | GCC |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | *EcoRI* | | *SacI* | | *KpnI* | *SmaI* | | *BamHI* | | *XbaI* | | *SalI* | | *PstI* | | *SphI* | | *HindIII* | | | | |
| | | | | | | | | | | *XmaI* | | | | | | *AccI* | | | | | | | | | | |
| | | | | | | | | | | | | | | | | *HincII* | | | | | | | | | | |

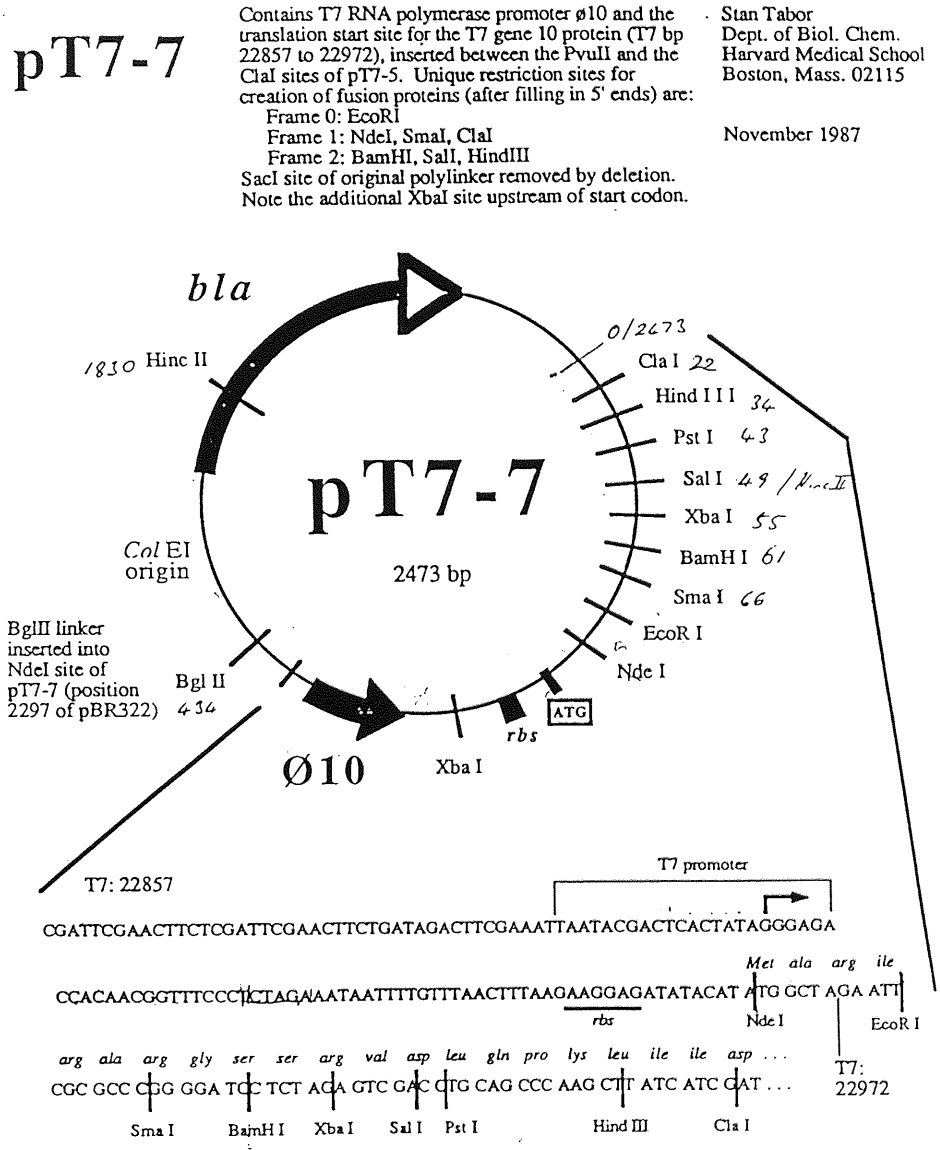
pUC19

| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Thr | Met | Ile | Thr | Pro | Ser | Leu | His | Ala | Cys | Arg | Ser | Thr | Leu | Glu | Asp | Pro | Arg | Val | Pro | Ser | Ser | Asn | Ser | Leu | Ala |
| ATG | ACC | ATG | ATT | ACG | CCA | AGC | TTG | CAT | GCC | TGC | AGG | TCG | ACT | CTA | GAG | GAT | CCC | CGG | GTA | CCG | AGC | TCG | AAT | TCA | CTG | GCC |
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | | | | *HindIII* | | *SphI* | | *PstI* | | | *SalI* | | *XbaI* | | *BamHI* | | *SmaI* | *KpnI* | | *SacI* | | *EcoRI* | | | |
| | | | | | | | | | | | | *AccI* | | | | | | *XmaI* | | | | | | | | |
| | | | | | | | | | | | | *HincII* | | | | | | | | | | | | | | |

In pUC18, the *Eco*RI site lies immediately downstream from $P_{lac}$.
In pUC19, the *Hind*III site lies immediately downstream from $P_{lac}$.

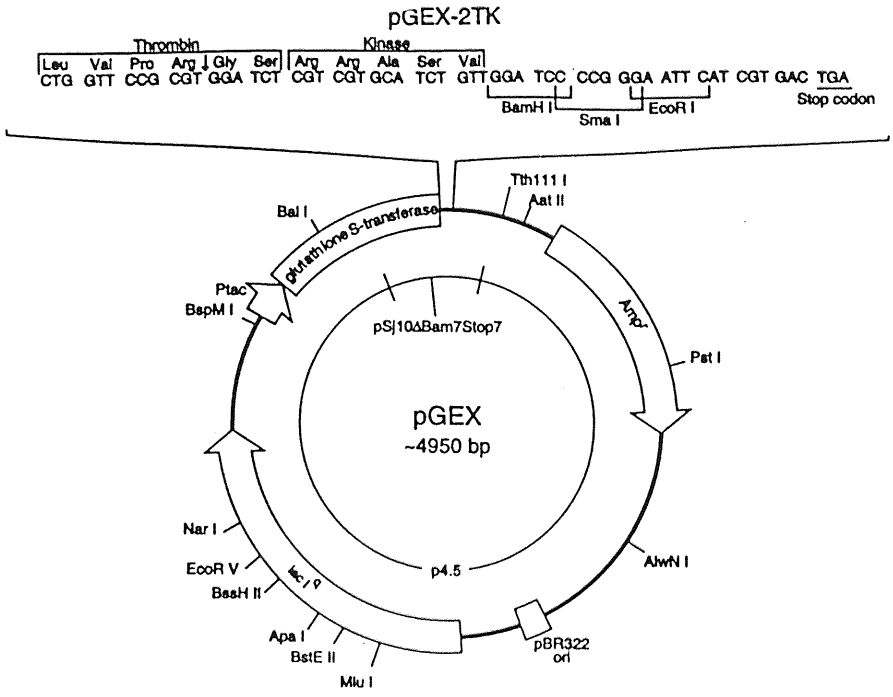
## APPENDIX III

Genetic map of pT7-7 showing major restriction sites and polycloning site (Tabor, 1990).



## APPENDIX IV

Genetic map of pGex-2TK showing major restriction sites and polycloning site. Taken from the Amersham Pharmacia Biotech catalogue 1998.



## APPENDIX V

Sequences of primers.

pUC19 Forward primer.

5' TTT TCC CAG TCA CGA C 3'

pUC19 Reverse primer.

5' CAG GAA ACA GCT ATG AC 3'

pT7-7 Forward primer.

5' AAG CTT GAG CAA ATC CCA TCT GGT TGG 3'

pT7-7 Reverse primer.

5' GGA TTT GCA CTC CGG ACA TTT GTA CG 3'

## APPENDIX VI

DNA annealing experiment using SPA (5.2.1, Table 5.1).

| Experiment | cpm |
|---|---|
| 32P only | 54.1 |
| | 64.1 |
| | 332.5 |
| | 48.3 |
| 32P TS1A | 645.5 |
| | 1013.9 |
| | 1274.3 |
| | 981.5 |
| 32P TS1B | 2895.7 |
| | 874.9 |
| | 90.3 |
| | 408.1 |
| Annealing experiment 1 | 2461.7 |
| | 2015.5 |
| | 2720.9 |
| | 2684.5 |
| Annealing experiment 2 | 4145.3 |
| | 3481.5 |
| | 2693.9 |
| | 3505.3 |

## APPENDIX VII

Optimisation of amount of SPA beads (mg) required/ pmol DNA in SPA method 1 (5.2.1, Fig. 5.3).

| Ratio of SPA beads to DNA | cpm | Ratio of SPA beads to DNA | cpm |
|---|---|---|---|
| 1pmol 33P/ 1mg SPA beads | | 7.0mg/pmol | 1212.0 |
| | 17.0 | | 1209.0 |
| | 519.0 | | 1107.0 |
| | 106.6 | | 626.8 |
| 0.1mg/pmol | 78.2 | 8.0mg/pmol | 843.6 |
| | 167.2 | | 1377.0 |
| | 196.2 | | 1310.0 |
| | 354.4 | | 1339.4 |
| 0.25mg/pmol | 306.6 | 9mg/pmol | 755.6 |
| | 690.8 | | 1215.8 |
| | 693.6 | | 1250.4 |
| | 799.0 | | 648.4 |
| 0.5mg/pmol | 415.8 | 10mg/pmol | 507.6 |
| | 623.2 | | 606.8 |
| | 526.8 | | 799.0 |
| | 458.4 | | 763.4 |
| 1mg/pmol | 539.8 | | |
| | 2607.6 | | |
| | 859.8 | | |
| | 876.0 | | |
| 1.5mg/pmol | 952.0 | | |
| | 1222.0 | | |
| | 1076.0 | | |
| | 1004.0 | | |
| 2.0mg/pmol | 1295.0 | | |
| | 1129.0 | | |
| | 1162.0 | | |
| | 952.6 | | |
| 3.0mg/pmol | 1013.2 | | |
| | 1368.8 | | |
| | 1032.2 | | |
| | 948.4 | | |
| 4.0mg/pmol | 1065.6 | | |
| | 1258.2 | | |
| | 1274.2 | | |
| | 977.8 | | |
| 5.0mg/pmol | 1076.8 | | |
| | 1294.4 | | |
| | 1244.4 | | |
| | 1244.6 | | |
| 6.0mg/pmol | 1018.4 | | |
| | 1275.6 | | |
| | 1379.0 | | |
| | 1286.6 | | |

## APPENDIX VIII

SPA method 1 pre-coupling DNA and beads (5.2.2, Table 5.2).
Total counts / assay = 1658

| Assay | cmp |
|---|---|
| No DNA (negative control) | 296.8 |
| | 314.5 |
| DNA (positive control) | 287.0 |
| | 251.0 |

SPA method 1 pre-coupling DNA and His-ZF (5.2.2, Table 5.3).
Total counts / assay = 1597.1

| Assay | cpm |
|---|---|
| No DNA (negative control) | 149.3 |
| | 119.8 |
| DNA (positive control) | 108.0 |
| | 104.1 |

SPA method 2, selection of assay format (5.3.1, Table 5.4).

| Assay format | cpm |
|---|---|
| Protein A coated yttrium silicate beads | 16316 |
| | 15321 |
| Protein A coated yttrium silicate beads + RGS His antibody | 7399 |
| | 7363 |
| Protein A coated PVT beads | 2529 |
| | 3227 |
| Protein A coated PVT beads + RGS His antibody | 3688 |
| | 2890 |
| Protein A coated PVT beads + penta His antibody | 3716 |
| | 5237 |
| Protein A coated PVT beads + tetra His antibody | 5813 |
| | 5886 |

## APPENDIX IX

Effect of pH on the binding of DNA to SPA beads (5.3.2, Figure 5.5).

| | Containing His-ZF | Negative control |
|---|---|---|
| pH6.8 | 99344 | 54915 |
| | 95805 | 50307 |
| | 95743 | 69755 |
| | 87987 | 70207 |
| pH7.0 | 90832 | 80287 |
| | 97051 | 85704 |
| | 99121 | 96982 |
| | 78806 | 99797 |
| pH7.2 | 87540 | 84050 |
| | 76466 | 92453 |
| | 82877 | 104536 |
| | 88088 | 107382 |
| pH7.4 | 85347 | 61803 |
| | 84528 | 64916 |
| | 66455 | 72512 |
| | 76023 | 75683 |
| pH7.6 | 80443 | 36566 |
| | 70516 | 32840 |
| | 80259 | 39073 |
| | 86924 | 34838 |
| pH7.8 | 66748 | 16203 |
| | 69455 | 19448 |
| | 70128 | 18018 |
| | 70144 | 21166 |
| pH8.0 | 63436 | 9764 |
| | 61626 | 10060 |
| | 63408 | 10023 |
| | 64494 | 10997 |

Statistical analysis (t-Test: Paired Two Sample for Means) of the effect of pH on the binding of DNA to SPA beads (5.3.2).

| pH6.8 | | | pH7.0 | | |
|---|---|---|---|---|---|
| | Negative control | With His-ZF | | Negative control | With His-ZF |
| Mean | 94719.75 | 61296 | Mean | 91452.5 | 90692.5 |
| Variance | 22979473 | 1.04E+08 | Variance | 83489359 | 85203098 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.63214 | | Pearson Correlation | -0.34929 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 4.862782 | | t Stat | 0.10075 | |
| P(T<=t) one-tail | 0.008305 | | P(T<=t) one-tail | 0.463052 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.016609 | | P(T<=t) two-tail | 0.926105 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| pH7.2 | | | pH7.4 | | |
|---|---|---|---|---|---|
| | Negative control | With His-ZF | | Negative control | With His-ZF |
| Mean | 83742.75 | 97105.25 | Mean | 78088.25 | 68728.5 |
| Variance | 29000306 | 1.18E+08 | Variance | 77919198 | 41725883 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.128413 | | Pearson Correlation | -0.77867 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | -2.32932 | | t Stat | 1.296579 | |
| P(T<=t) one-tail | 0.051105 | | P(T<=t) one-tail | 0.142749 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.102211 | | P(T<=t) two-tail | 0.285498 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| pH7.6 | | | pH7.8 | | |
|---|---|---|---|---|---|
| | Negative control | With His-ZF | | Negative control | With His-ZF |
| Mean | 79535.5 | 35829.25 | Mean | 69118.75 | 18708.75 |
| Variance | 45762774 | 6994302 | Variance | 2601081 | 4446829 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.422065 | | Pearson Correlation | 0.785192 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 14.245 | | t Stat | 77.16476 | |
| P(T<=t) one-tail | 0.000375 | | P(T<=t) one-tail | 2.4E-06 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.00075 | | P(T<=t) two-tail | 4.8E-06 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| pH8.0 | | |
|---|---|---|
| | Negative control | With His-ZF |
| Mean | 63241 | 10211 |
| Variance | 1414716 | 291916.7 |
| Observations | 4 | 4 |
| Pearson Correlation | 0.575839 | |
| Hypothesised Mean Difference | 0 | |
| df | 3 | |
| t Stat | 107.881 | |
| P(T<=t) one-tail | 8.78E-07 | |
| t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 1.76E-06 | |
| t Critical two-tail | 3.182449 | |

## APPENDIX X

Effect of KCl on the binding of DNA to SPA beads (5.3.2, Fig. 5.6).

| | His-ZF | Negative control |
|---|---|---|
| 0.0mM | 51392 | 12720 |
| | 49507 | 15375 |
| | 49988 | 13817 |
| | 50723 | 15288 |
| 2.5mM | 53144 | 15724 |
| | 56177 | 16814 |
| | 57166 | 17795 |
| | 56807 | 15837 |
| 25mM | 50004 | 49480 |
| | 50053 | 49801 |
| | 46742 | 52356 |
| | 51998 | 50493 |
| 250mM | 57090 | 93978 |
| | 56166 | 101883 |
| | 67723 | 92043 |
| | 72984 | 95361 |

Statistical analysis (t-Test: Paired Two Sample for Means) of the effect of KCl on the binding of DNA to SPA beads (5.3.2).

| 0.0mM KCl | | | 2.5mM KCl | | |
|---|---|---|---|---|---|
| | With His-ZF | Negative control | | With His-ZF | Negative control |
| Mean | 50402.5 | 14300 | Mean | 55823.5 | 16542.5 |
| Variance | 685187 | 1620486 | Variance | 3358087 | 936713.7 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.63559 | | Pearson Correlation | 0.61585 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 37.81896 | | t Stat | 54.08061 | |
| P(T<=t) one-tail | 2.03E-05 | | P(T<=t) one-tail | 6.96E-06 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 4.07E-05 | | P(T<=t) two-tail | 1.39E-05 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 25mM KCl | | | 250mMKCl | | |
|---|---|---|---|---|---|
| | With His-ZF | Negative control | | With His-ZF | Negative control |
| Mean | 49699.25 | 50532.5 | Mean | 63490.75 | 95816.25 |
| Variance | 4749197 | 1656520 | Variance | 67551766 | 18209762 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.72052 | | Pearson Correlation | -0.50336 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | -0.51558 | | t Stat | -5.87567 | |
| P(T<=t) one-tail | 0.320868 | | P(T<=t) one-tail | 0.004917 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.641736 | | P(T<=t) two-tail | 0.009835 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XI

Effect of $ZnCl_2$ on the binding of DNA to SPA beads (5.3.2, Fig. 5.7).

| [$ZnCl_2$] | His-ZF | Negative control |
|---|---|---|
| 0.0mM | 51392 | 12720 |
| | 49507 | 15375 |
| | 49988 | 13817 |
| | 50723 | 15288 |
| 2.5mM | 90921 | 89487 |
| | 88433 | 88444 |
| | 99212 | 84290 |
| | 77805 | 88936 |
| 25mM | 85259 | 70305 |
| | 89946 | 74877 |
| | 85482 | 70241 |
| | 81263 | 69878 |
| 250mM | 73776 | 50077 |
| | 76533 | 51494 |
| | 80448 | 50380 |
| | 75607 | 50180 |

Statistical analysis of the effect of $ZnCl_2$ on the binding of DNA to SPA beads (5.3.2).

| 0.0mM $ZnCl_2$ | | | 2.5mM $ZnCl_2$ | | |
|---|---|---|---|---|---|
| | Variable 1 | Variable 2 | | Variable 1 | Variable 2 |
| Mean | 50402.5 | 14300 | Mean | 89092.75 | 87789.25 |
| Variance | 685187 | 1620486 | Variance | 77863430 | 5623613 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.63559 | | Pearson Correlation | -0.7277 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 37.81896 | | t Stat | 0.24423 | |
| P(T<=t) one-tail | 2.03E-05 | | P(T<=t) one-tail | 0.411402 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 4.07E-05 | | P(T<=t) two-tail | 0.822803 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 25mM $ZnCl_2$ | | | 250mM $ZnCl_2$ | | |
|---|---|---|---|---|---|
| | Variable 1 | Variable 2 | | Variable 1 | Variable 2 |
| Mean | 85487.5 | 71325.25 | Mean | 76591 | 50532.75 |
| Variance | 12592288 | 5641990 | Variance | 7924098 | 426491.6 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.877481 | | Pearson Correlation | 0.178626 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 15.2672 | | t Stat | 18.78901 | |
| P(T<=t) one-tail | 0.000305 | | P(T<=t) one-tail | 0.000165 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.00061 | | P(T<=t) two-tail | 0.000329 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XII

Effect of the removal of the His-ZF from the buffer (5.4.2, Table 5.5).

| Assay | cpm |
|---|---|
| Negative control | 15174 |
| | 16459 |
| His-ZF buffer | 56457 |
| | 55796 |
| His-ZF | 52267 |
| | 56375 |

Statistical analysis (t-Test: Paired Two Sample for Means) of the effect of removing the His-ZF from the buffer (5.4.2).

| | With His-ZF | Buffer only |
|---|---|---|
| Mean | 56126.5 | 54321 |
| Variance | 218460.5 | 8437832 |
| Observations | 2 | 2 |
| Pearson Correlation | -1 | |
| Hypothesised Mean Difference | 0 | |
| df | 1 | |
| t Stat | 0.757182 | |
| P(T<=t) one-tail | 0.293709 | |
| t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.587418 | |
| t Critical two-tail | 12.70615 | |

## APPENDIX XIII

Effect of imidazole on the binding of DNA to SPA beads (5.4.3, Fig. 5.8).

| | His-ZF | His-ZF buffer | Negative control |
|---|---|---|---|
| 0.0mM | 47261 | 42836 | 11076 |
| | 50295 | 42735 | 12306 |
| | 45474 | 40678 | 12923 |
| | 47909 | 42968 | 11916 |
| 2.5mM | 57868 | 42225 | 11473 |
| | 46579 | 39277 | 10380 |
| | 42460 | 47697 | 10086 |
| | 48068 | 41109 | 10002 |
| 25mM | 35542 | 43401 | 6339 |
| | 35474 | 43139 | 5942 |
| | 52589 | 40453 | 6947 |
| | 49127 | 44088 | 6478 |
| 250mM | 15501 | 17289 | 3013 |
| | 15500 | 15140 | 5936 |
| | 12121 | 16573 | 2896 |
| | 15242 | 16149 | 3150 |

Statistical analysis of the effect of imidazole on the binding of DNA to SPA beads Anova: Two-Factor With Replication (5.4.3).

| SUMMARY | His-ZF | Buffer | Negative control | Total |
|---|---|---|---|---|
| 0.0mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 190939 | 169217 | 48221 | 408377 |
| Average | 47734.75 | 42304.25 | 12055.25 | 34031.42 |
| Variance | 3973558 | 1184519 | 598062.3 | 2.7E+08 |
| | | | | |
| 2.5mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 194975 | 170308 | 41941 | 407224 |
| Average | 48743.75 | 42577 | 10485.25 | 33935.33 |
| Variance | 42626744 | 13127776 | 459886.3 | 3.22E+08 |
| | | | | |
| 25mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 172732 | 171081 | 25706 | 369519 |
| Average | 43183 | 42770.25 | 6426.5 | 30793.25 |
| Variance | 80539178 | 2546645 | 171989.7 | 3.47E+08 |
| | | | | |
| 250mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 58364 | 65151 | 14995 | 138510 |
| Average | 14591 | 16287.75 | 3748.75 | 11542.5 |
| Variance | 2726361 | 806816.9 | 2137025 | 35202160 |
| | | | | |
| Total | | | | |
| Count | 16 | 16 | 16 | |
| Sum | 617010 | 575757 | 130863 | |
| Average | 38563.13 | 35984.81 | 8178.938 | |
| Variance | 2.35E+08 | 1.42E+08 | 12151386 | |

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Sample | 4.19E+09 | 3 | 1.4E+09 | 111.1859 | 2.9E-18 | 2.866265 |
| Columns | 9.08E+09 | 2 | 4.54E+09 | 361.1461 | 1.5E-24 | 3.259444 |
| Interaction | 1.18E+09 | 6 | 1.97E+08 | 15.67193 | 9.45E-09 | 2.363748 |
| Within | 4.53E+08 | 36 | 12574880 | | | |
| | | | | | | |
| Total | 1.49E+10 | 47 | | | | |

t-Test: Paired Two Sample for Means (5.4.3).

| 0.0mM imidazole | | | 2.5mM imidazole | | |
|---|---|---|---|---|---|
| | His-ZF buffer | His-ZF | | His-ZF buffer | His-ZF |
| Mean | 47734.75 | 42304.25 | Mean | 48743.75 | 42577 |
| Variance | 3973558 | 1184519 | Variance | 42626744 | 13127776 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.713391 | | Pearson Correlation | -0.38396 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 7.562324 | | t Stat | 1.434512 | |
| P(T<=t) one-tail | 0.002398 | | P(T<=t) one-tail | 0.123453 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.004795 | | P(T<=t) two-tail | 0.246905 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 25mM imidazole | | | 250mM imidazole | | |
|---|---|---|---|---|---|
| | His-ZF buffer | His-ZF | | His-ZF buffer | His-ZF |
| Mean | 43183 | 42770.25 | Mean | 14591 | 16287.75 |
| Variance | 80539178 | 2546645 | Variance | 2726361 | 806816.9 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.50334 | | Pearson Correlation | -0.20836 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 0.0836 | | t Stat | -1.66556 | |
| P(T<=t) one-tail | 0.46932 | | P(T<=t) one-tail | 0.097195 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.93864 | | P(T<=t) two-tail | 0.194391 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XIV

Effect of $NiCl_2$ on the binding of DNA to SPA beads (5.4.3, Fig. 5.9).

| [$NiCl_2$] | His-ZF | Buffer | Negative control |
|---|---|---|---|
| 0.0mM | 65529 | 36440 | 12423 |
| | 45454 | 42051 | 12120 |
| | 36379 | 47683 | 13147 |
| | 48269 | 41996 | 12164 |
| 0.01mM | 54385 | 53107 | 12168 |
| | 57487 | 58445 | 15129 |
| | 55320 | 49552 | 13051 |
| | 56706 | 37787 | 12781 |
| 0.1mM | 64377 | 59189 | 19768 |
| | 78005 | 62434 | 21043 |
| | 71191 | 64292 | 18015 |
| | 70987 | 64677 | 20288 |
| 1mM | 37969 | 37418 | 55768 |
| | 35805 | 45722 | 55201 |
| | 36001 | 46022 | 55038 |
| | 33842 | 40576 | 47008 |
| 10mM | 17682 | 15111 | 11136 |
| | 17366 | 15067 | 29223 |
| | 17508 | 14322 | 20779 |
| | 17251 | 12597 | 18224 |

Statistical analysis of the effect of $NiCl_2$ on the binding of DNA to SPA beads.
Anova: Two-Factor With Replication (5.4.3).

| SUMMARY | His-ZF | Buffer | Negative control | Total |
|---|---|---|---|---|
| 0.0mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 195631 | 168170 | 49854 | 413655 |
| Average | 48907.75 | 42042.5 | 12463.5 | 34471.25 |
| Variance | 1.49E+08 | 21068494 | 225501.7 | 3.19E+08 |
| | | | | |
| 0.01mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 223898 | 198891 | 53129 | 475918 |
| Average | 55974.5 | 49722.75 | 13282.25 | 39659.83 |
| Variance | 1925876 | 76674026 | 1652256 | 4.09E+08 |
| | | | | |
| 0.1mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 284560 | 250592 | 79114 | 614266 |
| Average | 71140 | 62648 | 19778.5 | 51188.83 |
| Variance | 30964135 | 6276685 | 1656198 | 5.62E+08 |
| | | | | |
| 1mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 143617 | 169738 | 213015 | 526370 |
| Average | 35904.25 | 42434.5 | 53253.75 | 43864.17 |
| Variance | 2845093 | 17432369 | 17435392 | 66128413 |

| 10mM | | | | |
|---|---|---|---|---|
| Count | 4 | 4 | 4 | 12 |
| Sum | 69807 | 57097 | 79362 | 206266 |
| Average | 17451.75 | 14274.25 | 19840.5 | 17188.83 |
| Variance | 34610.92 | 1381350 | 55764494 | 21265672 |
| | | | | |
| Total | | | | |
| Count | 20 | 20 | 20 | |
| Sum | 917513 | 844488 | 474474 | |
| Average | 45875.65 | 42224.4 | 23723.7 | |
| Variance | 3.78E+08 | 2.84E+08 | 2.52E+08 | |

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Sample | 7.85E+09 | 4 | 1.96E+09 | 76.67018 | 1.67E-19 | 2.578737 |
| Columns | 5.64E+09 | 2 | 2.82E+09 | 110.2375 | 4.54E-18 | 3.20432 |
| Interaction | 8.35E+09 | 8 | 1.04E+09 | 40.79396 | 4.03E-18 | 2.152134 |
| Within | 1.15E+09 | 45 | 25590697 | | | |
| | | | | | | |
| Total | 2.3E+10 | 59 | | | | |

t-Test: Paired Two Sample for Means (5.4.3).

| 0.0mM $NiCl_2$ | | | 0.01mM $NiCl_2$ | | |
|---|---|---|---|---|---|
| | His-ZF | His-ZF buffer | | His-ZF | His-ZF buffer |
| Mean | 48907.75 | 42042.5 | Mean | 55974.5 | 49722.75 |
| Variance | 1.49E+08 | 21068494 | Variance | 1925876 | 76674026 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.97599 | | Pearson Correlation | -0.02211 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 0.822341 | | t Stat | 1.405533 | |
| P(T<=t) one-tail | 0.235577 | | P(T<=t) one-tail | 0.127262 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.471153 | | P(T<=t) two-tail | 0.254524 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 0.1mM $NiCl_2$ | | | 1.0mM $NiCl_2$ | | |
|---|---|---|---|---|---|
| | His-ZF | His-ZF buffer | | His-ZF | His-ZF buffer |
| Mean | 71140 | 62648 | Mean | 35904.25 | 42434.5 |
| Variance | 30964135 | 6276685 | Variance | 2845093 | 17432369 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.518793 | | Pearson Correlation | -0.30786 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 3.558796 | | t Stat | -2.63252 | |
| P(T<=t) one-tail | 0.018926 | | P(T<=t) one-tail | 0.039078 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.037853 | | P(T<=t) two-tail | 0.078156 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 10mM $NiCl_2$ | | |
|---|---|---|
| | His-ZF | His-ZF buffer |
| Mean | 17451.75 | 14274.25 |
| Variance | 34610.92 | 1381350 |
| Observations | 4 | 4 |
| Pearson Correlation | 0.707474 | |
| Hypothesised Mean Difference | 0 | |
| df | 3 | |
| t Stat | 6.041223 | |
| P(T<=t) one-tail | 0.004548 | |
| t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.009096 | |
| t Critical two-tail | 3.182449 | |

## APPENDIX XV

Effect of EDTA on the binding of DNA to SPA beads (5.4.3, Fig. 5.10).

| [EDTA] | His-ZF | Buffer | Negative control |
|---|---|---|---|
| 0.0mM | 65529 | 39440 | 12423 |
| | 45454 | 42051 | 12120 |
| | 36379 | 47683 | 13147 |
| | 48269 | 41996 | 12164 |
| 0.01mM | 51172 | 37608 | 9767 |
| | 24967 | 37992 | 10122 |
| | 50439 | 37456 | 10075 |
| | 28818 | 41046 | 10385 |
| 0.1mM | 40157 | 41055 | 3708 |
| | 53076 | 46376 | 3882 |
| | 52217 | 50625 | 3376 |
| | 50336 | 50203 | 3474 |
| 1mM | 8166 | 12593 | 2074 |
| | 9110 | 12073 | 2076 |
| | 10281 | 11718 | 2094 |
| | 8736 | 11219 | 2008 |
| 10mM | 1712 | 1774 | 1413 |
| | 1595 | 1867 | 1315 |
| | 1765 | 1871 | 1287 |
| | 1881 | 1757 | 1400 |

Statistical analysis of the effect of EDTA on the binding of DNA to SPA beads.

Anova: Two-Factor With Replication (5.4.3).

| SUMMARY | His-ZF | Buffer | Negative control | Total |
|---|---|---|---|---|
| 0.0mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 195631 | 171170 | 49854 | 416655 |
| Average | 48907.75 | 42792.5 | 12463.5 | 34721.25 |
| Variance | 1.49E+08 | 12113494 | 225501.7 | 3.21E+08 |
| | | | | |
| 0.01mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 155396 | 154102 | 40349 | 349847 |
| Average | 38849 | 38525.5 | 10087.25 | 29153.92 |
| Variance | 1.93E+08 | 2874393 | 64190.92 | 2.52E+08 |
| | | | | |
| 0.1mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 195786 | 188259 | 14440 | 398485 |
| Average | 48946.5 | 47064.75 | 3610 | 33207.08 |
| Variance | 35644987 | 19705155 | 52280 | 4.94E+08 |
| | | | | |
| 1mM | | | | |
| Count | 4 | 4 | 4 | 12 |
| Sum | 36293 | 47603 | 8252 | 92148 |
| Average | 9073.25 | 11900.75 | 2063 | 7679 |
| Variance | 798950.3 | 335686.9 | 1425.333 | 18966767 |

| 10mM | | | | |
|---|---|---|---|---|
| Count | 4 | 4 | 4 | 12 |
| Sum | 6953 | 7269 | 5415 | 19637 |
| Average | 1738.25 | 1817.25 | 1353.75 | 1636.417 |
| Variance | 14100.92 | 3621.583 | 3868.917 | 50605.36 |
| | | | | |
| Total | | | | |
| Count | 20 | 20 | 20 | |
| Sum | 590059 | 568403 | 118310 | |
| Average | 29502.95 | 28420.15 | 5915.5 | |
| Variance | 4.87E+08 | 3.5E+08 | 21369874 | |

| ANOVA | | | | | | |
|---|---|---|---|---|---|---|
| Source of Variation | SS | df | MS | F | P-value | F crit |
| Sample | 1.15E+10 | 4 | 2.87E+09 | 104.0062 | 3.91E-22 | 2.578737 |
| Columns | 7.09E+09 | 2 | 3.55E+09 | 128.6474 | 2.44E-19 | 3.20432 |
| Interaction | 3.6E+09 | 8 | 4.51E+08 | 16.34119 | 5.14E-11 | 2.152134 |
| Within | 1.24E+09 | 45 | 27568893 | | | |
| | | | | | | |
| Total | 2.34E+10 | 59 | | | | |

*t*-test: Paired Two Sample for Means (5.4.3)

| 0.0mM EDTA | | | 0.01mM EDTA | | |
|---|---|---|---|---|---|
| | His-ZF | His-ZF buffer | | His-ZF | His-ZF buffer |
| Mean | 48907.75 | 42792.5 | Mean | 38849 | 38525.5 |
| Variance | 1.49E+08 | 12113494 | Variance | 1.93E+08 | 2874393 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.89529 | | Pearson Correlation | -0.58818 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 0.795149 | | t Stat | 0.043252 | |
| P(T<=t) one-tail | 0.242313 | | P(T<=t) one-tail | 0.484109 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.484626 | | P(T<=t) two-tail | 0.968218 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 0.1mM EDTA | | | 1.0mM EDTA | | |
|---|---|---|---|---|---|
| | His-ZF | His-ZF buffer | | His-ZF | His-ZF buffer |
| Mean | 48946.5 | 47064.75 | Mean | 9073.25 | 11900.75 |
| Variance | 35644987 | 19705155 | Variance | 798950.3 | 335686.9 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.829891 | | Pearson Correlation | -0.39424 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 1.116537 | | t Stat | -4.55253 | |
| P(T<=t) one-tail | 0.172782 | | P(T<=t) one-tail | 0.00993 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.345563 | | P(T<=t) two-tail | 0.01986 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 10.0mM EDTA | | |
|---|---|---|
| | His-ZF | His-ZF buffer |
| Mean | 1738.25 | 1817.25 |
| Variance | 14100.92 | 3621.583 |
| Observations | 4 | 4 |
| Pearson Correlation | -0.61358 | |
| Hypothesised Mean Difference | 0 | |
| df | 3 | |
| t Stat | -0.97073 | |
| P(T<=t) one-tail | 0.201641 | |
| t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.403282 | |
| t Critical two-tail | 3.182449 | |

## APPENDIX XVI

Effect of de-salting the His-ZF and His-ZF buffer used in the assay (5.4.4, Table 5.6).

| Assay | cpm |
|---|---|
| Negative control | 12914 |
| | 7440 |
| | 8116 |
| | 7699 |
| His-ZF buffer (de-salted) | 9481 |
| | 10242 |
| | 9542 |
| | 10236 |
| His-ZF (de-salted) | 9439 |
| | 13134 |
| | 12387 |
| | 12440 |

t-Test: Paired Two Sample for Means (5.4.4).

| | Negative control | His-ZF buffer | | His-ZF | His-ZF buffer |
|---|---|---|---|---|---|
| Mean | 9042.25 | 9875.25 | Mean | 11850 | 9875.25 |
| Variance | 6739971 | 177044.9 | Variance | 2699349 | 177044.9 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.69877 | | Pearson Correlation | 0.701729 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | -0.57334 | | t Stat | 2.860625 | |
| P(T<=t) one-tail | 0.303287 | | P(T<=t) one-tail | 0.032272 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.606575 | | P(T<=t) two-tail | 0.064544 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XVII

Effect of various concentrations of de-salted His-ZF on the binding of DNA to SPA beads (5.4.4, Fig. 5.12).

| His-ZF concentration | Negative control | Positive |
|---|---|---|
| 0pmol | 9376 | 9359 |
| | 9330 | 9552 |
| | 9410 | 9155 |
| | 9234 | 11868 |
| 1pmol | 9460 | 12383 |
| | 9079 | 10766 |
| | 9005 | 12622 |
| | 9814 | 10997 |
| 2pmol | 8265 | 13116 |
| | 9878 | 13563 |
| | 9832 | 12499 |
| | 9823 | 14841 |
| 4pmol | 8021 | 12065 |
| | 11404 | 15544 |
| | 11790 | 12901 |
| | 10074 | 13016 |

Statistical analysis of the effect of various concentrations of de-salted His-ZF on the binding of DNA to SPA beads (5.4.4).

| 0pmol | | | 1pmol | | |
|---|---|---|---|---|---|
| | Negative | Positive | | Negative | Positive |
| Mean | 9337.5 | 9983.5 | Mean | 9339.5 | 11692 |
| Variance | 5835.667 | 1604648 | Variance | 139807 | 894294 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.95044 | | Pearson Correlation | -0.29825 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | -0.96449 | | t Stat | -4.21669 | |
| P(T<=t) one-tail | 0.202973 | | P(T<=t) one-tail | 0.012187 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.405946 | | P(T<=t) two-tail | 0.024375 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 2pmol | | | 4pmol | | |
|---|---|---|---|---|---|
| | Negative | Positive | | Negative | Positive |
| Mean | 9449.5 | 13504.75 | Mean | 10322.25 | 13381.5 |
| Variance | 624153.7 | 983872.3 | Variance | 2893951 | 2258016 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.255145 | | Pearson Correlation | 0.619957 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | -7.37885 | | t Stat | -4.3456 | |
| P(T<=t) one-tail | 0.002573 | | P(T<=t) one-tail | 0.011249 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.005146 | | P(T<=t) two-tail | 0.022499 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XVIII

Effect of the variation of protein concentration on the binding of DNA to SPA beads (7.2.1, Table 7.1).

| [Protein] | GST-ZF | GST |
|---|---|---|
| 8.2nM | 31774 | 6816 |
| | 31857 | 6225 |
| | 43109 | 6226 |
| | 30737 | 12280 |
| 16.4nM | 32618 | 5562 |
| | 46991 | 3405 |
| | 37351 | 5142 |
| | 39344 | 4395 |
| 32.8nM | 53577 | 4870 |
| | 44694 | 4829 |
| | 48756 | 3142 |
| | 51099 | 4872 |
| 65.3nM | 45771 | 5458 |
| | 53100 | 3809 |
| | 53891 | 4434 |
| | 39648 | 5749 |
| 98.4nM | 38001 | 15771 |
| | 17614 | 11293 |
| | 31716 | 12096 |
| | 47516 | 13770 |

Statistical analysis of the effect on DNA binding to the SPA beads due to the addition of the GST zinc finger protein, compared with the effect of the GST protein alone. t-Test: Paired Two Sample for Means (7.2.1, Table 7.2).

| 8.2nM protein | | | 16.4nM protein | | |
|---|---|---|---|---|---|
| | GST-ZF | GST | | GST-ZF | GST |
| Mean | 34369.25 | 7886.75 | Mean | 39076 | 4626 |
| Variance | 34207731 | 8655552 | Variance | 35800146 | 895518 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.45559 | | Pearson Correlation | -0.98084 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 6.922366 | | t Stat | 9.965336 | |
| P(T<=t) one-tail | 0.00309 | | P(T<=t) one-tail | 0.001075 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.00618 | | P(T<=t) two-tail | 0.00215 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 32.8nM protein | | | 65.6nM protein | | |
|---|---|---|---|---|---|
| | GST-ZF | GST | | GST-ZF | GST |
| Mean | 49531.5 | 4428.25 | Mean | 48102.5 | 4862.5 |
| Variance | 14275311 | 735698.9 | Variance | 45132067 | 811325.7 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.158559 | | Pearson Correlation | -0.916 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 24.12308 | | t Stat | 11.45163 | |
| P(T<=t) one-tail | 7.81E-05 | | P(T<=t) one-tail | 0.000715 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.000156 | | P(T<=t) two-tail | 0.001429 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 98.4nM protein | | |
|---|---|---|
| | GST-ZF | GST |
| Mean | 33711.75 | 13232.5 |
| Variance | 1.57E+08 | 3928727 |
| Observations | 4 | 4 |
| Pearson Correlation | 0.694415 | |
| Hypothesised Mean Difference | 0 | |
| df | 3 | |
| t Stat | 3.637985 | |
| P(T<=t) one-tail | 0.017896 | |
| t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.035792 | |
| t Critical two-tail | 3.182449 | |

## APPENDIX XIX

Effect of the addition of DTT on the binding of DNA to SPA beads (7.2.2, Table 7.2).

| [DTT] | GST-ZF | GST |
|---|---|---|
| 0.0mM | 45164 | 5756 |
| | 43395 | 4970 |
| | 46683 | 6538 |
| | 40273 | 6621 |
| 0.5mM | 38469 | 3955 |
| | 36632 | 5493 |
| | 36929 | 6347 |
| | 38359 | 6846 |
| 5.0mM | 34281 | 4008 |
| | 31788 | 5496 |
| | 47496 | 5853 |
| | 35708 | 5869 |

Statistical analysis of the effect of DTT on the binding of DNA to SPA beads.
t-Test: Paired Two Sample for Means (7.2.2, Table 7.2).

| Effect of 0.5mM DTT on GST-ZF assay | | | Effect of 0.5mM DTT on GST assay | | |
|---|---|---|---|---|---|
| | No DTT | O.5mM DTT | | No DTT | O.5mM DTT |
| Mean | 43878.75 | 37597.25 | Mean | 5971.25 | 5660.25 |
| Variance | 7583711 | 906158.9 | Variance | 597404.9 | 1604493 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.3857 | | Pearson Correlation | 0.57681 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 3.874804 | | t Stat | 0.600623 | |
| P(T<=t) one-tail | 0.015215 | | P(T<=t) one-tail | 0.295218 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.030429 | | P(T<=t) two-tail | 0.590436 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| Effect of 5.0mM DTT on GST-ZF assay | | | Effect of 5.0mM DTT on GST assay | | |
|---|---|---|---|---|---|
| | No DTT | 5.0mM DTT | | No DTT | 5.0mM DTT |
| Mean | 43878.75 | 37318.25 | Mean | 5971.25 | 5306.5 |
| Variance | 7583711 | 48662684 | Variance | 597404.9 | 779027 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.574664 | | Pearson Correlation | 0.373781 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 2.244719 | | t Stat | 1.428294 | |
| P(T<=t) one-tail | 0.055239 | | P(T<=t) one-tail | 0.124259 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.110477 | | P(T<=t) two-tail | 0.248519 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XX

Effect of BSA on the binding of DNA to SPA beads. Total cpm / assay = 127307 (7.2.2, Table 7.3).

| | cpm | |
|---|---|---|
| BSA concentration | GST | GST-ZF |
| 0.0mg/ml | 6843 | 35464 |
| | 14884 | 71783 |
| | 3515 | 30295 |
| | 5845 | 75677 |
| 5.0mg/ml | 2877 | 43515 |
| | 3084 | 45198 |
| | 3229 | 34740 |
| | 3229 | 43871 |

Statistical analysis of the effect of BSA on the binding of DNA to SPA beads, t-Test: Paired Two Sample for Means (7.2.2, Table 7.3).

| Effect of BSA on GST assay | | | Effect of BSA on GST-ZF assay | | |
|---|---|---|---|---|---|
| BSA conc. | 0.0mg/ml | 5.0mg/ml | BSA conc. | 0.0mg/ml | 5.0mg/ml |
| Mean | 7771.75 | 3104.75 | Mean | 53304.75 | 41831 |
| Variance | 24426321 | 27725.58 | Variance | 5.63E+08 | 22872142 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.2853 | | Pearson Correlation | 0.707703 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 1.869658 | | t Stat | 1.112519 | |
| P(T<=t) one-tail | 0.079154 | | P(T<=t) one-tail | 0.17352 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.158307 | | P(T<=t) two-tail | 0.347041 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XXI

Effect of $ZnCl_2$ on the binding of DNA to SPA beads Total CPM/ assay = 112862 (7.2.2, Table 7.4, Fig. 7.2).

| [$ZnCl_2$] | GST-ZF | GST |
|---|---|---|
| 0.0mM | 43270 | 5400 |
| | 35620 | 5512 |
| | 31200 | 4293 |
| | 46450 | 2970 |
| 0.1mM | 30371 | 13218 |
| | 34320 | 12098 |
| | 33984 | 13887 |
| | 28991 | 13971 |
| 1.0mM | 105104 | 55701 |
| | 84982 | 57298 |
| | 111074 | 57566 |
| | 112097 | 55120 |
| 10.0mM | 98231 | 94675 |
| | 132239 | 105717 |
| | 136057 | 104079 |
| | 135200 | 103690 |

Statistical analysis of the effect of $ZnCl_2$ on the binding of DNA to SPA beads, t-Test: Paired Two Sample for Means (7.2.2, Table 7.4).

| 0.0mM $ZnCl_2$ | | | 0.1mM $ZnCl_2$ | | |
|---|---|---|---|---|---|
| | GST-ZF | GST | | GST-ZF | GST |
| Mean | 39135 | 4543.75 | Mean | 31916.5 | 13293.5 |
| Variance | 48642300 | 1403412 | Variance | 6999496 | 748723 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.37863 | | Pearson Correlation | -0.51133 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 9.220037 | | t Stat | 11.72596 | |
| P(T<=t) one-tail | 0.001349 | | P(T<=t) one-tail | 0.000666 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.002699 | | P(T<=t) two-tail | 0.001333 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 1.0mM $ZnCl_2$ | | | 10.0mM $ZnCl_2$ | | |
|---|---|---|---|---|---|
| | GST-ZF | GST | | GST-ZF | GST |
| Mean | 103314.3 | 56421.25 | Mean | 125431.8 | 102040.3 |
| Variance | 1.59E+08 | 1430385 | Variance | 3.32E+08 | 24881192 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.44019 | | Pearson Correlation | 0.965809 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 7.118556 | | t Stat | 3.477721 | |
| P(T<=t) one-tail | 0.002853 | | P(T<=t) one-tail | 0.020061 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.005705 | | P(T<=t) two-tail | 0.040121 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XXII

Effect of order of addition and volume on the binding of DNA to SPA beads. Total cpm / assay = 242691.7 (7.2.3, Table 7.5).

| | Order of addition A | | Order of addition B | |
|---|---|---|---|---|
| | 50µl | 100µl | 50µl | 100µl |
| DNA | | | 5502 | 22084.4 |
| | | | 20035.9 | 20149 |
| Antibody + DNA | | | 4680.3 | 2752.1 |
| | | | 4256.4 | 2931.2 |
| GST+ Antibody+ DNA | 6140.9 | 4833.8 | 5876.4 | 3431.3 |
| | 15447.7 | 4556.1 | 5937.5 | 3774.1 |
| GST-ZF+ Antibody+ DNA | 42105 | 22832.3 | 32087.5 | 22892.9 |
| | 27152.3 | 21727 | 28974 | 25989.7 |

Statistical analysis of the effect of order of addition and - volume on the binding of DNA to SPA beads, t Test: Paired Two Sample for Means (7.2.3).

| Effect of addition of antibody | | | | | |
|---|---|---|---|---|---|
| 50µl | | | 100µl | | |
| | No antibody | Antibody | | No antibody | Antibody |
| Mean | 12768.95 | 4468.35 | Mean | 10794.3 | 2841.65 |
| Variance | 1.06E+08 | 89845.6 | Variance | 43308263 | 16038.41 |
| Observations | 2 | 2 | Observations | 2 | 2 |
| Pearson Correlation | -1 | | Pearson Correlation | 1 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 1 | | df | 1 | |
| t Stat | 1.109869 | | t Stat | 1.742531 | |
| P(T<=t) one-tail | 0.233439 | | P(T<=t) one-tail | 0.165836 | |
| t Critical one-tail | 6.313749 | | t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.466879 | | P(T<=t) two-tail | 0.331673 | |
| t Critical two-tail | 12.70615 | | t Critical two-tail | 12.70615 | |

| Order of addition | | | | | |
|---|---|---|---|---|---|
| 50µl GST | | | 50µl GST-ZF | | |
| | Order A | Order B | | Order A | Order B |
| Mean | 10794.3 | 5906.95 | Mean | 34628.65 | 30530.75 |
| Variance | 43308263 | 1866.605 | Variance | 1.12E+08 | 4846941 |
| Observations | 2 | 2 | Observations | 2 | 2 |
| Pearson Correlation | 1 | | Pearson Correlation | 1 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 1 | | df | 1 | |
| t Stat | 1.057216 | | t Stat | 0.69226 | |
| P(T<=t) one-tail | 0.241149 | | P(T<=t) one-tail | 0.307259 | |
| t Critical one-tail | 6.313749 | | t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.482299 | | P(T<=t) two-tail | 0.614519 | |
| t Critical two-tail | 12.70615 | | t Critical two-tail | 12.70615 | |

| 100µl GST | | |
|---|---|---|
| | Order A | Order B |
| Mean | 4694.95 | 3602.7 |
| Variance | 38558.64 | 58755.92 |
| Observations | 2 | 2 |
| Pearson Correlation | -1 | |
| Hypothesised Mean Difference | 0 | |
| df | 1 | |
| t Stat | 3.520548 | |
| P(T<=t) one-tail | 0.088095 | |
| t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.176189 | |
| t Critical two-tail | 12.70615 | |

| Effect of volume | | | | | |
|---|---|---|---|---|---|
| GST order A | | | GST order B | | |
| | 50µl | 100µl | | 50µl | 100µl |
| Mean | 10794.3 | 4694.95 | Mean | 5906.95 | 3602.7 |
| Variance | 43308263 | 38558.64 | Variance | 1866.605 | 58755.92 |
| Observations | 2 | 2 | Observations | 2 | 2 |
| Pearson Correlation | -1 | | Pearson Correlation | 1 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 1 | | df | 1 | |
| t Stat | 1.272753 | | t Stat | 16.3596 | |
| P(T<=t) one-tail | 0.211981 | | P(T<=t) one-tail | 0.019433 | |
| t Critical one-tail | 6.313749 | | t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.423963 | | P(T<=t) two-tail | 0.038866 | |
| t Critical two-tail | 12.70615 | | t Critical two-tail | 12.70615 | |
| GST-ZF order A | | | | | |

| | | |
|---|---|---|
| | 50µl | 100µl |
| Mean | 34628.65 | 22279.65 |
| Variance | 1.12E+08 | 610844 |
| Observations | 2 | 2 |
| Pearson Correlation | 1 | |
| Hypothesised Mean Difference | 0 | |
| df | 1 | |
| t Stat | 1.783584 | |
| P(T<=t) one-tail | 0.162656 | |
| t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.325311 | |
| t Critical two-tail | 12.70615 | |

## APPENDIX XXIII

Effect of amount of antibody on the binding of DNA to SPA beads. Total cpm / assay = 242691.7 (7.2.3, Table 7.6).

| Antibody concentration | DNA only | GST | GST-ZF |
|---|---|---|---|
| 0.0mg/ml | 5789.8 | 2669.5 | 3368.3 |
| | 19484.2 | 19284 | 16826.1 |
| 1.0mg/ml | 2275.8 | 2651.8 | 26826.8 |
| | 5701.8 | 5994.4 | 24599.4 |
| 1.5mg/ml | 2488.3 | 2307.3 | 27231.1 |
| | 3663.7 | 3520 | 20674.5 |
| 2.0mg/m; | 2319.1 | 2315.2 | 31133.8 |
| | 3382.1 | 4229 | 25141.8 |
| 3.0mg/ml | 2492.4 | 3892.4 | 10214.2 |
| | 3539.5 | 3522 | 33527.1 |

Statistical analysis of the effect of amount of antibody on the binding of DNA to SPA beads, t Test: Paired Two Sample for Means (7.2.3).

| Effect of the addition of antibody to GST | | | | | |
|---|---|---|---|---|---|
| | No antibody | 1.0mg/ml | | 1.0mg/ml | 2.0mg/ml |
| Mean | 10976.75 | 4323.1 | Mean | 4323.1 | 3272.1 |
| Variance | 1.38E+08 | 5586487 | Variance | 5586487 | 1831315 |
| Observations | 2 | 2 | Observations | 2 | 2 |
| Pearson Correlation | 1 | | Pearson Correlation | 1 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 1 | | df | 1 | |
| t Stat | 1.002667 | | t Stat | 1.471165 | |
| P(T<=t) one-tail | 0.249576 | | P(T<=t) one-tail | 0.190029 | |
| t Critical one-tail | 6.313749 | | t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.499152 | | P(T<=t) two-tail | 0.380058 | |
| t Critical two-tail | 12.70615 | | t Critical two-tail | 12.70615 | |

| Effect of the addition of antibody to GST-ZF | | | | | |
|---|---|---|---|---|---|
| 0.0mg/ml and 1.0mg/ml | | | 1.0mg/ml and 2.0mg/ml | | |
| | No antibody | 1.0mg/ml | | 1.0mg/ml | 2.0mg/ml |
| Mean | 10097.2 | 25713.1 | Mean | 25713.1 | 28137.8 |
| Variance | 90556190 | 2480655 | Variance | 2480655 | 17952032 |
| Observations | 2 | 2 | Observations | 2 | 2 |
| Pearson Correlation | -1 | | Pearson Correlation | 1 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 1 | | df | 1 | |
| t Stat | -1.99116 | | t Stat | -1.28816 | |
| P(T<=t) one-tail | 0.148148 | | P(T<=t) one-tail | 0.210124 | |
| t Critical one-tail | 6.313749 | | t Critical one-tail | 6.313749 | |
| P(T<=t) two-tail | 0.296296 | | P(T<=t) two-tail | 0.420247 | |
| t Critical two-tail | 12.70615 | | t Critical two-tail | 12.70615 | |

## APPENDIX XXIV

Demonstration of the specificity of binding of TS DNA to GST-ZF. Total counts = 314931.5 (7.3, Fig. 7.3).

| No competition | |
|---|---|
| GST | GST-ZF |
| 14485 | 57886 |
| 14895 | 46138 |
| 15899 | 28774 |
| 7456 | 48424 |

| 1xTS competition | | 10xTS competition | |
|---|---|---|---|
| GST | GST-ZF | GST | GST-ZF |
| 10930 | 20332 | 6593 | 11054 |
| 12631 | 25105 | 6087 | 11483 |
| 10402 | 26816 | 6025 | 12305 |
| 11965 | 21184 | 11038 | 11214 |

| 1xC1 competition | | 10xC1 competition | |
|---|---|---|---|
| GST | GST-ZF | GST | GST-ZF |
| 10631 | 40873 | 6369 | 22713 |
| 12669 | 45112 | 5994 | 22694 |
| 10264 | 43603 | 9196 | 23619 |
| 9772 | 44102 | 6249 | 22524 |

| 1xC2 competition | | 10xC2 competition | |
|---|---|---|---|
| GST | GST-ZF | GST | GST-ZF |
| 5269 | 21425 | 4834 | 8498 |
| 5282 | 19222 | 6245 | 9527 |
| 5185 | 21248 | 5014 | 9618 |
| 8818 | 20677 | 5334 | 10179 |

| 1xC3 competition | | 10xC3 competition | |
|---|---|---|---|
| GST | GST-ZF | GST | GST-ZF |
| 9720 | 17397 | 4988 | 8322 |
| 5001 | 20947 | 4737 | 9135 |
| 5236 | 16452 | 5154 | 8478 |
| 4970 | 16261 | 4544 | 8629 |

| 1xC4 competition | | 10xC4 competition | |
|---|---|---|---|
| GST | GST-ZF | GST | GST-ZF |
| 4249 | 19182 | 4756 | 8370 |
| 6311 | 19311 | 5453 | 8271 |
| 5436 | 18661 | 5170 | 8856 |
| 5547 | 18084 | 4771 | 9911 |

Statistical analysis t-test (Paired Two Sample for Means) demonstrating the significance of the differences between assay with no competition and those with competing cold DNA (7.3, Table. 7.7).

| 1xTS | | | 10xTS | | |
|---|---|---|---|---|---|
| | No competitor | 1xTS | | No competitor | 10xTS |
| Mean | 45305.5 | 23359.25 | Mean | 45305.5 | 11514 |
| Variance | 1.47E+08 | 9630906 | Variance | 1.47E+08 | 309414 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.88988 | | Pearson Correlation | -0.97876 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 2.932713 | | t Stat | 5.32874 | |
| P(T<=t) one-tail | 0.030434 | | P(T<=t) one-tail | 0.006458 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.060867 | | P(T<=t) two-tail | 0.012915 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 1xC1 | | | 10xC1 | | |
|---|---|---|---|---|---|
| | No competitor | 1xC1 | | No competitor | 10xC1 |
| Mean | 45305.5 | 43422.5 | Mean | 45305.5 | 22887.5 |
| Variance | 1.47E+08 | 3282887 | Variance | 1.47E+08 | 245039 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.47793 | | Pearson Correlation | -0.8645 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 0.287464 | | t Stat | 3.567419 | |
| P(T<=t) one-tail | 0.396235 | | P(T<=t) one-tail | 0.018811 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.792471 | | P(T<=t) two-tail | 0.037621 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 1xC2 | | | 10xC2 | | |
|---|---|---|---|---|---|
| | No competitor | 1xC2 | | No competitor | 10xC2 |
| Mean | 45305.5 | 20643 | Mean | 45305.5 | 9455.5 |
| Variance | 1.47E+08 | 999315.3 | Variance | 1.47E+08 | 490592.3 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | -0.03408 | | Pearson Correlation | -0.48683 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 4.038799 | | t Stat | 5.738864 | |
| P(T<=t) one-tail | 0.013654 | | P(T<=t) one-tail | 0.005253 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.027309 | | P(T<=t) two-tail | 0.010506 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 1xC3 | | | 10xC3 | | |
|---|---|---|---|---|---|
| | No competitor | 1xC3 | | No competitor | 10xC3 |
| Mean | 45305.5 | 17764.25 | Mean | 45305.5 | 8641 |
| Variance | 1.47E+08 | 4748844 | Variance | 1.47E+08 | 124170 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.189474 | | Pearson Correlation | -0.07363 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 4.621585 | | t Stat | 6.025982 | |
| P(T<=t) one-tail | 0.009535 | | P(T<=t) one-tail | 0.00458 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.01907 | | P(T<=t) two-tail | 0.009161 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| 1xC4 | | | 10xC4 | | |
|---|---|---|---|---|---|
| | No competitor | 1xC4 | | No competitor | 10xC4 |
| Mean | 45305.5 | 18809.5 | Mean | 45305.5 | 8852 |
| Variance | 1.47E+08 | 312887 | Variance | 1.47E+08 | 563794 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.260021 | | Pearson Correlation | -0.1211 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 4.414346 | | t Stat | 5.950907 | |
| P(T<=t) one-tail | 0.010787 | | P(T<=t) one-tail | 0.004745 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.021574 | | P(T<=t) two-tail | 0.00949 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

Comparison of the effect of different DNA sequences at equi-molar concentration to the labelled TS DNA t-test; Paired Two Sample for Means (7.3, Table. 7.8).

| Comparison of TS and C1 | | | Comparison of TS and C2 | | |
|---|---|---|---|---|---|
| | 1xTS | 1xC1 | | 1xTS | 1xC2 |
| Mean | 23359.25 | 43422.5 | Mean | 23359.25 | 20643 |
| Variance | 9630906 | 3282887 | Variance | 9630906 | 999315.3 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.581743 | | Pearson Correlation | -0.30414 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | -15.8967 | | t Stat | 1.53548 | |
| P(T<=t) one-tail | 0.000271 | | P(T<=t) one-tail | 0.111119 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.000541 | | P(T<=t) two-tail | 0.222238 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

| Comparison or TS and C3 | | | Comparison of TS and C4 | | |
|---|---|---|---|---|---|
| | 1xTS | 1xC3 | | 1xTS | 1xC4 |
| Mean | 23359.25 | 17764.25 | Mean | 23359.25 | 18809.5 |
| Variance | 9630906 | 4748844 | Variance | 9630906 | 312887 |
| Observations | 4 | 4 | Observations | 4 | 4 |
| Pearson Correlation | 0.266253 | | Pearson Correlation | 0.156049 | |
| Hypothesised Mean Difference | 0 | | Hypothesised Mean Difference | 0 | |
| df | 3 | | df | 3 | |
| t Stat | 3.408402 | | t Stat | 2.967611 | |
| P(T<=t) one-tail | 0.0211 | | P(T<=t) one-tail | 0.029591 | |
| t Critical one-tail | 2.353363 | | t Critical one-tail | 2.353363 | |
| P(T<=t) two-tail | 0.0422 | | P(T<=t) two-tail | 0.059181 | |
| t Critical two-tail | 3.182449 | | t Critical two-tail | 3.182449 | |

## APPENDIX XXV

Effect of a wide range of [DNA], cpm/pmol=351976.5 (7.4.1, Fig 7.4).

| [DNA] | pmol/assay | Total cpm / assay | cpm / Assay | |
|---|---|---|---|---|
| | | | GST-ZFG | GST |
| 0.1nM | 0.005pmol | 1759.9 | 635 | 266 |
| | | | 415 | 350 |
| | | | 231 | 1173 |
| | | | 591 | 509 |
| 1.0nM | 0.05pmol | 17598.8 | 3190 | 347 |
| | | | 2294 | 542 |
| | | | 2299 | 415 |
| | | | 3396 | 312 |
| 10nM | 0.5pmol | 175988.2 | 12602 | 2127 |
| | | | 11336 | 3649 |
| | | | 13274 | 1919 |
| | | | 12824 | 2401 |
| 20nM | 1.0pmol | 351976.5 | 17006 | 3415 |
| | | | 16042 | 3601 |
| | | | 13971 | 4540 |
| | | | 24108 | 4535 |
| 30nM | 1.5pmol | 527964.7 | 17057 | 6064 |
| | | | 17139 | 5182 |
| | | | 19319 | 5688 |
| | | | 19128 | 5396 |
| 40nM | 2.0pmol | 703953 | 15899 | 9051 |
| | | | 19674 | 7301 |
| | | | 17508 | 8199 |
| | | | 17284 | 8724 |
| 50nM | 2.5pmol | 879941.2 | 21441 | 10361 |
| | | | 17043 | 9223 |
| | | | 17291 | 10140 |
| | | | 22071 | 8523 |
| 60nM | 3.0pmol | 1055930 | 17374 | 11339 |
| | | | 18217 | 11786 |
| | | | 18963 | 10610 |
| | | | 17372 | 10048 |
| 70nM | 3.5pmol | 1231917 | 16276 | 11592 |
| | | | 19831 | 12353 |
| | | | 17244 | 12232 |
| | | | 18960 | 12486 |
| 80nM | 4.0pmol | 1407906 | 19017 | 14375 |
| | | | 22345 | 15604 |
| | | | 21627 | 16721 |
| | | | 21082 | 18100 |
| 90nM | 4.5pmol | 1583894 | 21723 | 16060 |
| | | | 21253 | 16746 |
| | | | 19553 | 16660 |
| | | | 18950 | 14784 |
| 100nM | 5.0pmol | 1759882 | 17413 | 18573 |
| | | | 23456 | 18493 |
| | | | 23043 | 14029 |
| | | | 20386 | 18620 |

## APPENDIX XXVI

Effect of variation of protein concentration on the binding of DNA to SPA beads, Total CPM/ Assay = 68375 (7.4.1, Fig 7.5).

| [Protein] | cpm / assay GST | cpm/ assay GST-ZF |
|---|---|---|
| 5mM | 7425 | 29702 |
| | 5725 | 30736 |
| | 5372 | 30470 |
| | 3112 | 20089 |
| 7.5mM | 5812 | 28830 |
| | 5142 | 37599 |
| | 6461 | 34165 |
| | 8470 | 38853 |
| 10mM | 6716 | 37794 |
| | 7176 | 40594 |
| | 8755 | 43502 |
| | 4694 | 42111 |
| 12.5mM | 7724 | 44775 |
| | 8918 | 40459 |
| | 6022 | 42658 |
| | 7840 | 42837 |
| 15mM | 10324 | 58506 |
| | 6590 | 49990 |
| | 10609 | 47796 |
| | 6365 | 46466 |
| 17.5mM | 5835 | 42073 |
| | 11464 | 48403 |
| | 19885 | 49072 |
| | 9996 | 49791 |
| 20mM | 12351 | 50821 |
| | 9277 | 45287 |
| | 13469 | 45314 |
| | 13166 | 50209 |

## APPENDIX XXVII

Effect of a variation of DNA concentration over a low range (7.4.1, Fig 7.6).

| [DNA] | Total cpm / assay | cpm GST assay | cpm GST-ZF assay |
|---|---|---|---|
| 2nM | 6167 | 826 | 5711 |
| | | 274 | 3675 |
| | | 938 | 3020 |
| | | 271 | 5318 |
| 4nM | 12334 | 808 | 5448 |
| | | 558 | 5578 |
| | | 924 | 5434 |
| | | 997 | 5643 |
| 6nM | 18500 | 877 | 5445 |
| | | 958 | 5246 |
| | | 1016 | 5284 |
| | | 841 | 5306 |
| 8nM | 24667 | 1280 | 9379 |
| | | 1245 | 7816 |
| | | 1402 | 7650 |
| | | 2568 | 9408 |
| 10nM | 30834 | 1661 | 9769 |
| | | 1347 | 10101 |
| | | 1370 | 10237 |
| | | 1594 | 10376 |
| 12nM | 37000 | 1389 | 12035 |
| | | 1504 | 10086 |
| | | 1500 | 11662 |
| | | 1888 | 10668 |
| 14nM | 43168 | 1486 | 11443 |
| | | 1810 | 11751 |
| | | 1859 | 11982 |
| | | 1840 | 12482 |
| 16nM | 49335 | 1925 | 12247 |
| | | 2450 | 12071 |
| | | 1622 | 12116 |
| | | 1845 | 12946 |

## APPENDIX XXVIII

Variation of concentration of DNA 1-16nM CPM/pmol = 247469 (7.4.2, Fig. 7.7).

| DNA concentration (nM) | GST | GST-ZF | DNA concentration (nM) | GST | GST-ZF |
|---|---|---|---|---|---|
| 1 | 587 | 4761 | 9 | 3695 | 19641 |
| | 1112 | 4313 | | 3424 | 19286 |
| | 352 | 4196 | | 3743 | 19604 |
| | 1216 | 3981 | | 3349 | 19854 |
| 2 | 1278 | 7032 | 10 | 3948 | 19976 |
| | 1157 | 7635 | | 4332 | 19935 |
| | 1066 | 6948 | | 3974 | 19171 |
| | 1087 | 7883 | | 3479 | 20031 |
| 3 | 1492 | 9166 | 11 | 4752 | 19420 |
| | 1527 | 9813 | | 4615 | 19846 |
| | 1521 | 9871 | | 4961 | 20417 |
| | 1693 | 8847 | | 5197 | 19434 |
| 4 | 2104 | 11178 | 12 | 4275 | 20392 |
| | 1697 | 11033 | | 4144 | 20052 |
| | 2593 | 11717 | | 3961 | 19770 |
| | 1914 | 10773 | | 4044 | 20609 |
| 5 | 2362 | 13357 | 13 | 4532 | 20430 |
| | 2771 | 12726 | | 4697 | 19471 |
| | 2310 | 13037 | | 4998 | 20130 |
| | 2952 | 13362 | | 6538 | 20768 |
| 6 | 3282 | 14998 | 14 | 4928 | 20494 |
| | 2347 | 14602 | | 4502 | 20540 |
| | 4264 | 15035 | | 4429 | 21052 |
| | 2715 | 15087 | | 6931 | 20571 |
| 7 | 3420 | 16490 | 15 | 5205 | 20398 |
| | 2977 | 15920 | | 5583 | 21128 |
| | 3289 | 17038 | | 5639 | 21223 |
| | 3633 | 16546 | | 5417 | 20303 |
| 8 | 3483 | 18611 | 16 | 5172 | 20451 |
| | 3701 | 19243 | | 5708 | 20698 |
| | 3199 | 18643 | | 5961 | 21136 |
| | 3095 | 18535 | | 7060 | 21313 |

Calculation of bound and free ligand for Kd plots (7.4.2, Fig. 7.8 and 7.9).

| | | | | | Background not subtracted (Fig. 7.8) | | Background subtracted (Fig. 7.9) | |
|---|---|---|---|---|---|---|---|---|
| [DNA] nM | Totals | Mean GST cpm | Mean GST-ZF cpm | Mean GST-ZF cpm - Mean GST cpm | nM DNA bound | nM DNA free | nM DNA bound | nM DNA free |
| 1 | 10609 | 816.75 | 4312.75 | 3496 | 0.34855 | 0.65145 | 0.28254 | 0.71746 |
| 2 | 22404 | 1147 | 7374.5 | 6227.5 | 0.595996 | 1.404004 | 0.503295 | 1.496705 |
| 3 | 31882 | 1558.25 | 9424.25 | 7866 | 0.761654 | 2.238346 | 0.635716 | 2.364284 |
| 4 | 44213 | 2077 | 11175.25 | 9098.25 | 0.903167 | 3.096833 | 0.735304 | 3.264696 |
| 5 | 54483 | 2598.75 | 13120.5 | 10521.75 | 1.06038 | 3.93962 | 0.850349 | 4.149651 |
| 6 | 65492 | 3152 | 14930.5 | 11778.5 | 1.206661 | 4.793339 | 0.951917 | 5.048083 |
| 7 | 78041 | 3329.75 | 16498.5 | 13168.75 | 1.333385 | 5.666615 | 1.064275 | 5.935725 |
| 8 | 86385 | 3369.5 | 18758 | 15388.5 | 1.515994 | 6.484006 | 1.243671 | 6.756329 |
| 9 | 100921 | 3552.75 | 19596.25 | 16043.5 | 1.58374 | 7.41626 | 1.296607 | 7.703393 |
| 10 | 112793 | 3933.25 | 19778.25 | 15845 | 1.598449 | 8.401551 | 1.280564 | 8.719436 |
| 11 | 122826 | 4881.25 | 19779.25 | 14898 | 1.59853 | 9.40147 | 1.20403 | 9.79597 |
| 12 | 135198 | 4106 | 20205.75 | 16099.75 | 1.632999 | 10.367 | 1.301153 | 10.69885 |
| 13 | 142313 | 5191.25 | 20199.75 | 15008.5 | 1.632514 | 11.36749 | 1.21296 | 11.78704 |
| 14 | 155767 | 5197.5 | 20664.25 | 15466.75 | 1.670054 | 12.32995 | 1.249995 | 12.75001 |
| 15 | 170638 | 5461 | 20763 | 15302 | 1.678035 | 13.32196 | 1.23668 | 13.76332 |
| 16 | 185004 | 5975.25 | 20899.5 | 14924.25 | 1.689067 | 14.31093 | 1.206151 | 14.79385 |