Regulation of the *Xenopus* gene *Xsna*

Gwyndaf Thomas Roberts

1996

Division of Developmental Biology
National Institute for Medical Research
The Ridgeway
Mill Hill
London
NW7 1AA

A dissertation submitted to the University of London for the degree of Doctor of Philosophy
"...in the evening light this country can be beautiful....
shadows soften the hillsides; there is a blending of
lines and folds.... so that one comes to bless the
absolute bareness, feeling that here is a pure beauty
of form, a kind of ultimate harmony."

George Leigh Mallory, 1921.
Acknowledgements

If I wrote a few words to thank, one by one, everyone who have helped, supported and cajoled me through the last four years I would need a separate volume of this thesis to contain the extra leaves of paper. Therefore, it is to those who do not find their names below, but who deserve a mention at least as much, that I thank first. The students and staff who have become firm friends, the NIMROD Hill Walkers, both official and unofficial, and the Welsh exiles have all provided those distractions that are so important for a balanced life.

I thank the MRC, Sir John Skehl and Dr Rod King for allowing me to follow my studies at NIMR, and the past and present members of the Division of Developmental Biology, in particular, Drs. Jim Smith, Tim Mohun, Alison Snape and Helen New, who have shown immense patience whenever I have asked for help, advice, or their opinion on my latest discovery in the library. I am indebted to Dr. Michael Sargent for providing me with the opportunity to work in his lab, but more importantly for his tutelage, and infinite patience and optimism. The other members of the lab, Mike, Richard, Linda and Roberto, have all contributed in numerous ways to the happy memories that I will carry with me in the future. Who knows, maybe Crewe Alexandra will win the FA Cup. And, maybe George Morgan will teach his dad not to ignore his vintage Celtic blood.

I don't think I have ever properly thanked my parents for their unconditional love and support over the past 26 years. If I have been able to make them proud, then I may have gone some way to repaying an impossible debt. I am grateful to my brother, Gethin, for attempting to keep me in touch with the Scarlets, and my "sister" Sara for showing me there are better diets than mashed potato and luncheon meat. Finally, I thank Robyn for trying very hard to appreciate the natural flair and talent inherent in all Welsh rugby teams. "Howay the lads!"

October, 1996.

Barnet, Hertfordshire.
I fy Mam a Nhad, diolch am pobpeth.

For Mum and Dad, thanks for everything.
Abstract

*Xsna* is a zinc finger transcription factor that is expressed in the developing *Xenopus* embryo in the mesoderm from stage 9 and the prospective neural crest from stage 11. The promoter of this gene contains no sequences that resemble the motifs of known transcription factors. However, it has previously been shown that two adjacent elements within the promoter (approximately 100 base pairs from the transcriptional start site), are responsible for the differential expression pattern of the gene.

Using the electrophoretic mobility shift assay (EMSA) specific DNA binding activities have been associated with the sequences within these regions. Their binding preferences have been further defined using a range of mutated oligonucleotides in the EMSA, and by DNA protection analysis. Binding activity has been detected in extracts prepared from ovary, unfertilised eggs and embryos up to stage 35, indicating that their presence is in itself insufficient to initiate *Xsna* transcription. A combination of gel filtration and UV crosslinking analyses suggest that both activities are about 370 kDa in size, with ~97 kDa in direct contact with the DNA. Several purification techniques have been considered as possible components of a purification protocol but an inherent instability of the complexes has limited the use of these conventional steps.

To overcome these problems, a new approach to DNA affinity purification has been developed using biotinylated double-stranded oligonucleotides and streptavidin coated magnetic beads. The protein-DNA complex is released from the beads by reducing a disulphide bond positioned between the biotin moiety and the specific binding motif. The optimal conditions for this technique have been established, and it will now be possible to isolated sufficient material for N-terminal sequencing, from which the proteins responsible for the regulation will be identified.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>APS</td>
<td>ammonium persulphate</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>BrdU</td>
<td>bromodeoxyuridine</td>
</tr>
<tr>
<td>Bio-E</td>
<td>biotinylated probe E</td>
</tr>
<tr>
<td>Bio-M</td>
<td>biotinylated probe M</td>
</tr>
<tr>
<td>Bio-P</td>
<td>biotinylated probe P</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyl</td>
</tr>
<tr>
<td></td>
<td>transferase</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>DEAE</td>
<td>diethyl amino ethyl</td>
</tr>
<tr>
<td>dH₂O</td>
<td>distilled H₂O</td>
</tr>
<tr>
<td>DMS</td>
<td>dimethylsulphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethyl sulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic</td>
</tr>
<tr>
<td>DNase</td>
<td>deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleoside</td>
</tr>
<tr>
<td></td>
<td>triphosphate</td>
</tr>
<tr>
<td>dsDNA</td>
<td>double-stranded DNA</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E</td>
<td>ectodermal</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tereacetic</td>
</tr>
<tr>
<td></td>
<td>acid</td>
</tr>
<tr>
<td>eIF</td>
<td>eukaryotic translation</td>
</tr>
<tr>
<td></td>
<td>initiation factor</td>
</tr>
<tr>
<td>EMSA</td>
<td>electrophoretic mobility</td>
</tr>
<tr>
<td></td>
<td>shift assay</td>
</tr>
<tr>
<td>fmol</td>
<td>femtomole</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>λ</td>
<td>lambda</td>
</tr>
<tr>
<td>M</td>
<td>mesodermal</td>
</tr>
<tr>
<td>MPC</td>
<td>magnetic particle collector</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>NAM</td>
<td>normal amphibian medium</td>
</tr>
<tr>
<td>nmol</td>
<td>nanomole</td>
</tr>
<tr>
<td>NP40</td>
<td>Nonidet P-40</td>
</tr>
<tr>
<td>OP-Cu²⁺</td>
<td>Copper-Phenanthroline</td>
</tr>
<tr>
<td>PAG</td>
<td>non-denaturing polyacrylamide gel</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PKC</td>
<td>protein kinase C</td>
</tr>
<tr>
<td>pmol</td>
<td>picomole</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>ribonuclease</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SDS-PAG</td>
<td>denaturing polyacrylamide</td>
</tr>
<tr>
<td></td>
<td>gel</td>
</tr>
<tr>
<td>³⁵SLR</td>
<td>sulphur labelling reagent</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-borate-EDTA buffer</td>
</tr>
<tr>
<td>TEMED</td>
<td>N ,N ′ ,N ″ -tetramethylene-</td>
</tr>
<tr>
<td></td>
<td>diamine</td>
</tr>
<tr>
<td>TFIIIA</td>
<td>transcription factor IIIA</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris (hydroxymethyl)-</td>
</tr>
<tr>
<td></td>
<td>aminomethane</td>
</tr>
<tr>
<td>UFE</td>
<td>unfertilised eggs</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
</tbody>
</table>
## Contents

<table>
<thead>
<tr>
<th>Quotation</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iv</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>v</td>
</tr>
<tr>
<td>Contents</td>
<td>vi</td>
</tr>
<tr>
<td>Tables</td>
<td>x</td>
</tr>
<tr>
<td>Figures</td>
<td>x</td>
</tr>
</tbody>
</table>

### Chapter 1 - General Introduction

1. **Xenopus laevis**
   1.1 Xenopus laevis
   1.2 Establishment of the Embryonic Body Plan in Xenopus
   1.3 Mesoderm Induction
   1.4 Neural Induction
   1.5 The Neural Crest
   1.6 Xsna and Xslu
   1.7 Drosophila snail
   1.8 The snail Family
   1.9 Establishment of the Dorsal-Ventral axis in Drosophila

### Chapter 2 - Materials and Methods

2. **Chemicals, Buffers and Media**
   2.1 Chemicals, Buffers and Media
      2.1.1 General
      2.1.2 Unfertilised Eggs and Embryos
      2.1.3 DNA Manipulations
      2.1.4 Protein Manipulations
      2.1.5 RNA Manipulations
## 2.2 Unfertilised Eggs and Embryos
- 2.2.1 Collection of Unfertilised Eggs and Embryos
- 2.2.2 Preparation of Whole Cell Extract
- 2.2.3 Microinjection of Xenopus Embryos

## 2.3 DNA Manipulations
- 2.3.1 Gel Electrophoresis of DNA
- 2.3.2 Polymerase Chain Reaction
- 2.3.3 dsDNA Sequencing

## 2.4 Synthesis of Radiolabelled Probes
- 2.4.1 EMSA Probes
- 2.4.2 Copper-Phenathroline Protection Assay Probe
- 2.4.3 Biotinylated Probes
- 2.4.4 UV Crosslinking Probes
- 2.4.5 RNAse Protection Probes

## 2.5 Protein Purification
- 2.5.1 Ammonium Sulphate Precipitation
- 2.5.2 Hydrophobic Interaction Chromatography
- 2.5.3 Ion Exchange Chromatography
- 2.5.4 DNA Affinity Chromotography

## 2.6 Characterisation of Proteins
- 2.6.1 Electrophoretic Mobility Shift Assay
- 2.6.2 Crosslinking
- 2.6.3 Gel Filtration
- 2.6.4 DNA Footprinting
- 2.6.5 Denaturing Polyacrylamide Gel Electrophoresis (SDS-PAG) for Protein Separation
- 2.6.6 Post Synthetic Labelling of DNA-bound Protein with \(^{35}\)SLR
- 2.6.7 Protease Clipping

## 2.7 RNA Manipulations
- 2.7.1 Preparation of Capped RNA for Injection
- 2.7.2 Preparation of RNA from the Embryos
- 2.7.3 RNAse Protection Assay
Chapter 7 - Biochemical Characterisation of the Xsna Promoter Binding Activities 153

7.1 Introduction 154
7.2 Determination of the Size of the Complexes Involved 154
7.3 Analysis of Individual Peptides 155
7.4 A Comparison of the M and E Complexes Using a Protease Clipping Assay 166
7.5 Discussion 169
7.6 Summary 169

Chapter 8 - The Regulation of Xsna 171

8.1 Introduction 172
8.2 Xsna Induction is Regulated by PKCα 174
8.3 Discussion 174
8.4 Summary 179

Chapter 9 - General Discussion 180

9.1 Binding Activities Associated with the Xsna Promoter 181
9.2 Transcriptional Control of Xsna 183
9.3 Purification of Xsna Promoter Binding Proteins 185
9.4 Similarities Between the M and E Complexes 187
9.5 Speculation about the Regulation of Xsna 188

Appendix A - Summary of the Probes 190

References 194
Tables
1.1 The snail Family 34
3.1 The Effect of Components of the EMSA on the Binding Activities 75

Figures
1.1 The Three Signal Model 5
1.2 Vertical and Planar Routes of Neural Induction 12
1.3(a) Fold Experiments of Nieuwkoop Testing Planar Induction 16
(b) Nieuwkoop and Nigtevecht's Notochord Sandwich 16
1.4 Keller Explants 18
1.5 Deletion Analysis of the Xsna Promoter 28

2.1 Biotinylated Probes 54

3.1 The Effect of Increasing Concentrations of Lambda DNA as Non-specific Competitor in the EMSA 77
3.2 The Sensitivity of the Binding Activity to Storage at Low Temperatures 80
3.3 EMSA's Showing the Change in M and E Binding Activities After 48 hrs. 83
3.4 Developmental Profiles of the Binding Activities 86

4.1 The Binding Activities Observed Using Probe Set E 95
4.2 The Binding Activities Observed Using Probe Set M 97
4.3 DMS Protection Assay of the Promoter Region 101
4.4 Copper-Phenanthroline Footprinting Assay of the Promoter Region 104

5.1 Hydrophobic Interaction Chromatography 114
5.2 Ion Exchange Chromatography 116
5.3 DNA Affinity Chromatography 121

6.1 Isolation of Specific Protein-DNA Complexes Using Streptavidin Coated Magnetic Beads with Bio-M and Bio-E 128
6.2 Isolation of Specific Protein-DNA Complexes Using Bio-P 130
6.3 The Effect of (NH4)2SO4 on the Capacity of Streptavidin Coated Magnetic Beads to Bind Bio-M 134
6.4 Ammonium Sulphate Fractionation of Unfertilised Egg Extract 138
6.5 optimum Conditions for Washing the Streptavidin Coated Beads 142
6.6 An Attempt to Elute Binding Activity from the Streptavidin Coated Beads 142
6.7(a) The Effect of DTT on the Radiolabelling of Proteins by [35S]-SLR 145
(b) Labelling Proteins Using [35S]-SLR in the Presence of Toluene 145
6.8(a) Specific Competition Using the TFIIIA Oligonucleotide 148
(b) The Eluate from Bio-P Radiolabelled by [35S]-SLR 148
7.1(a) The Elution from a Sephacryl S400 Gel Filtration Column 156
(b) Molecular Weight Determination 156
7.2 The 2-D Resolution of UV Crosslink Protein-DNA Complexes 159
7.3 UV Crosslinking after Isolation Using the Streptavidin Coated Beads 161
7.4 Testing the Efficiency of the UV Crosslinking Reaction 164
7.5 Protease Clipping Assays 167
8.1 An RNAse Protection Assay Demonstrating that Xsna can Regulate its Own Transcription 175
8.2 The Order of Molecular Events Leading to Neural Crest Induction 177
General Introduction

1.1 Xenopus laevis

1.2 Establishment of the Embryonic Body Plan in Xenopus

1.3 Mesoderm Induction

1.4 Neural Induction

1.5 The Neural Crest

1.6 Xsna and Xslug

1.7 Drosophila snail

1.8 The snail Family

1.9 Establishment of the Dorsal-Ventral axis in Drosophila

2.0 Xsna and snail

2.1 A Brief Overview of Eukaryotic Transcription

2.2 Project Aims
1.1 **Xenopus laevis**

*Xenopus laevis*, the African clawed toad, has a special place in the study of vertebrate development because of the accessibility of the embryonic stages for dissection and the convenience with which substantial amounts of experimental material can be obtained for biochemical investigations. In the wild the eggs are laid in jelly coats simultaneously with fertilisation, and then the embryos develop outside of the mother. In the laboratory the female can be induced to spawn with the injection of human gonadotropic hormones so that her eggs can be fertilised artificially, and thereby begin their development synchronously. In recent times the traditional histological approach for describing the dynamics of embryonic development have been augmented with very efficient methods of *in situ* hybridisation, which reveal unprecedented details of embryonic organisation. Although classical genetics has limited value as a result of its long generation time of eighteen months and pseudotetraploidy, it has been possible to make incisive investigations into inductive events using growth factors in conjunction dominant negative interference.

1.2 **Establishment of the Embryonic Body Plan in Xenopus**

The first phase of vertebrate embryonic development is the establishment of the primary axes: the anterior-posterior and dorsal-ventral axis (a-p and d-v respectively). This begins during the six month period of oogenesis when maternal RNA and proteins are accumulated and localised within the oocyte, resulting in a cytoplasm rich animal pole and a yolky vegetal pole.

The d-v axis is determined during the first cell cycle after fertilisation when a cytoplasmic shift, called the cortical rotation, takes place. This involves the movement of the plasma membrane and cortex relative to the internal cytoplasm. The point at which the
edge of the vegetal cortex meets the animal cytoplasm indicates the prospective dorsal side, and the prospective ventral side occurs where the animal cortex meets the vegetal cytoplasm (Gerhart et al., 1989: figure 1.1). The net result is that the dorsal midline is positioned approximately 180° from the point of sperm entry. Ultra violet irradiation of the egg during the first cell cycle prevents the cortical rotation and results in a ventralised phenotype in which the dorsal mesodermal derivatives such as notochord, somites and neural plate are not formed. Manual rotation of the eggs restores the dorsal axis, with the new d-v polarity corresponding to the direction of imposed rotation (Gerhart et al., 1981).

1.3 Mesoderm Induction

The second phase of vertebrate development is the induction of mesoderm between the ectoderm of the animal pole and the endoderm of the vegetal pole. After fertilisation rapid cell division occurs creating a hollow ball of cells called the blastula, in which the animal cap cells are smaller than the yolky vegetal pole cells. The area around the equator of the embryo between these cell types is called the marginal zone, and is where the mesoderm is induced. The mesoderm first appears on the dorsal side at stage 9, and on the ventral side by stage 11. Shortly after being specified, the mesodermal cells begin to involute in a process known as gastrulation. This is the first instance of cell movement within the developing embryo.

Nieuwkoop (1968) demonstrated that the source of the induction was the vegetal pole. He prepared conjugates of ectoderm from animal caps, that would normally form epidermis if left alone to develop in the absence of any external stimuli. However, if cells of vegetal origin were juxtaposed beneath the caps, they produced mesodermal tissue. This was visualised by the elongation of the explants but is more accurately demonstrated
by histology and analysis with molecular markers (Dale et al., 1985). The type of mesoderm induced depended on the source of the vegetal explants: ventral vegetal explants induced ventral mesodermal cell types such as blood and mesenchyme, whereas dorsal explants induced dorsal derivatives such as notochord and somites (Dale et al., 1985). It has also been demonstrated that a dorsal vegetal blastomere transplanted from a healthy donor can rescue an embryo ventralised by UV irradiation, and induce a secondary axis in a healthy recipient (Gimlich and Gerhart, 1984).

This experiment confirmed and extended the experiments of Spemann and Mangold (1924) who showed that a dorsal marginal zone could induce a secondary axis when transplanted into a normal embryo. The dorsal marginal zone, defined as the organiser by Spemann and Mangold, is the most dorsal tissue in the embryo and lies directly above the dorsal vegetal cells which have been termed the Nieuwkoop Centre (Gerhart et al., 1989). As alluded to above, vegetal cells induce either ventral or dorsal mesodermal derivatives, but never lateral mesoderm alone (muscle without notochord) (Dale and Slack, 1987b). The fate map of *Xenopus* indicates most muscle arises from the ventral mesoderm (Dale and Slack, 1987a), but if the embryo is dissected into dorsal and ventral halves muscle is seen only in the former (Dale and Slack, 1987b), suggesting that a signal emanating from the organiser region is required to dorsalise the ventral mesoderm in order to produce lateral cell types, and thereby generate the complete mesodermal pattern.

These observations were incorporated by Smith and Slack (1983) into the three signal model (figure 1.1). In this model the first signal originating from the ventral vegetal cells induces ventral mesoderm; the second signal arising from the dorsal vegetal cells (the Nieuwkoop Centre) induces the dorsal cell types and consequently the organiser; and the third signal from the organiser dorsalises the induced mesoderm and establishes the correct
Figure 1.1 The Three Signal Model.
(See text for details)

(From Ladher, R., 1996, after modification from Smith and Slack, 1985)
Chapter 1 - General Introduction

(a) **Cortical Rotation**

- Animal pole
- Vegetal pole

(b) **Mesoderm Induction**

- Induction of Spemann's Organiser

(b) **Dorsalisation**

- Induction of Different Domains of Mesoderm.
  - VM Ventral mesoderm
  - VLM Ventrolateral mesoderm
  - DLM Dorsolateral mesoderm
  - DM Dorsal mesoderm

- Animal pole
- Vegetal pole

Marginal Zone

Nieuwkoop Centre no longer active

Dorsalisation occurring by morphogenetic gradient
pattern. Animal cap assays are used as a versatile test for any protein or mRNA that can modify the fate of naive ectoderm (Dawid, 1991). Animal cap explants can either be cultured in a medium containing the putative factor, or its mRNA can be injected into the embryos at the one cell stage and the animal explants prepared just before mesoderm induction begins, at stage 8. The explants are then examined morphologically, histologically and tested with molecular markers.

The three signal model was given powerful support when the growth factors activin, a TGFβ superfamily member, and members of the fibroblast growth factor (FGF) family were shown to produce phenotypes in explanted ectoderm, similar to that predicted by the theory.

It was suspected that FGF had an important role to play in mesoderm induction since RNA transcripts and the protein of basic FGF are present in the oocyte and early embryo (Kimelman et al., 1988; Slack and Isaacs, 1989). The capacity of bFGF to induce mesoderm of a ventral histology in animal cap explants (Kimelman and Kirschner, 1987; Slack et al., 1987) suggested that it was the vegetal ventralising component in the three signal model. However, it is now quite clear that there is no FGF localised in the vegetal pole to fulfil this role (Isaacs et al., 1992), in fact high levels of maternal FGF-2 are to be found in the animal pole of the blastula stage embryo (Song and Slack, 1994). It was also recently shown that after high levels of bFGF RNA are injected in to the animal hemispheres of one cell embryos, their excised animal caps form significant amounts of dorsal mesoderm derivatives. The consensus of opinion now is that, while the FGF family have a role to play in mesoderm development, they are more likely to be involved as competence factors for later genes (Schulte-Merker and Smith, 1995; Isaacs et al., 1995) rather than the initial induction.
Activin, which was known previously to have a role in blood cell development, was shown to be a mesoderm inducing factor after the observation that media from XTC cells, a *Xenopus* cell line, could induce animal cap explants to form mesodermal derivatives (Smith, 1989). In addition, when animal cells treated with activin were transplanted into the ventral region of a host embryo, a secondary axis was formed, mimicking the action of a transplanted organiser (Cooke, 1989). The unusual feature of this factor was that a spectrum of responses could be demonstrated in animal cap assays, with increasing doses of activin. Low doses of activin resulted in ventral mesodermal derivatives, and higher doses gave more dorsal tissues (Green et al., 1990). An equivalent concentration dependent effect was found with FGF (Green et al., 1992). It has been suggested that concentration gradients of the two growth factors within the embryo can account for the complete pattern of mesodermal tissues (Smith et al., 1993).

The role of activin in mesoderm induction is now in doubt for three main reasons: (i) mRNA's encoding the activins A and B are not present in *Xenopus* oocytes and embryos in significant amounts (Dohrmann et al., 1993; Rebagliati and Dawid, 1993); (ii) ectopic-expression of the activin-specific binding protein follistatin has no effect on the developing embryo (Schulte-Merker et al., 1994; Kessler and Melton, 1995); (iii) activin knockout mouse embryos form mesoderm normally (Matzuk et al., 1995). There are arguments against the significance of these observation (Fukui et al., 1994; Oda et al., 1995) but attention now seems to have switched to two other TGFβ superfamily members, Vg1 mRNA, encoding a protein of the TGFβ family, and bone morphogenetic protein 4 (BMP-4).

Vg1 is a maternal mRNA that is localised exclusively in the vegetal hemisphere of the oocyte (Weeks and Melton, 1987), but virtually no mature protein has been detected
in the embryo, and injection of the mRNA into the embryo yields no ectopic mesoderm (Dale et al., 1989; Tannahill and Melton, 1989). This seemed to preclude a role for Vg1 as a primary mesoderm inducer. However, Thomsen and Melton (1993) used a chimeric (BMP)-Vg1 molecule to show that a processed Vg1 is a powerful axial mesoderm inducer, resistant to inhibition by follistatin. Although it is still unclear whether Vg1 is processed in vivo, it remains a tantalising prospect.

BMP-4 is expressed throughout the mesoderm except in the most dorsal tissue, the organiser (Fainsod et al., 1995). It induces ventral mesoderm in animal cap assays and ectopic expression of BMP-4 mRNA results in a ventralised embryo (Dale et al., 1992; Jones et al., 1992). Interference with the BMP-4 signalling system by expressing the truncated receptor results in complete dorsalisation (Graff et al., 1994). One interpretation of these findings was that the role of BMP-4 was to limit the influence of the organiser, by inhibiting the dorsal organiser-inducing signal (Jones et al., 1992). This has now been shown not to be the case: BMP-4 exerts its influence during gastrula stages (Jones et al., 1996), which raises the possibility that ventral mesoderm formation requires an active signal and is not just a default state.

Two more genes with a role in embryonic patterning were discovered when Smith and Harland (1991, 1992) undertook an expression screen for genes that could rescue UV ventralised embryos. These were Xwnt-8, a relative of the Drosophila gene wingless (wg) and noggin, a novel secreted protein. Noggin possesses a number of characteristics that suggest it may fit the role of the dorsalising signal in the three signal model (figure 1.1). The timing and position of its expression perfectly match the pattern expected of the dorsal signal according to the three signal model, that is, in the organiser during gastrulation, and later in the notochord and prechordal plate, which are derived from the organiser. Noggin
is unique among the known mesoderm inducers in being able to induce muscle in ventral marginal zone explants from gastrula (Smith et al., 1993). Although noggin is an important candidate to fill the role of the dorsalising cytokine, the nodal related genes, Xnr-1, Xnr-2 and Xnr-3 (Jones et al., 1995; Smith et al., 1995), and chordin (Sasai et al., 1994) are also relevant. chordin is particularly intriguing as it shares strong homology with the Drosophila gene short gastrulation, the product of which inhibits the action of decapentaplegic, a TGFβ superfamily member closely related to BMP-4, and can indeed substitute for it (Sasai et al., 1994; François et al., 1994).

Although Xwnt-8 can rescue UV ventralised embryo and therefore must exercise dorsalising activity (Smith and Harland, 1991), its normal expression pattern is restricted to the ventral side of the embryo at late blastula and gastrula stages (Christian et al., 1991). Interestingly, Xwnt-8 was shown to inhibit the formation of the organiser (Christian et al., 1991). Xwnt-11, another member of the Wnt family, is localised in the vegetal hemisphere of the oocyte, in much the same way as Vg-1, and by late blastula stage is strongly expressed in the axial mesoderm, with a bias towards the dorsal side (Ku and Melton, 1993). Injected Xwnt-11 mRNA will partially rescue a UV ventralised embryo but it is not able to induce mesoderm itself. Although there is no evidence for localisation to the dorsal side of the maternal transcripts, it has been postulated that Xwnt-11 is the signal emanating from the Nieuwkoop Centre that is responsible for the induction of the organiser (Slack, 1994).

What is becoming clear from this exploration of the mechanisms of mesoderm development, is that there are several instances where unrelated molecules appear to direct similar aspects of the pattern formation. This suggests that mesoderm induction, and, by
extension, other inductive processes in the embryo, are subtle and elegant processes in which molecules work in concert and exert their influence in overlapping zones of activity.

There are several molecules that appear to be important to mesoderm induction, some of which have been described above, and there are others that hint at an involvement but have yet to have satisfactory functions assigned to them. Based on the current knowledge, modified versions of the three signal model (figure 1.1: Smith and Slack, 1983) have been proposed, in which a fourth ventralising signal (BMP-4 or Xwnt-8) acts to counter the dorsalising signal (noggin, Xnr-1,2 or 3, or chordin) that originates from the Xwnt-11 induced organiser (Sive, 1993; Cunliffe and Smith, 1994; Conlon et al., 1996; Jones et al., 1996). Pierce and Kimelman (1995) have even proposed a model in which FGF is a fifth signal arising from the animal pole, which would delimit the axial mesoderm to the region in which it overlaps with a vegetal signal, such as Vg-1.

1.4 Neural Induction

Neural development in the *Xenopus* embryo is induced in the ectoderm by signals from the dorsal mesoderm during gastrulation. This was demonstrated by the experiments described above in which the transplantation of an organiser region (Spemann and Mangold, 1924) into ventral regions of a host embryo results in the formation of a secondary axis from the host tissue. Spemann suggested that the signals that cause neural induction could follow different routes: a planar route, in which inductive signals pass through the plane between the organiser and the ectoderm, and, a vertical route whereby the signals pass to the ectoderm from the underlying involuted mesoderm (figure 1.2).

The experiments of Holtfreter (1933) and Mangold (1933) strongly suggested that vertical inductive signals were solely responsible for neuralisation of the ectoderm.
Figure 1.2  Vertical and Planar Routes of Neural Induction:
Schematic diagram of a sagittal section through a mid-gastrula embryo, showing the
relationship between the involuting dorsal mesoderm (striped) and the presumptive
neuroectoderm (dark shaded).
Orientation: animal pole, left; vegetal pole, right; dorsal, up. A = anterior, P = posterior.

(From, Doniach, 1993)
Holtfreter found that a urodele blastulae cultured in a high salt medium would exogastrulate, that is, the mesoderm and the endoderm would move outward rather than involute during gastrulation. This allowed him to observe the effect of a mesodermal signal on ectoderm during the period of normal neural induction, without the presence of underlying mesoderm to supply any vertical signals, but with the planar route still available. The result was the ectoderm remained a bag of undifferentiated epidermis, from which he concluded that planar signals were not sufficient for neural induction.

Mangold's experiment involved the transplantation of involuted mesoderm from a urodele neurula stage embryo into a host blastocoel; a procedure known as the Einsteck method. During gastrulation the transplant came into contact with the ventral ectoderm, inducing a neural fate. However, the most compelling feature of this result was that the anterior-posterior (a-p) characteristics of the neural tissue corresponded to that of the transplanted mesoderm. For example, anterior mesoderm induced brain and eye like tissues, and posterior mesoderm induced spinal cord like tissue. This suggested that not only were vertical signals primarily responsible for the induction, but also that the a-p neural pattern was conferred by qualitatively different signals positioned along the underlying mesoderm.

These experiments seem to preclude the existence of any significant contribution by a planar signal, but later experiments directed by Nieuwkoop indicated that planar induction does occur. Nieuwkoop investigated the planar signal hypothesis in two sets of experiments conducted in urodele embryos. In the first, folded flaps of competent ectoderm were inserted perpendicularly along the a-p axis of a presumptive neural plate (Nieuwkoop et al., 1952a, b). The expectation, if planar signals were transmitted into the flaps, was that histological characteristics of the implants would reflect their a-p position in
the neural plate. This turned out to be the case (figure 1.3a). The other experiment involved sandwiching a small piece of anterior notochord between two larger sheets of competent ectoderm (Nieuwkoop and Nigtevecht, 1954: figure 1.3b). They found evidence of spreading neural induction in parts of the ectoderm not in contact with the notochord.

More recently experiments using molecular markers such as *en-2* (Hemmati-Brivanlou et al., 1990), *XIF3*, *XIHbox6* and *XIF6* (Sharpe and Gurdon, 1990) have shown that although the inductive signals in the dorsal mesoderm do appear to be regionalised, the information is fairly roughly distributed. For example, both anterior and posterior mesoderm can induce the anterior neural marker *XIF3*, but only posterior mesoderm can induce the posterior neural marker *XIHbox6* and the pan-neural marker *XIF6*. Again, the inference was that an additional complementary system is required for the establishment of the complete neural pattern.

The lack of neural development in exogastrulae appeared to be compelling evidence against the existence of a primary planar signal, until molecular analyses were carried out using techniques pioneered by Keller. Keller and Danilchik (1988) excised a strip of dorsal ectoderm from early gastrulae, before the involuting mesoderm had come into contact with the ectoderm, and then cultured them flat either as a single sheet (open face explants) or as two sheets with their their inner surfaces in contact (sandwich explants: figure 1.4). The presumptive a-p axes of the ectoderm and mesoderm point in opposite directions, with their posterior ends at their common boundary, and since mesodermal involution is prevented, the cell movement of convergence-extension extend both parts of the explant. Both types of explant show evidence of neural differentiation in the ectoderm (Wilson and Keller, 1991), which suggests a planar signal, and the sandwich
Chapter 1 - General Introduction

Figure 1.3 (a) Fold experiments of Nieuwkoop et al. (1952a, b) testing planar induction of a-p pattern:
Schematic diagram of a lateral view of a tailbud embryo with folds of competent
ectoderm that have been inserted perpendicularly at different a-p positions. The a-p
neural pattern in each fold is related to its position in the a-p axis of the host; anterior
(4) is always distal, posterior proximal.

(b i) Nieuwkoop and Nigtevecht's (1954) notochord sandwich:
diagram of a cross-section of notochord sandwiched between two sheets of competent
ectoderm; arrows depict hypothetical path of neural inducers.

(b ii) a-p neural pattern found in above sandwiches:
anterior neural tissue (4) is distal, posterior (2) proximal to notochord.

(From, Doniach, 1993)
Chapter 1 - General Introduction

(a) Notochord competent ectoderm

(b) Notochord
   competent ectoderm

(bii)
Figure 1.4 Keller explants:

Regions early gastrula used to make Keller explants. On the left is a stage 10+ embryo in sagittal cross section, dorsal to the right, animal pole up. The shading approximates tissue types predicted by the fate map (Keller, 1975, 1976; Keller et al., 1992):

The animal hemisphere consists of ectoderm that will give rise to epidermis (white) and neuroectoderm (light stippling). Dorsal mesoderm is shown in dark grey and the archenteron roof is stripped.

The layer of endoderm is not shown in the explants depicted on the far right. In these, the white column down the centre of the mesoderm represents the notochord.

Scale bar = 500 μm.

(From, Doniach, 1992)
Chapter 1 - General Introduction

SANDWICH

OPEN FACE

cut x2
explant undergoes the convergence and extension movements that are characteristic of the developing neural plate. It is not known why, but this doesn't occur in the open face explants. Further, it was noticed that the extension and convergence movements of the sandwich explants were also dependent on mesodermal contact (Keller and Danilchik, 1988).

RNAse protection analysis showed that expression of the pan-neural markers \( NCAM \) and \( NF3 \) could be found in the ectoderm in Keller sandwiches only if it was in direct contact with mesoderm (Dixon and Kibrer, 1989). The ventral epidermal marker \( Epi-1 \), which is never expressed in the neural plate, is suppressed in the ventral ectoderm by the planar contact with mesoderm (Savage and Phillips, 1989).

Using molecular markers to analyse the behaviour of the explant, it is evident that the explants generate a pattern of gene expression which reflects the pattern seen \textit{in vivo}. For example, Doniach et al. (1992) used a panel of markers that are expressed in an anterior to posterior pattern (\textit{en-2}, \textit{Krox-20}, \textit{XIHbox1} and \textit{XIHbox6}) to probe both types of explant with \textit{in situ} hybridisation and wholemount immunostaining. They found that in both cases the pattern of the four markers was equivalent to the a-p pattern seen in the embryo. It was also apparent that the extension and convergence movements of the neural plate were not integral to the expression.

What is abundantly clear from these investigations over the last seventy years is that neither vertical nor planar signals on their own provide enough information to construct the embryological nervous system. Vertical signals seem to lack the finesse to focus the correct a-p pattern of neural markers, and planar signals appear unable to induce markers at the anterior limit of the nervous system, including the forebrain. In particular, it has been shown that the floor plate of the neural tube, a structure never detected in
Keller explants, is completely dependent on vertical signals emanating from the notochord (Placzek et al., 1990).

Although the phenomenon of neural induction has been recognised and investigated for many years, it is only relatively recently that credible molecular candidates for the inductive role have emerged, for example, follistatin, noggin and FGF (Ferreiro and Harris, 1994). Interest in follistatin as a neural inducer first arose when it was discovered that the over-expression of truncated activin receptors in *Xenopus* animal caps promoted the formation of neural tissue. However, it was known that this modified receptor also blocks other TGFβ family members such as Vg1 and BMP-4 (Sculle-Merker et al., 1994), therefore the effect of the apparently activin-specific antagonist, follistatin, was similarly tested (Hemmati-Brivanlou and Melton, 1994). Follistatin RNA is found in the embryo in the correct position, the organiser, to be involved in neural induction, and it did replicate the action of the truncated activin receptor when injected into blastula stage ectoderm (Hemmati-Brivanlou and Melton, 1994). But it failed to maintain a neuralising effect through gastrulation as would be expected of the endogenous factor (Mariani and Hemmati-Brivanlou, unpublished data cited in Lamb and Harland, 1995). This does not preclude the possibility of a synergistic relationship with another co-factor.

In addition to a role in the dorsalisation of mesoderm, the secreted protein noggin can directly induce neural tissue of anterior character without first inducing mesoderm (Lamb et al., 1993). This suggests that another signal is necessary to provide the full a-p neural axis. It has recently been suggested that FGF can act synergistically with noggin to make a more complete neural structure (Lamb and Harland, 1995). Embryonic FGF (eFGF) is also expressed in the embryonic tissues with neuralising activity, the organiser and the posterior notochord (Isaacs et al., 1992), and has recently been nominated to
partner noggin in its neuralising activities (Lamb and Harland, 1995). Treatment of blastula and gastrula stage animal caps with bFGF, which accurately mimics the action of eFGF, resulted in the induction of posterior neural markers (Lamb and Harland, 1995). When this was repeated in combination with noggin (Lamb et al., 1993) a more complete a-p neural axis was generated (Lamb and Harland, 1995). Although the generated neural pattern is still deficient in some respects compared to that seen \textit{in vivo}, this does provide compelling evidence for the involvement of these two molecules.

The observations that neural induction occurs when blastula and gastrula ectodermal cells are dissociated in low divalent ion media (Grunz and Tacke, 1989), and that activin is neutralised by follistatin (Hemmati-Brivanlou and Melton, 1994), have added to the opinion that neural induction involves the derepression of a default state (Green, 1994). Recent work has suggested that the neuralising properties of chordin are mediated by an antagonistic relationship with BMP-4 (Sasai et al., 1995), which prevents BMP-4 promoting an epidermal fate in presumptive ectoderm (Wilson and Hemmati-Brivanlou, 1995). This raises the intriguing possibility that these two molecules counteract each other during the induction of both the mesoderm and nervous system, to reveal the default state of the tissues involved.

After the formation of the neural plate, a cascade of pro-neural and neurogenic genes interact to form the nervous system (Chitnes and Kitner, 1995). In many instances evolutionary conserved proteins are active in functionally homologous roles in \textit{Drosophila} and vertebrates, and have thereby contributed to the molecular understanding of the vertebrate equivalent (Kelly and Melton, 1995). For example, \textit{Xash1} is one of several \textit{Xenopus} homologues of the \textit{Drosophila} pro-neural gene \textit{achaete-scute} (Ferreiro et al.,
1993), and XNJ (formerly known as Xotch) is one of the Xenopus homologues of the neurogenic gene Notch.

1.5 The Neural Crest

After the neural plate has completed its convergence and extension movements, the longitudinal neural folds come together at the dorsal midline to form the neural tube (the prospective central nervous system) and separate from the flanking ectoderm which later becomes epidermis. In the chick the extreme margins of the neural plate form a distinctive tissue called the neural crest, which comes to lie on the dorsal aspect of the neural tube. The cells in this tissue eventually migrate away from the neural tube to form the peripheral nervous system and other tissues (Le Douarin, 1982). In Xenopus the tissue which will become the neural crest is evident for many hours before migration, and it also has an apparently more complex organisation than in chick. Anurans, such as Xenopus, possess two layers of neuroectoderm, a feature not seen in other vertebrates (Schroeder, 1970). In the longitudinal neural folds, the superficial cells are the prospective roof of the neural tube, and the deep cells are the prospective neural crest. The region surrounding the neural plate from which the crest cells emerge has been termed the neural plate border (Essex et al., 1993). The cephalic crest is much more developed than the trunk crest and always remains at the side of the neural tube, while the trunk crest lies on top (Schroeder, 1970). The crest cells migrate, initially from the anterior, to various parts of the developing embryo, along paths characteristic of the a-p position along the body axis, to form particular tissues. Their migratory pathways lead to diverse sites within the embryo, where they differentiate into a variety of tissues, including most of the peripheral nervous
system, pigment cells, the adrenal medullary cells, facial cartilage and dorsal fin (Le Douarin, 1982).

The inductive influences that are responsible for the specification of the neural crest remain obscure. When amphibian neural plate is transplanted into non-neural ectoderm, there is induction of crest histological markers at the junction of the tissues (Moury and Jacobson, 1990), and similarly, when young chicken neural plate and non-neural ectoderm are artificially juxtaposed, neural crest markers are expressed (Liem et al., 1995). These data suggest that this tissue is induced on the border of two interacting and distinct domains. In chicks BMP-4 and BMP-7 are candidates for the signal emanating from the non-neural tissue (Dickinson et al., 1995). However, these studies involve the use of mature neural plate and may not reflect the molecular events that occur at neural crest induction, which is presumed to occur simultaneously with that of the neural plate itself (Mayor et al., 1995).

Recently during an expression cloning study designed to isolate molecules involved in the induction of the neural crest, it transpired that the RNA helicase translation initiation factors eIF4AI and eIF4AII were able to upregulate early crest markers without expression of neural plate or mesoderm markers (Morgan, R. and Sargent, M., submitted). Further, these factors also sensitise the responding tissue to noggin, so that the neural plate markers *NCAM* and *Xash* are induced at much lower concentrations. It is also apparent that an increase in the activity of the enzyme protein kinase Ca, which has been implicated in the acquisition of competence for neural induction in *Xenopus* (Otte et al., 1988, 1990, 1991; Otte and Moon, 1992), is a crucial step in this process. When PKCa is specifically inhibited, the induction of neural crest markers by eIF4AI or II is inhibited (Morgan, R. and Sargent, M., submitted).
1.6  *Xsna* and *Xslu*

The *Xsna* gene encodes a putative transcription factor that contains five zinc-finger DNA binding motifs. It is the *Xenopus* homologue of the *Drosophila* gene *snail* and was cloned from a stage 17 *Xenopus* embryo library (Kintner and Melton, 1987) using the *Drosophila snail* cDNA as a probe (Sargent and Bennett, 1990). The cDNA clone is 1.8 kb in length and encodes 259 amino acid residues. The zinc finger regions of *snail* and *Xsna* are 60% homologous, with no significant homology outside this region (Sargent and Bennett, 1990).

The expression pattern of *Xsna* reflects dynamically the important events that occur in early *Xenopus* embryogenesis (Essex et al., 1993). *Xsna* transcripts are localised in the mesoderm and neuroectoderm. The mesodermal expression begins at stage 9 on the dorsal side of the embryo when the cells inside the embryo begin to move towards the animal pole, at a position sometimes referred to as the internal blastopore lip. At stage 10, when the blastopore lip is forming, *Xsna* mRNA begins to appear on the ventral side. Expression of *Xsna* on the ventral side appears to be dependent on a signal emanating from the organiser region (Sargent and Bennett, 1990). The timing of the *Xsna* response indicates that it is not an "immediate early gene" like *Mix-1* (Rosa, 1989) and is indeed regulated by *Xbra* mRNA (Cunliffe and Smith, 1992). By stage 11, when the blastopore is formed, *Xsna* is expressed uniformly in the mesoderm from dorsal to ventral, except for one break seen inside the embryo above the dorsal blastopore lip that corresponds to the prospective notochord. Down regulation of the mesodermal *Xsna* precedes the histological differentiation of prospective notochord, somites and pronephros.

*Xsna* is also expressed in the deep and superficial layers of the neuroectoderm from stage 11, in a subset of cells that surround the neural plate and have therefore been
designated the neural plate border (Essex et al., 1993; Mayor et al., 1993). The deep cells, prospective neural crest, continue to express \textit{Xsna} until after migration, while the superficial cells, the prospective roof of the neural tube, cease expression soon after neurulation. Expression in the transverse neural fold, which becomes forebrain, ceases early in neurulation.

The promoter region of \textit{Xsna} is compact compared with that of \textit{Drosophila snail} (Mayor et al., 1993). It was identified by its capacity to drive the expression of constructs containing the chloramphenicol transferase (CAT) reporter gene, when injected into embryos. The start of transcription was mapped by primer extension to be 33 bases upstream of the translational start site (Mayor et al., 1993). A sequence comparison with a transcription factor database of the region 5' to this site has revealed recognition sequences similar to the binding site for AP4 (-55 relative to the transcriptional start site: Mermod et al., 1988) and the B2 promoter element of TFIIIA (-91: Scotto et al., 1989) as well as sites for SP1 (-129: Briggs et al., 1985) and Oct1 (-142: Mattaj et al., 1985; La Bella et al., 1988) all in reverse of their usual orientation. But there are no TATA or CAAT boxes to be seen at typical sites for these motifs (-32 or -100 respectively, from the start site), and unlike many TATA-less promoters the \textit{Xsna} promoter is not GC rich and does not have SP1 sites near the 3' end (Blake et al., 1990; et al., 1992). In the absence of a TATA box there is usually no precisely determined start sequence (Blake et al., 1990) or, occasionally, there is a specific sequence at "start" that is required for initiation (O'Shea-Greenfield and Smale, 1992; Wiley et al., 1992). However, the \textit{Xsna} promoter initiates transcription at a fixed position, and if the 45 bases upstream of this point are removed, no reduction in expression is seen.
Deletion analysis using β-Gal and CAT constructs containing the promoter region identified specific elements that regulate the spatial and temporal expression of the gene (figure 1.5). When the sequence -112 to -160 relative to the transcriptional start site was removed, both the mesodermal and ectodermal expression decreased by equivalent amounts, possibly indicating the presence of an upregulator element that affected all expression. Removal of the sequence -96 to -112 almost abolished mesodermal expression, but the ectodermal component was maintained until a further truncation down to -93 was made, which almost completely abolished it. An internal deletion, -89 to -92, within the minimal promoter was all that was required to prevent any expression. Consequently, the sequence from -96 to -112 was designated the mesodermal (M) region and from -45 to -96 the ectodermal (E) region. Attempts to produce mesodermal expression alone proved impossible suggesting that the M element is dependent on an aspect of the sequence within the E element, or a factor that associates with it (Mayor et al., 1993).

*Xslu, Xenopus slug,* is a gene closely related to *Xsna,* which encodes a protein containing five zinc-fingers. It was cloned using the chicken *Slug* cDNA as a probe (Mayor et al., 1995). It contains 267 amino acid residues and is very similar to *Xsna* in the first 40 and last 123 amino acids of the zinc finger region. The *Xenopus* and chicken orthologues of Slug show 90 % conservation over their entire lengths (Mayor et al., 1995).

*Xslu* expression, examined by digoxygenin whole mount in situ hybridisation (Mayor et al., 1995), was found only in the ectoderm in a pattern similar to *Xsna.* The expression begins slightly later than *Xsna* at the open neural plate stage (12), in both the superficial and deep layers of the neuroectoderm, with no expression in the transverse...
Figure 1.5 β-gal expression in mesoderm and ectoderm of embryos obtained from fertilised eggs injected with truncated promoter constructs. β-gal constructs injected into embryos are on the left. Frequency of occurrence of expression in mesoderm and neural folds at stage 18 is shown on the right. Embryos were scored as positive if any β-gal positive cells were found in mesoderm or neural folds by dissection.

ID -45/+75, internal deletion. Bar on histogram, s.e.m.

(From, Mayor et. al., 1993)
Chapter 1 - General Introduction

I

A

-160 -112 -96 -93 -89 -45 +41

B

0 50 100%

-160

-112

-96

-93

ID-92/-89

Mesoderm Expression Element
● Neural Fold Expression Element
■ Possible Upregulator Element

Mesoderm
● Neural Fold
folds and only weak expression in the trunk folds. By stage 16, Xslu expression is clearly
organised into the characteristic premigratory aggregates of cranial crest, which become
the branchial arches. Mesodermal expression of Xslu, which is limited to the lateral plate
mesoderm, begins weakly at stage 17 and increases slightly up to stage 26. It might be
reasonable to expect to find similar nucleotide sequences within the promoters of Xsna and
Xslu, that direct their similar ectodermal expression patterns. However, extensive
sequence comparison have yielded no such cis-region (Sargent, unpublished data).

The expression of Xsna and Xslu in the prospective neural crest are the earliest
sign of this tissue, and therefore they are important markers for studying the induction of
this unique cell population (Mayor et al., 1995), and subsequent early events.

1.7 Drosophila snail

In 1978 the recent Nobel Laureates Christiane Nüsslein-Volhard and Eric Wieschaus
began searching the Drosophila genome for mutations that altered the pattern of
embryonic development. The consequences of this mammoth undertaking were not only
the definition of the gap, pair rule and segmentation polarity classes of genes
(Nüsslein-Volhard and Wieschaus, 1980), but also the identification of nineteen maternal
and zygotic genes that are involved in dorsal-ventral (d-v) patterning (Simpson, 1983). A
mutation in one of these loci resulted in the snail phenotype.

The Drosophila snail gene was cloned after a 90 kb chromosome walk along
chromosome 2 (Boulay, et al. 1987). Subsequent analysis revealed an open reading frame
of 1.2 kb that encoded a protein of 390 amino acid residues containing five zinc finger
DNA binding motifs similar to the Xenopus transcription factor TFIIIA (Miller et al.,
1985). This strongly suggested that Snail was itself a transcription factor and was the first
indication of the nature of the interactions that occurred during the establishment of the
d-v axis (Boulay, et al. 1987).

The binding site for Snail was defined using recombinant protein to select specific
binding sequences from a pool of random sequence oligonucleotides (Mauhin et al., 1993).
The consensus sequence was interpreted to be $^{5}$G/A A/t G/A A CAGGTG C/t A C$^{3}$.
The underlined core, which is highly conserved, has no significant homology with any other
*Drosophila* zinc finger transcription factors, but is identical to the core motif of
helix-loop-helix binding site (Mauhin et al., 1993).

Mutational analysis of its locus (Simpson, 1983), indicated that snail was a key
mesoderm determinant in the embryo, responsible for the differentiation of the ventral
mesoderm (Leptin, 1991). It acts by preventing the presumptive mesodermal expression
of genes that are destined to be active only in more lateral or dorsal regions. As a
consequence of repressing the neuroectodermal regulatory genes, the sharp limit of snail
expression assists in the definition of the mesoderm-neuroectoderm boundary (Kosman et
al., 1991). Subsequently, snail expression begins in the developing nervous system, and
appears to be independent of the proneural genes (Ip et al., 1994). Moreover, deletion
analysis of the promoter region responsible for this pan-neural expression, showed that it
was possible to separate the central and peripheral nervous system components from each
other (Ip et al., 1994).

Two other snail related genes have been isolated in *Drosophila*. escargot (esc)
encodes a protein containing 470 amino acid residues, accommodating five zinc finger
motifs, that shared a 76 % amino acid identity with the DNA binding region of Snail. The
eexpression of escargot is primarily ectodermal (Whiteley et al., 1992), while its role
appears to be the control of polyploidy during the growth of the precursors of the insect
appendages, the imaginal discs (Fuse et al., 1994). The normal cell cycle alternates between the S phase (DNA replication) and the M phase (mitosis). However, after the completion of embryogenesis, most of the larval cells in Drosophila miss out the M phase to become large polyploid cells that have a limited ability to differentiate, and are mostly histolyzed during metamorphosis. At the onset of metamorphosis the cells of the imaginal discs replace the larval cells and, along with other diploid cells, form the adult fly (Cohen, 1993). The maintenance of ploidy is therefore essential, and it has recently been shown that Escargot has an intrinsic role in a Cdc2 dependent cell cycle checkpoint, inhibiting the transition from G2 into S phase (Hayashi, 1996).

scratch (*scrt*: Roark et al., 1995) was isolated after a screen for pan-neural expressed genes in which modified transposable elements, P elements containing lacZ, were used to identify enhancers that directed interesting expression patterns (Bier et al., 1989). The sequence of the gene revealed that it coded for a protein of 664 amino acids which contained five zinc finger motifs with a high similarity to Snail. For example, a 76% identity over the most conserved stretch of 46 amino acid. Each zinc finger in Scrt most resembles its counterpart in both the Sna and Esc proteins (Roark et al., 1995). Functional analysis of Scrt revealed that it has an important role in promoting nervous system development by working in concert with Deadpan (Dpn) to repress the expression of non-neuronal genes (Roark et al., 1995).

Snail has been shown to function as a short range repressor (Gray et al., 1994), acting by either directly interfering with the transcriptional machinery, or by quenching the effect of activators whilst occupying neighbouring sites (Gray and Levine, 1996). The type of repression employed depends on the location of its binding site. When located within an upstream enhancer, it quenches nearby activators and allows other enhancers to
interact with the transcriptional machinery. In contrast, when it binds proximal to the promoter, it acts in a dominant fashion blocking all transcriptional activity (Gray and Levine, 1996). This flexibility in the mode of repression allows Snail to establish localised patterns of gene expression. It will be interesting to discover whether the above mentioned Escargot and Scratch function in an analogous manner.

1.8 The snail Family

The snail family (table 1.1) is made up of genes that encode homologous zinc finger transcription factors. Their first four zinc fingers show the canonical DNA binding motif CxxC (12 αα) HxxxxH (El Baradi and Pieler, 1991), but their fifth shows a characteristic variant motif CxxC (12 αα) HxxxxH. Members of this family have been isolated in several phyla including the arthropods, annelids, molluscs and the chordates (Sommer et al., 1992; table 1.1), and therefore, there can be no doubting that this evolutionarily conserved family play an important role in developmental systems.

The known Drosophila family members, snail (Boulay et al., 1987), escargot (Whiteley et al., 1992) and scratch (Roark et al., 1995), have already been described. Using the zinc finger region of snail to probe a Southern blot containing Drosophila genomic DNA, Whiteley et al. (1992) recorded data that there maybe a number of related genes present. This raises an interesting evolutionary question: since it is accepted that gene duplication is more extensive in vertebrates than in flies, (for example, the hox genes (Hunt and Krumlauf, 1992); wingless and the wnt family (Moon, 1993); decapentaplegic and the TGFβ family (Wall and Hogan, 1994)), and there are at least three snail family members in Drosophila, is it reasonable to expect to find several homologues in each vertebrate species?
Table 1.1  The *snail* Family

(References:
1 Boulay et al., 1987: 2 Whiteley et al., 1992: 3 Roark et al., 1995: 4 Sargent and Bennett, 1990:
5 Mayor et al., 1995: 6 Nieto et al., 1992; Smith et al., 1992: 7 Sargent and Nieto, personal
communication: 8 Nieto et al., 1994: 9 A. Isaac, personal communication: 10 Thisse et al., 1993;
Hammerschmidt and Nüsslein-Volhard, 1993: 11, 12, 13 Thisse et al., 1995: 14 Sommer et al.,
1992)
<table>
<thead>
<tr>
<th>Species</th>
<th>Gene</th>
<th>Method of Isolation</th>
<th>Role / Expression</th>
<th>No. of Zinc Fingers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila</td>
<td>snail&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Mutagenic Screen.</td>
<td>Repression of non-mesodermal gene exp. Subsequent exp. in developing nervous sys.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>escargot&lt;sup&gt;2&lt;/sup&gt;</td>
<td>Enhancer trap.</td>
<td>Controls polyploidy in imaginal discs.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>scratch&lt;sup&gt;3&lt;/sup&gt;</td>
<td>Enhancer trap.</td>
<td>Repression of non-neurogenic gene exp.</td>
<td>5</td>
</tr>
<tr>
<td>Xenopus</td>
<td>Xsna&lt;sup&gt;4&lt;/sup&gt;</td>
<td>Library screen with snail probe.</td>
<td>Expressed in mesoderm from stage 9, and in prospective neural crest from stage 11.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Xslug&lt;sup&gt;5&lt;/sup&gt;</td>
<td>Library screen with Slug probe.</td>
<td>Expressed in prospective neural crest from stage 12.</td>
<td>5</td>
</tr>
<tr>
<td>Mouse</td>
<td>Sna&lt;sup&gt;6&lt;/sup&gt;</td>
<td>Library screen with Xsna probe.</td>
<td>Expressed in mesoderm and prospective neural crest.</td>
<td>4 (fingers 2-5)</td>
</tr>
<tr>
<td></td>
<td>Slug&lt;sup&gt;7&lt;/sup&gt;</td>
<td>Library screen with Slug probe.</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Chicken</td>
<td>Snr-1&lt;sup&gt;8&lt;/sup&gt; (snail related-1)</td>
<td>Library screen with Xsna probe.</td>
<td>Expressed in mesoderm and prospective neural crest.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Slug&lt;sup&gt;9&lt;/sup&gt;</td>
<td>Library screen with Xsna probe.</td>
<td>Essential for cell migration.</td>
<td>5</td>
</tr>
<tr>
<td>Zebrafish</td>
<td>snail&lt;sup&gt;10&lt;/sup&gt;</td>
<td>Library screen with snail probe.</td>
<td>Expressed in mesoderm and prospective neural crest.</td>
<td>4 (fingers 2-5)</td>
</tr>
<tr>
<td></td>
<td>snail&lt;sup&gt;11&lt;/sup&gt;</td>
<td>Library screen with snail probe.</td>
<td>Expressed in prospective neural crest.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>pointillé&lt;sup&gt;12&lt;/sup&gt;</td>
<td>Library screen with snail probe.</td>
<td>Expressed in specific subsets of neurons.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>pointillé&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Library screen with snail probe.</td>
<td>Expressed in specific subsets of neurons.</td>
<td>5</td>
</tr>
<tr>
<td>Tribolium</td>
<td>snail&lt;sup&gt;14&lt;/sup&gt;</td>
<td>PCR cloning.</td>
<td>Expressed in mesoderm in a pattern analogous to that of Drosophila snail.</td>
<td>5</td>
</tr>
<tr>
<td>(Flour beetle)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Although *Xsna* was isolated from a stage 17 library using *snail* cDNA as the probe (Sargent and Bennett, 1990), it now transpires that its sequence over the conserved zinc finger region is more similar to *escargot* (67 % rather than 74 % nucleic acid identity). However, *snail* and *Xsna* are both expressed in mesoderm and in the developing nervous system (Boulay et al., 1987; Essex et al., 1993), whereas *escargot* is expressed predominantly in ectoderm (Whiteley et al., 1992). But the question remains as to which *Drosophila* gene is *Xsna* the true orthologue. This may indicate that the full complement of *Drosophila* and *Xenopus* snail related genes have yet to be isolated. Although it may also illustrate evolutionary divergence from a distant common ancestor, during which time the genes have been selected and modified to accomplish diverse results: an arthropod and a vertebrate.

As well as *Xsna* and *Xslu*, *snail* vertebrate homologues with developmental roles have also been isolated in mouse (*Sna*: Neito et al., 1993; Smith et al., 1993: *Slu*: Sargent and Neito, personal communication), chicken (*Slug*: Neito et al., 1994: *Snr-1*: A. Isaac, personal communication), and zebrafish (*snail1*: Hammerschmidt and Nüsslein-Volhard, 1993; Thisse et al., 1993: *snail2*, *pointillé 1* and *pointillé 2*: Thisse et al., 1995). The products of these genes share an additional conserved sequence at the amino terminus, MPRSFLV(K/R)K, which is unique to the vertebrate *snail*-related genes. A brief summary of either their developmental role, or expression pattern, is included in table 1.1.

It may be possible to rationalise apparent differences in the number of genes, and their roles, between the *Drosophila* and *Xenopus* groups of the *snail* family by virtue of evolutionary divergence, but it is harder for this to explain the differences found between the different vertebrates, as these species are more closely related. The most obvious difference is that four homologues have been isolated in zebrafish and only two in the
other vertebrate species. However, *pointillé*1 and *pointillé*2 are the most divergent of the
*snail* family and their expression is only found in specific subsets of neurons (Thisse et al.,
1995). It is therefore possible that their orthologues have not been isolated for technical
rather than biological reasons. There are more subtle differences that need explaining. For
example, mirroring the relationship between *snail*, *esc* and *Xsna*, *snail*2 from zebrafish is
more similar at the nucleotide level to *Xsna* and the chicken *Slug* than to *Sna* from mouse;
but *snail*2 has no mesodermal expression, unlike the other three which show analogous
expression patterns.

Unlike the other family members, *snail*1 from zebrafish and the murine *Sna* lack the
first out of the five characteristic zinc fingers. This has led to the suggestion that there are
two sub-groups within the vertebrate *snail* family based on the number of encoded zinc
fingers (Thisse et al., 1993, 1995). This would currently result in one group containing
eight members and the other only two; with zebrafish being the only species with a
member in each group (table 1.1). The imbalance is explained by the assumption that not
all of the vertebrate genes have yet been described (Thisse et al., 1995).

The attempt to deduce the relationship between these genes highlights the problem
in determining the best criteria on which to base the decision. It is clear that organising the
genes of this family into sub-groups based on structural properties ignores similarities and
differences, in their expression patterns. It may be better in the short term, to look upon
the *snail*-related genes as a family of putative specialised short range regulators, whose
adventiveness has been taken advantage of by several developmental processes, and in the
process they have diverged in order to better fulfil their roles. Only after further
investigations into the various *snail* related genes will a more definitive classification be
able to be justified.
1.9 Establishment of the Dorsal-Ventral axis in Drosophila

The foundations of the Drosophila embryonic body plan are laid down by the mother while the oocyte is surrounded by 15 nurse cells, which are all interlinked via cytoplasmic channels that allow the transfer of intracellular materials. These 16 cells are encapsulated within a bag of 1600 follicle cells. Four determining systems are initiated, each of which is responsible for a cascade of molecular events that focus the body plan so that each cell within the embryo is aware of its developmental fate. The activation or repression of a gene is determined by the concentration and specific mix of transcription factors at that precise point in the embryo, as well as, possibly, by the binding affinity between the gene and these factors.

The anterior-posterior (a-p) system is defined by the localisation of maternal bicoid RNA at the anterior pole of the embryo. This is translated by the zygotic machinery which results in a gradient of the protein along the a-p axis. The posterior system allows for the formation of the abdomen. It is co-ordinated by the nanos gene product, which functions by countering the anteriorising effects of bicoid, and a mad protein called hunchback; that is, it preserves the default posterior state of the embryo. The terminal system defines the head and the tail of the embryo. It depends on the localised action of the receptor protein torso which, although found uniformly around the embryo, is only active at the poles.

The dorsal-ventral (d-v) system is controlled by Dorsal, a member of the rel family of transcription factors which also includes NF-κB. Dorsal protein is distributed uniformly in the oocyte cytoplasm where it is held by Cactus, a protein that contains multiple ankyrin repeats found in all members of the IkB family (Kidd 1992). The Dorsal-Cactus interaction occurs through the rel-homology domain masking the nuclear localisation
signal (Geisler et al., 1992). Nuclear localisation of Dorsal across the d-v axis is generated in response to a signal mediated by the membrane bound protein Toll, a receptor activated along the ventral midline. The cytoplasmic domain of Toll has homology to the intracellular domain of the interleukin-1 receptor which is the mediator of NF-κB. The proposed mechanism of action of Dorsal is to activate both snail and twist expression. Twist is thought to be responsible for activating and maintaining the expression of the early mesoderm specific genes, including itself and snail, while, as described above, the mesodermal role of Snail is to repress the expression of non-mesoderm specific genes (Leptin, 1991).

2.0 Xsna and snail

When considering a vertebrate gene which has a relatively well defined homologue in Drosophila, the possibility arises that their control mechanisms may also be homologous. Three Xenopus members of the c-rel family of transcription factors, to which dorsal and NFκB also belong, have been reported (XrelA.1 and XrelA.2: Richardson et al., 1994: Xrell: Kao and Hopwood, 1991). The expression pattern of XrelA suggests that it has a role in the patterning of the embryo termini rather than in the dorsal-ventral system (Richardson et al., 1995). Transcripts of Xrell are wide-spread in the embryo which would seem to preclude a role in specifying a discreet region, unless a similar mechanism to that of the dorsal-cactus interaction is also conserved (Kao and Hopwood, 1991).

A Xenopus homologue of Drosophila twist has been isolated (Hopwood et al., 1989). Xtwi encodes a putative transcription factor that contains the basic helix-loop-helix (bHLH) DNA binding domain similar to twist. Although it is expressed in the mesoderm and prospective neural crest, its expression pattern differs from Xsna in that it appears only
in somatic mesoderm somewhat later, and in the neural crest in significant amounts only at stage 15 (Hopwood et al., 1989), although traces are present at stage 12 (Essex et al., 1993).

Snail, Escargot and Scratch all seem to act by repressing inappropriate gene expression (Leptin, 1991; Fuse et al., 1994; Roark et al., 1995). Scratch works in concert with bHLH activator proteins (genes of the *achaete-scute* complex: Roark et al., 1995), in the same manner that Snail and Twist operate together (Leptin, 1991). It is appealing to speculate that *Xsna* is also a short range repressor that performs its role in conjunction with a bHLH activator, possibly *Xtwi*, under the direction of a *c-rel* family member.

### 2.1 A Brief Overview of Eukaryotic Transcription

Each eukaryotic cell contains the same genome, but each cell type utilises, at precise times, only a subset of the available genes to produce a very specific mixture of proteins. This differential gene activation involves the regulated assembly of multi-protein complexes on enhancers and promoters. It would be impossible to review such a vast subject, but it is possible to highlight some of the major milestones in its development.

Biochemical efforts to understand eukaryotic transcription began with the discovery of a mammalian nuclear RNA polymerase activity (Weiss and Gladstone, 1959), and the subsequent purification of three distinct classes of DNA dependent RNA polymerases (Roeder and Rutter, 1969); RNA Polymerases II being responsible for the transcription of all cytoplasmic mRNA. It was soon realised that other factors were required to initiate this process, and largely through the efforts of Roeder and colleagues, these general transcription factors were purified by chromatographic fractionation.
Assembly of these factors \textit{in vitro} produced a low invariable level of transcription, but it was apparent that this basal complex did not regulate its own activity (Buratowski, 1994).

Examination of the sequences surrounding the coding regions by genetic and biochemical methods showed that distinct cis-elements controlled transcription. Closest to the coding region was the core promoter, which often contained elements such as TATA and CCAAT boxes, which seemed to determine the initiation site. It was then discovered that elements, called enhancers and silencers, could influence the level of transcription from sites positioned several thousand nucleotides, either upstream or downstream, from the core promoter (Serfling et al., 1985). Specific transcription factors termed activators and repressors, could bind to these sites and dictate the level of gene transcription, providing a complex combinatorial control over the specificity of gene.

The first human transcription factor to be isolated was specificity protein 1 (SP1; Dynan and Tijan, 1983). As this protein was known to bind specifically to a GC rich region of DNA it was purified by affinity chromatography on chemically synthesised GC rich DNA anchored to a solid matrix. When crude cell extract was passed over the matrix, the SP1 bound to the DNA, allowing it to be concentrated and freed from contaminants (Kadonaga and Tijan, 1986). This pioneering DNA affinity chromatographic procedure has since become a powerful tool in the purification of DNA binding proteins. SP1 contained a zinc finger DNA binding domain (see below), and a glutamine rich activation domain which was later shown to interact with the basal complex through a class of proteins known as co-activators (Dynlacht et al., 1991). This was one of the first descriptions of what appears to be a universal mode of eukaryotic gene regulation (Tijan, 1995).
Shortly before the identification of SP1 the structure of an RNA polymerase III transcription factor from *Xenopus* called TFIIIA was elucidated (Miller et al., 1985). This protein contained a repeated motif (nine copies) that forms finger-like structures, stabilised by the interaction of a zinc ion with two cysteine and two histidine residues, through which the protein could interact with DNA. When it was realised that other factors, including SP1, also contained this motif, it became apparent that this was a major class of DNA binding proteins. Subsequently, several other types of DNA binding motifs have come to light including, the homeodomain, leucine-zipper associated domains and the basic helix-loop-helix domain (Mitchell and Tijan, 1989).

Studies of the yeast transcription factors GAL4 and GCN4 (Ptashne, 1987) demonstrated for the first time the existence of a defined transcriptional activation region, and suggested this class of proteins may be organised into distinct domains with particular functions. It was discovered that the activation domains of these proteins could be exchanged and paired with different DNA binding domains, to produce chimeric transcription factors. This led to the development of the yeast two-hybrid system as an invaluable tool for studying processes which depend on two components, and indeed for identifying hitherto unknown partners in an interaction.

Strict regulation of transcription is essential to ensure that gene expression is promptly initiated when required, and also to avoid over expression of a particular gene. There are many examples where a cascade of gene transcription unfolds after the expression of one gene produces the activator for another, which in turn encodes for another, as in the anterior-posterior system in *Drosophila* (Lawrence, 1992). However, this kind of process requires de novo protein synthesis before a gene may respond to a stimulus. It is now clear that a number of other ways of regulating transcription occur.
Post-translational modifications, such as phosphorylation and dephosphorylation, can change the activity of a transcription factor very rapidly (Berk, 1989). In some cases factors are sequestrated in the cytoplasm by an association with another protein. For example, steroid receptors associate with the heat shock protein hsp90 (Tsai and O'Malley, 1994), and the transcription factor NF-κB is paired with the inhibitor IkB (Liou and Baltimore, 1993). External stimuli will induce the release of these factors, allowing them to enter the nucleus and activate their target genes (Johnson and McKnight, 1989). A substantial component of transcriptional regulation is increasingly perceived to depend on the interplay of transcription factors and chromatin structure (Wolfe, 1994a). Selective wrapping of specific DNA sequences around histones has been shown to have a key role in modulating gene expression (Wolfe, 1994b).

An extra level of regulation that plays a key role in establishing localised patterns of gene expression is provided by transcriptional repressors. For example, the activity of snail in Drosophila development (Leptin, 1991). Three modes of transcriptional repression have been described (Levine and Manley, 1989): repression by competition involves the prevention of the activator-DNA interaction by a repressor interfering with the binding domain; direct repression involves interference with the formation or activity of the basal transcription complex; and finally, the activity of an activator may be masked, or quenched, by a repressor occupying a neighbouring binding site (Levine and Manley, 1989). Recent work involving the repressive function of Snail has suggested that a key distinction between repressors is their range of their action, with short-range repressors such as Snail providing the definitive level of control (Gray and Levine, 1996).
2.2 Project Aims

*Xsna* is a zinc finger transcription factor that is expressed during *Xenopus* development in the mesoderm from stage 9 and in the prospective neural crest from stage 11. Two contiguous elements within the promoter, approximately 100 base pairs from the transcriptional start site, appear to be involved in directing the specificity of the temporal and spatial expression. The *Drosophila* homologues of this gene have been shown to act as negative regulators in several important developmental pathways, and so it is likely that *Xsna* has a major role in the correct development of mesoderm and neural crest.

Therefore, the aim of this project was to gain further insight into the mechanism through which *Xsna* is regulated in these two populations of cells, at separate but precisely defined times.

The starting point for this investigation was to test and optimise the electrophoretic mobility shift assay (EMSA), a powerful method with which to detect protein-DNA interactions, so that the binding activities could be reliably monitored. The EMSA could then be used, in conjunction with protection assays, to further define the important sequences within the two control elements using radiolabelled probes with altered base sequences. This information could then be used to design the most effective DNA-affinity purification matrices. It was planned to investigate biochemical properties of these binding activities with the aim of identifying differences that may reflect, or be responsible for, the regulation of *Xsna* transcription at crucial phases of development. By studying properties such as charge, size and substrate affinity, it was anticipated that a purification protocol could be devised to isolate sufficient material to obtain an N-terminal sequence.
Materials and Methods

2.1 Chemicals, Buffers and Media 46
2.2 Unfertilised Eggs and Embryos 49
2.3 DNA Manipulations 50
2.4 Synthesis of Radiolabelled Probes 52
2.5 Protein Purification 59
2.6 Characterisation of Proteins 62
2.7 RNA Manipulations 70
2.1 Chemicals, Buffers and Media

Unless stated, all chemicals were obtained from Sigma Chemical Company.

2.1.1 General

**Elution Buffer** - 0.1% (w/v) SDS; 10 mM MgCl₂; 0.5 M NH₄Ac

**20x TBE (1 litre)** - 121.1 g Tris; 61.83 g Borate; 18.6 g EDTA

2.1.2 Unfertilised Eggs and Embryos

**Cocktail of protease inhibitors** - 2 μg/ml Aprotinin; 1 ng/ml Cystatin; 10 μM E64; 100 μM Leupeptin; 1 mM amidino phenyl methyl sulphonyl fluoride (Pefabloc, Pentapharm AG); 1 mM phenyl methyl sulphonyl fluoride

**Homogenising Buffer** - 50 mM Tris-HCl, pH 7.5; 50 mM KCl; 2.5 mM MgCl₂; (0.5 mM DTT, unless using cleavable probes)

2.1.3 DNA Manipulations

**4x EMSA Buffer** - 50 mM Tris-HCl, pH 7.5; 50 mM KCl; 6% (v/v) glycerol

**EMSA Loading buffer** - 1 in 200 dilution of load buffer (see below)

**Copper-Phenanthroline Protection Assay Solutions:**

- **OP Solution** - 40 mM 1,10-phenanthroline (80 mg of 1,10-phenanthroline monohydrate dissolved in 10 ml of absolute ethanol)
- **Cu²⁺ Solution** - 9 mM CuSO₄ (72 mg of anhydrous copper (II) sulphate dissolved in 50 ml of dH₂O)
- **[OP]₂Cu²⁺** - 2 ml of OP solution mixed with 2 ml of Cu²⁺ solution, made up to 40 ml with dH₂O
- **MPA Solution** - 58 mM 3-mercaptopropionic acid
- **STOP Solution** - 28 mM 2,9-dimethyl-1,10-phenanthroline (Neocuproine) in absolute ethanol

**Denaturing Polyacrylamide** - 10% (w/v) 19:1 acrylamide : methylene bisacrylamide in urea

**Load Buffer** - 2% (w/v) Bromophenol Blue; 10 mM EDTA; 40% (w/v) Sucrose
10x Mg buffer - 15 mM MgCl₂; 2.5 mM DTT; 8 mM EDTA

Polyacrylamide - 40% (w/v) 19:1 acrylamide : bisacrylamide

**Sequencing Solutions:**

- **5x Sequencing Buffer** - 200 mM Tris.HCl, pH 7.5; 100 mM MgCl₂; 250 mM NaCl
- **Labelling mix** - 0.5 μM each of dCTP, dGTP and dTTP
- **Termination mix** - 80μM dTTP, dCTP, dGTP, dATP, 50 mM NaCl and 10 μM of a single dideoxynucleotide
- **STOP Buffer** - 0.05 % (w/v) Bromophenol Blue; 0.05 % (w/v) Xylene; 20 mM EDTA; 45 % (v/v) Formamide

**2.1.4 Protein Manipulations**

- **Coomasie Stain** - 0.2 % (w/v) Coomasie Blue in 40 % (v/v) ethanol and 10 % (v/v) acetic acid
- **Coomasie De-stain** - 10 % (v/v) acetic acid and 5 % (v/v) methanol
- **PPO/DMSO Solution** - 22 % (w/v) solution of 2,5-diphenyloxazole in dimethyl sulphoxide

**SDS-PAGE** (Laemmli, 1970):

- **Resolving Gel Mixture** - 10 % (w/v) 19:1 acrylamide : bisacrylamide stock solution;
  3 M Tris-HCl, pH 8.9; 0.1 % (w/v) SDS; 0.25 % (v/v) TEMED; 0.5 % (w/v) APS
- **Stacking Gel mixture** - 10 % (w/v) 19:1 acrylamide : bisacrylamide stock solution;
  0.5 M Tris-HCl, pH 6.7; 0.1 % (w/v) SDS; 0.25 % (v/v) TEMED; 0.5 % (w/v) APS
- **Loading Buffer** - 2 % (w/v) Bromophenol Blue; 10 mM EDTA; 40 % (w/v)
  Sucrose; 1 % (w/v) SDS; 1 mM DTT

**Running Buffer (1 litre)** - 6 g Tris-HCl, pH 8.3; 28.8 g Glycine; 0.1 % (w/v) SDS
Chapter 2 - Materials and Methods

Silver Stain (Wray et al., 1981)

Silver Staining Solution - Add 1 ml of 2 M NaOH, 1.4 ml of 14.8 M NH₄OH and 1 ml of 4.7 M AgNO₃ to 100 ml of dH₂O

Developer - Add 1 ml of 2.5 % (w/v) citrate and 0.25 ml of 38 % (v/v) formaldehyde to 500 ml of dH₂O

2.1.5 RNA Manipulations

DNA digestion mix - 1x One Phor All buffer (Pharmacia); 1 u DNAse I (Cambio)

Embryo Homogenisation Buffer (RNA) - 1x TEN; 1 % (w/v) SDS; 1 % (v/v) β-mercaptoethanol

10x MOPS buffer - 4 M NaCl; 0.4 M MOPS, pH 6.4; 10 mM EDTA

RNase digestion buffer - 10 mM Tris-HCl, pH 7.5; 1 mM EDTA; 10 mM MgCl₂; 300 mM NaCl

STOP/LOAD Buffer - 0.05 % (w/v) Bromophenol Blue; 0.05 % (w/v) Xylene;

20 mM EDTA; 45 % (v/v) Formamide

TE/S - 10 mM Tris-HCl, pH 8.0; 1 mM EDTA; 0.1 % (w/v) SDS

-U NTPs - 10 mM ATP, CTP, GTP
2.2 Unfertilised Eggs and Embryos

2.2.1 Collection of Unfertilised Eggs and Embryos

Adult females were primed by injection of 800 units of human gonadotrophin (HCG) 12 hours before unfertilised eggs were required. The eggs were expelled by gentle peristalsis of the females' ventrolateral surface, transferred into a petri dish and were fertilised by rubbing them with the testis of a sacrificed male (Smith and Slack, 1983). After 5 min. they were flooded in 10 % (v/v) Normal Amphibian Medium (NAM; Slack, 1984). The efficiency of the fertilisation was judged by the number of embryos that had rotated so that the animal hemisphere was uppermost. Embryos were dejellied using 2 % (w/v) cysteine hydrochloride at pH 7.8 - 8.1. Large amounts of unfertilised eggs were obtained by allowing gravid females to lay their eggs in 1x NAM supplemented with 1 mM CaCl$_2$ and 1 mM MgCl$_2$. These were separated from the mothers every 2-3 hours and were sorted manually to select perfect specimens. They were frozen in batches of up to 1 ml at -70 °C.

2.2.2 Preparation of Whole Cell Extracts

After collection, embryos, staged according to Nieuwkoop and Faber (1967), or unfertilised eggs (ufe), were homogenised at room temperature in 50 mM Tris-HCl, pH 7.5, 50 mM KCl, 0.5 mM DTT and 2.5 mM MgCl$_2$, using four volumes of homogenising buffer to one volume of embryos/ufe in the presence of a cocktail of protease inhibitors (described above). The crude extract was cleared by centrifugation at 10,000 g for 10 min.

2.2.3 Microinjection of Xenopus Embryos

Dejellied embryos were transferred into a petri dish lined with 1.2 % (w/v) agarose containing 75 % (v/v) NAM (Slack, 1984) supplemented with 5 % (w/v) ficoll. Embryos were injected at the one to four cell stage using an air-driven injection system.
(Inject+matic). Typically volumes of 10 nl were delivered into the embryo. The volume was calculated by injecting the liquid into oil and measuring the diameter of the drop using a graticule. The volume was determined by treating the drop as a sphere.

2.3 DNA Manipulations

2.3.1 Gel Electrophoresis of DNA

2.3.1.1 Non-denaturing polyacrylamide gel electrophoresis

6 - 12 % (w/v) polyacrylamide gels were made up from a 40 % (w/v) 19:1 acrylamide : bisacrylamide stock solution (NBL), and 20x TBE. The polyacrylamide/TBE mixture containing 25 \( \mu l \) TEMED and 250 \( \mu l \) 10% (w/v) APS, was poured between two scrupulously clean glass plates. The gel was run at 200 V in 1x TBE until the DNA Fragments were resolved.

2.3.1.2 Denaturing polyacrylamide gel electrophoresis

6 % (w/v) polyacrylamide gels were made up using the Sequagel stock solutions supplied by National Diagnostics. For a 30 ml gel, 7.2 ml of concentrate (25% (w/v) 19:1 acrylamide : methylene bisacrylamide in 8.3 M urea) was mixed with 21.3 ml of diluent (8.3 M urea) and 1.5 ml of 20x TBE. Polymerisation of the mixture was initiated with the addition of 25 \( \mu l \) TEMED and 250 \( \mu l \) 10% (w/v) APS, and then it was poured between two scrupulously clean glass plates. These gels were run at 33 W in 1x TBE for 1 - 2 hours.

2.3.1.3 Agarose gel electrophoresis

10 \( \mu l \) DNA samples mixed with 2 \( \mu l \) of loading buffer (40% (w/v) sucrose, 2% (w/v) bromophenol blue, 10 mM EDTA) were resolved on 1 - 2 % (w/v) agarose TBE gels run at 100 V in 1x TBE. \( \Phi X \) plasmid DNA cut with Hae III (Boehringer Mannheim) and/or \( \lambda \)
DNA (NEB) cut with BstE II (NBL) were routinely run alongside the samples as size standards on all gels.

### 2.3.2 Polymerase Chain Reaction

Amplification of specific DNA fragments by PCR was performed using 200 ng of each oligonucleotide primer, 1 U of AmpliTaq DNA Polymerase (Perkin Elmer) and 0.2 mM dNTPs in PARR buffer (Cambio Ltd.) to a final volume of 100 μl. The samples were overlaid with 30 μl of mineral oil (Sigma) before being placed in a thermal cycler and exposed to the following conditions: 30 cycles of 1 min. at 92 °C, 1 min. at ~60 °C (variable primer annealing temperatures) and 1 min. at 74 °C (10 min. for the last cycle). The PCR products were resolved on a 2% (w/v) agarose TBE gel and isolated using the Qiaex agarose gel extraction kit (Qiagen).

### 2.3.3 dsDNA Sequencing

The method used was based on the dideoxy-mediated chain termination method (Sanger et al., 1977). Briefly, 1 μg of DNA was incubated with 0.2 M NaOH in a volume of 10 μl for 5 min, before being precipitated by the addition of 50 μl of ethanol and 3 μl of 7 M ammonium acetate at 0 °C. The DNA was dried under vacuum and then rehydrated in 7 μl of dH₂O and 2 μl of 5x sequencing buffer (200 mM Tris.HCl, pH 7.5, 100 mM MgCl₂, 250 mM NaCl). 1 μl of primer (OD 0.2) was added and, after the mixture was heated to 65 °C for 5 min, it was then allowed to cool to <37 °C to promote the annealing of the primer to the template DNA. Meanwhile, 2.5 μl of each termination mix (80 μM dTTP, dCTP, dGTP, dATP, 50 mM NaCl and 10 μM of a single dideoxynucleotide) was placed into separate labelled wells of a microtitre plate, and the T7 DNA polymerase (Pharmacia) was diluted 1:8 in T7 dilution buffer (Pharmacia). The sequencing reaction was then initiated by adding 2 μl of the labelling mix (0.5 μM each of dCTP, dGTP and dTTP), 1 μl 0.1 M
DTT, 0.5 µl [α-32P] dATP and 2 µl of the diluted T7 DNA polymerase, and left at room temperature for 5 min. The reaction was terminated by aliquoting 3.5 µl into each pre-prepared termination mix and incubating at 37 °C for 5 min, after which 4 µl of STOP buffer was added. Following 2 min at 80 °C, the sequence was resolved on a 6% denaturing polyacrylamide gel at 33 W until the bromophenol blue reached the bottom of the gel. The gel was fixed in 10% (v/v) glacial acetic acid for 10 min and 20% (v/v) ethanol for 15 min, dried and exposed to XAR Kodak autoradiographic film overnight at room temperature.

2.4 Synthesis of Radiolabelled Probes

A summary of the probes used during this investigation is given in appendix A.

2.4.1 EMSA Probes

Oligonucleotides were synthesised using a Beckman Oligo 1000 DNA Synthesiser by the Sequencing and Synthesis Service, NIMR. The 5' end of these oligonucleotides were radiolabelled with [γ-32P] dATP (Amersham) using T4 polynucleotide kinase (Gibco-BRL) at 37 °C for 10 - 60 min. in a reaction mixture containing the manufacturers recommended 5x forward reaction buffer (Gibco-BRL). The reaction was terminated by incubating at 65 °C for 10 min. An equal amount of the non-radioactive complementary oligonucleotide was added and, after the mixture was heated to 80 °C for 5 min, it was then allowed to cool to <37 °C to promote the annealing of the two strands of DNA. The labelled probes were gel purified by electrophoresis on a 12% (w/v) non-denaturing polyacrylamide gel, which was then exposed to x-ray film for 2 min. From the resulting autoradiograph it was possible to determine whether the proper duplex had been formed and to excise the probe. Elution from the gel slice was carried out overnight at 55 °C in the elution buffer and the
eluate was extracted with phenol/chloroform (1:1 v/v), to ensure the removal of any polyacrylamide particles. The DNA was ethanol precipitated, dried and rehydrated in dH₂O.

2.4.2 Copper-Phenanthroline Protection Assay Probes

The region of DNA containing the sequence under investigation was amplified from plasmid DNA by PCR using two flanking primers. The primer sites were situated 100 base pairs upstream on both strands. Unlike the standard PCR protocol described above one of the primers was radiolabelled. The result of this was that the amplified DNA fragments were only labelled on one strand. 500 ng of the primers (synthetic oligonucleotides) were radiolabelled with [γ-³²P] dATP (Amersham) using T4 polynucleotide kinase (Gibco-BRL) as described for the EMSA probe above. After the reaction had been terminated the primers were gel purified by electrophoresis on a 12% (w/v) non-denaturing polyacrylamide gel, as described above. This procedure yields approximately 200 ng of the primers. The PCR product was resolved on a 2% (w/v) agarose gel in TBE buffer, visualised in UV light after staining in ethidium bromide, excised and recovered using the Qiaex agarose gel extraction kit (Qiagen).

2.4.3 Biotinylated Probes

2.4.3.1 Non-cleavable Biotinylated Probes

A DNA molecule containing the putative DNA binding sequence of interest was made by annealing two short (< 20 bp) complementary oligonucleotides together in the manner described for the EMSA probes. These molecules of DNA possessed a four base overhang at each end which allowed them to be concatemerised using a modified version of the method described by Kadonga (1991).
Figure 2.1  Biotinylated Probes.
(a) **Biotinylated Probes**

(i) **Non-cleavable Biotinylated Probes**

A

Biotin-16-U

(ii) **Cleavable Biotinylated Probes**

Biotin-ss-NH₂ → NH₂

**Key:**

--- = Oligonucleotides

--- = Region filled in with Klemow (see text)

(b) **Biotinylated UV Probes**

(i) **Mesodermal Probe**

Biotin-ss-NH₂ → NH₂

(ii) **Ectodermal Probe**

Biotin-ss-NH₂ → NH₂

**Key:**

Thick line = Oligonucleotides

Thin line = Region filled in with T7 DNA Polymerase (see text)

- = Putative "Ectodermal" Binding Site

- = Putative "Mesodermal" Binding Site
35 µg of DNA monomers were phosphorylated using 50 u of T4 polynucleotide kinase (Gibco BRL) in the presence of 20 mM ATP, and the kinase/ligase buffer (50 mM Tris (HCl), pH 7.5, 10 mM MgCl₂, 5 mM DTT, 0.1 mM spermidine and 0.1 mM EDTA), at 37 °C for 2 hours. The reaction mixture was heated to 65 °C for 15 min. to inactivate the kinase and then allowed to cool. 15 u of ligase (Boehringer Mannheim) was then added along with an extra 4 mM of spermidine and the reaction was left at 14 °C overnight. The success of the concatemerisation ( ~ 20’mer) was assessed by running 5 µl of the reaction on a 1% (w/v) agarose TBE gel electrophoresis.

Biotin was incorporated into the ends of the concatemer using biotin-16-UTP (Boehringer Mannheim), 20 mM -T dNTP mix, and 15u of the Klenow fragment of DNA polymerase I at 14 °C for 1 hour (refer to figure 2.1ai). The efficiency of the biotinylation was assessed by 1% (w/v) agarose TBE gel electrophoresis of two aliquots, one loaded directly as a control and the other loaded after mixing with 10 µg of streptavidin. The streptavidin bound to the biotinylated molecules retarded migration of the modified molecules. When the reaction was complete, the free biotin was removed using a G-50 Sephadex (Pharmacia) column and, following ethanol precipitation, washing in 70 % (v/v) ethanol and drying under vacuum, the probe was then rehydrated in dH₂O. Finally, the concatemerised DNA was coupled to streptavidin coated streptavidin coated magnetic beads (Dynalbeads M-280 Streptavidin, Dynal) that had been prewashed in 1 % (w/v) Triton and 0.25 M NaCl.

2.4.3.2 Cleavable Biotinylated Probes
A DNA molecule containing the putative DNA binding sequence of interest was made by annealing two complementary oligonucleotides together in the manner described for the EMSA probes. The lower strand was shorter than its complement resulting in a 4 base
overhang which was filled in with 20 μM -A dNTP mix and [α-32P] dATP (Amersham) using 15μ of the Klenow fragment of DNA polymerase I at 14 °C for 1 hour (refer to figure 2.1aii). The radiolabelled probe was purified on a G-50 Sephadex (Pharmacia) column before being ethanol precipitated, washed in 70 % (v/v) ethanol and dried under vacuum.

The upper strand was modified at the 5' end by the incorporation of a 12 carbon spacer arm and a primary amine, which is able to react with sulfosuccinimidyl-2-(biotinamido) ethyl-1,3-dithiopropioate (NHS-ss-Biotin: Pierce) to form a stable amide bond. In preparation for the reaction, 1 mg NHS-ss-Biotin was dissolved in 150 μl 0.1 M borate buffer, pH 8.5. After the radiolabelled probe was rehydrated in dH2O, borate buffer, pH 8.5, was added to a final concentration of 0.1 M, followed by 2 μl of the prepared NHS-ss-Biotin. The reaction was left at 37 °C overnight after which the biotinylated radiolabelled probe was purified on a G-50 Sephadex (Pharmacia) column. Following ethanol precipitation, washing in 70 % (v/v) ethanol and drying under vacuum, the probe was then rehydrated in dH2O.

2.4.4 UV Crosslinking Probes
2.4.4.1 Non-biotinylated Probes
The region of DNA containing the sequence under investigation was amplified from plasmid DNA by PCR using two flanking primers. Unlike the standard PCR protocol described above 0.2 mM Br-dUTP (Boehringer Mannheim) was used in place of dTTP, and [α-32P] dATP (Amersham) instead of dATP. The result of this was that the amplified DNA fragments were radiolabelled and had Br-deoxyuridine incorporated in place of thymidine across the sequence of interest. The PCR product was resolved on a 2% (w/v)
agarose gel in TBE buffer, visualised in UV light after staining in ethidium bromide, 
excised and recovered using the Qiaex agarose gel extraction kit (Qiagen).

2.4.4.2 Biotinylated Probes

A DNA molecule containing the putative DNA binding sequences of interest was made by 
annealing two complementary oligonucleotides together in the manner described for the 
EMSA probes. The upper strand was modified at the 5' end by the incorporation of a 12 
carbon spacer arm and a primary amine, which is able to react with NHS-ss-Biotin 
(N-Hydroxysuccinimide: Pierce) to form a stable amide bond. The probes were 
biotinylated as described above (see section 2.4.3.2).

The mesodermal probe was made up of a long modified oligonucleotide, that 
contained both the mesodermal and ectodermal regions, and a shorter oligonucleotide that 
contained only ectodermal region (refer to figure 2.1bi). The overhang was filled in with 
20 mM CTP, GTP, Br-dUTP (Boehringer Mannheim) and 20 μCi [α-32P] dATP 
(Amersham) using T7 DNA polymerase in 5x sequencing buffer (200 mM Tris.HCl, pH 
7.5, 100 mM MgCl2, 250 mM NaCl) at 37 °C. This resulted in only the mesodermal 
region containing bromodeoxyuridine (BrdU).

The ectodermal probe was made up of a short modified oligonucleotide, that 
contained only the mesodermal region, and a longer oligonucleotide that contained both 
regions (refer to figure 2.1bii). When the overhang was filled in as described above the 
result was that only the ectodermal region contained BrdU.

2.4.5 RNAse Protection Probes

Linear template DNA (1 μg) was incubated for 2 hours at 37 °C with 1 μl DTT (100 mM), 
1 μl -U NTPs (10 mM ATP, CTP, GTP), 2.0 μl 5x Transcription buffer (Cambio), 0.5 μl 
RNasin (RNAse inhibitor: Cambio), 1 μl of RNA polymerase and 3.5 μl of [α-32P]UTP
(Amersham). If EF1α probe was being made, only 1 μl [α-32P] UTP (Amersham) was used with 2.5 μl UTP (20 μM). 5 μl of STOP/LOAD buffer (0.05 % (w/v) Bromophenol Blue; 0.05 % (w/v) Xylene; 20 mM) was added and the radiolabelled RNA probes were gel purified by electrophoresis on a 6 % (w/v) denaturing polyacrylamide gel, run at 25 W until the bromophenol blue had reached the bottom. This was then exposed to x-ray film for 1 min. and the band cut out. The probe was eluted from the gel slice in TE/S (10 mM Tris-HCl, pH 8.0; 1 mM EDTA; 0.1 % (w/v) SDS) by first freezing on dry ice and then incubating at 50 °C overnight.

2.5 Protein Purification

2.5.1 Ammonium Sulphate Precipitation

The clarified embryo or unfertilised egg cell extract was brought up to an ammonium sulphate concentration of 1.36 M (40 % saturation) by the addition of the solid salt. This was incubated on ice for 15 min. and then centrifuged at 10,000 g for 10 min. to separate the precipitated protein from the active supernatant.

2.5.2 Hydrophobic Interaction Chromatography

A 250 μl aliquot taken from the supernatant after an ammonium sulphate precipitation, was carefully layered onto a 1 ml phenyl sepharose (Pharamcia) column pre-equilibrated in homogenisation buffer containing 50 % (w/v) ammonium sulphate. After the matrix was washed with four column volumes of this buffer, a stepwise elution was performed with the sequential addition of 1 ml aliquots of buffer containing decreasing concentrations of ammonium sulphate (25 - 0 % (w/v)) and increasing concentrations of the non-ionic detergent NP40 (0 - 0.5% (w/v)). After each step the hydrophobicity of the matrix was
therefore reduced. The flow through, first 1 ml of the wash and each fraction of eluate were tested using the EMSA.

2.5.3 Ion Exchange Chromatography

During this part of the investigation a total of four gel matrices were tested: two cationic exchangers, phosphocellulose and heparin (Pharmacia); and two anionic exchangers, DEAE-cellulose and DEAE-sephadex (Pharmacia).

Two approaches were used in trials to choose a suitable matrix. The first was a batch method where 250 µl of the matrix, pre-equilibrated in homogenising buffer, was added to 1 ml of the embryo or unfertilised egg cell extract. After rigorous mixing followed by centrifugation, the supernatant (the flow through) was removed. The pellet was then washed in 1 ml of normal homogenising buffer, followed by similar aliquots of buffer containing increasing concentrations of salt (0.1 - 2 M NaCl). Between each wash the mixture was pelleted in order for the supernatant to be removed for testing using the EMSA.

For the second approach a column containing 1 ml of matrix, pre-equilibrated in homogenising buffer, was prepared. The sample (1ml) was layered onto the surface of the bed and allowed to drain into the column. The matrix was washed in four column volumes of homogenising buffer, the first 1 ml of which was collected. A stepwise elution was then carried out and all aliquots were tested using EMSA.

2.5.4 DNA Affinity Chromatography

2.5.4.1 Using Non-cleavable Biotinylated Probes

Prewashed DNA coupled streptavidin coated magnetic beads (see section 2.4.3.2) were added to 500 µl of the embryo or unfertilised egg cell extract, and left gently shaking at room temperature for 10 min. for any binding reactions to occur. The beads were
collected using a magnetic particle concentrator (MPC, Dynal) and then washed once in homogenising buffer containing 1% (w/v) BSA and 10x sonicated calf thymus (sCT) DNA, and once in homogenising buffer containing 10x sCT DNA. Proteins still associated with the DNA were then eluted into a small volume using elevated salt concentrations. The experimental variables tested included the pH of the homogenising buffer (7.4 - 10.0), the pH (7.4 - 10.0) and ionic strength (0.25 - 2.0 M Na/KCl) of the elution buffer and the effect of a non-ionic detergent (1 % (v/v) NP40). All fractions were tested for binding activity using EMSA.

**2.5.4.2 Using Cleavable Biotinylated Probes**

The probe was added to 500 µl of the embryo or unfertilised egg cell extract, and left at room temperature for 10 min. for any binding reactions to occur. 20 µl of the streptavidin coated magnetic beads (10 mg/ml) was added to the mixture which was then left gently shaking for 30 min. in order for the biotinylated probes to bind to the streptavidin coated beads. The beads were collected using a magnetic particle concentrator (MPC, Dynal) and then washed three times in EMSA buffer. The probe-protein complex was released by incubating the beads in Tris, pH 8.5, with 0.2 M DTT. The eluate was loaded onto a non-denaturing 6% polyacrylamide gel alongside aliquots from each stage of the procedure, and resolved by electrophoresis in 1x TBE running buffer at room temperature. The gel was fixed in 10 % (v/v) glacial acetic acid for 10 min. and 20 % (v/v) ethanol for 15 min., dried and exposed to an autoradiograph at -70 °C. The autoradiographic image was enhanced with the use of an intensifying screen and by pre-exposing the film to a hypersensitising light flash (Laskey, 1980).
2.6 Characterisation of Proteins

2.6.1 Electrophoretic Mobility Shift Assay

The standard binding reaction (10 µl) contained 2.5 µl 4x EMSA buffer (50 mM Tris-HCl, pH 7.5, 50 mM KCl, 6% (v/v) glycerol), 1 µl 10x Mg buffer (15 mM MgCl₂, 2.5 mM DTT, 8 mM EDTA), 1.5 mM spermidine, 5µg/ml lambda DNA and 1 µl of embryo or unfertilised egg cell extract. Non-specific protein-DNA complexes were allowed to form at room temperature for 5 min. before radiolabelled DNA probe was added. This was left for 10 min. at room temperature. For competition analysis, various amounts of unlabelled probe were incubated with the binding reaction for 5 min. at room temperature prior to the addition of the radiolabelled probe. Any DNA-protein complexes that may have formed were resolved by non-denaturing 6% (w/v) polyacrylamide gel electrophoresis in a 1x TBE running buffer at room temperature. Prior to loading onto the gel, 1.5 µl of EMSA loading buffer was added to the reaction mixture. The gel was fixed, dried and exposed as described above.

2.6.2 Crosslinking

2.6.2.1 Ultra Violet Light Induced Crosslinking

Three approaches were taken for this investigation. The first required preparative EMS assays (5x the normal reaction volume) to be set up using the unbiotinylated BrdU containing probes. After the DNA-protein complex had formed the EMSA reactions were transferred to wells in a microtitre plate positioned on top of some ice to avoid overheating the reactions during the irradiation. This was irradiated from above (2 cm) using a UV transilluminator (λ = 305 - 310 nm). Samples were removed at intervals to monitor the progress of the reaction. On completion, the samples were mixed with an equal volume of SDS-PAGE loading buffer, heated to 80 °C for 1 min. and resolved by 10
% (w/v) denaturing polyacrylamide gel electrophoresis (SDS-PAGE), either at 100 v for 6 hours or 8 mA overnight. Afterwards, gels were fixed, dried and exposed as described above.

For the second approach a normal EMSA reaction was performed, again using unbiotinylated BrdU containing probes. Four assays were performed; plus and minus extract using both probes. After the DNA-protein complex had formed, the reactions were resolved by non-denaturing 6 % (w/v) polyacrylamide gel electrophoresis. The gel was then positioned on top of some ice 2 cm beneath the UV transilluminator (λ = 300 - 310 nm) and irradiated for 30 min. The individual lanes containing the four assays were then excised from the rest of the gel and placed in the stacking portion of a 10 % (w/v) denaturing polyacrylamide gel (SDS-PAG) before it had polymerised. This allowed the DNA-protein complexes to be electrophoresed directly into the SDS-PAG, after which the gel was fixed, dried and exposed as described above.

In the third approach, biotinylated BrdU containing probes were used in preparative EMSA reactions (5x the normal reaction volume). After the formation of the DNA-protein complex, 20 µl of the streptavidin coated magnetic beads (10 mg/ml) was added to the reactions, which were then gently shaken for 30 min. to allow the biotinylated probes to bind to the streptavidin coated beads. The beads were collected using a magnetic particle concentrator (MPC, Dynal), washed three times in EMSA buffer before the probe-protein complex was released by incubating the beads in Tris, pH 8.5, containing 0.2 M DTT. The eluate was placed on top of some ice and irradiated for various times (5 - 30 min.) using the UV transilluminator (λ = 300 - 310 nm) positioned 2 cm above. The samples were then treated in the same manner as described for the first approach.
2.6.2.2 Chemically Induced Protein-Protein Crosslinking

The procedure for UV crosslinking using the biotinylated probes was repeated with the following addition. After the UV irradiation the samples were split in two, one of which was stored temporarily, while to the other 1 μl of 10 mM BS3, a protein crosslinking agent, was added. This was incubated at room temperature for 30 min., after which all the samples were resolved by 10 % (w/v) denaturing polyacrylamide gel electrophoresis (SDS-PAGE) as described above.

2.6.3 Gel Filtration

A 50 ml Sephacryl S400 (Pharmacia) gel filtration column with a diameter of 1.6 cm equilibrated with homogenising buffer. The flow rate was maintained at 6 ml hr\(^{-1}\) by a peristaltic pump and the column was calibrated with cytochrome C (12 kDa), bovine serum albumin (67 kDa), thyroglobulin (669 kDa) and dextran blue (>1000 kDa).

Care was taken to apply the sample (500 μl) evenly to the column when there was no liquid above the matrix, to ensure that the starting plate occupied the minimum volume. A UV monitor, set to record absorption at A\(_{280}\)nm, recorded the distribution of the protein in the outflow. An automatic fraction collector was used to collect 1 ml fractions into eppendorf tubes containing the cocktail of protease inhibitors (described above). These were later tested for binding activity using the EMSA.

2.6.4 DNA Footprinting

2.6.4.1 DMS Protection Assay

The method described here contains elements from the techniques described by Kuwabara and Sigman (1987) and Sasse-Dwight and Gralla (1991).

A 235 base pair sequence containing both the suggested mesodermal and ectodermal binding sites (Mayor et al., 1993), from +74 to -160 bases relative to the
transcription start site of the \textit{Xsna} promoter, was amplified by PCR. Some of the DNA was radiolabelled with [γ\textsuperscript{32}P] ATP (Amersham) and T4 polynucleotide kinase (Gibco BRL). This probe was used to indicate on a mobility shift assay the position of its binding activity in the gel, so that unlabelled DNA, used in binding reactions run in neighbouring lanes could be excised. The binding reactions were carried out as for the mobility shift assay described above. The excised complexes were immersed in 100 ml of 50 mM Tris-HCl, pH 8.0.

Freshly prepared dimethyl sulphate (DMS : 6 μl of 150 mM) was added to the immersed excised binding activities and incubated for 1 or 2.5 min. The reaction was quenched with 200 ml of cold stop buffer (3 M ammonium acetate, 1 M 2-mercaptoethanol, 20 mM EDTA) and the DNA left to elute from the gel slice overnight at room temperature. The DNA was precipitated by the addition of 2 volumes of ethanol, washed in 70% (v/v) ethanol, dried and resuspended in 100 μl of 1 M piperidine. The sample was then heated to 90 °C in a clamped rack for 30 min. before being cooled on ice and passed through a Sephadex G50-80 spin column at 500 g for 2 min. to give approximately 70 μl of cleaved DNA in water.

The cleaved DNA was split equally between two reaction samples with either the 5' or 3' primer so that only one strand was amplified in each case. The reaction samples contained the 35 μl of cleaved DNA, 1 μl of \textsuperscript{32}P end labelled primer, 10 μl of Taq reaction buffer (Advanced Biotechnologies), 5 μl of MgCl\textsubscript{2}, 4 μl of 5 mM dNTP mix and 44.5 μl of distilled water. After the samples had been boiled for 2 minute 0.5 μl of Taq polymerase (Perkin Elmer) was added. They were then overlaid with mineral oil and incubated for 30 cycles of 1 minute at 94 °C, 2 min. at 57 °C and 3 min. at 72 °C before being left to cool to room temperature.
The amplified DNA was extracted with phenol/chloroform (1:1 v/v) and precipitated with 2.5 volumes of 95% (v/v) ethanol and 1/3 volume quench (4 mM ammonium acetate, 20 mM EDTA) at -20 °C overnight. It was then pelleted by centrifugation at 4 °C, washed in 70% (v/v) ethanol and dried before being resuspended in loading buffer. After a 2 minute boil, the samples were fractionated on a 6% acrylamide sequencing gel together with a sequence standard at 33 watts for 90 min. Gels were fixed, dried and exposed as described above.

2.6.4.2 Copper-phenanthroline Protection Assay
A preparative EMSA assay (5x the normal reaction volume) was carried out using the copper-phenanthroline protection assay probes, and resolved on a 6% (w/v) non-denaturing gel until the bromophenol blue had reached the bottom. Four assays were performed; plus and minus extract using both probes. They were loaded onto the gel so that the plus extracts were on one half, and the minus extracts on the other, so that the gel could be split into two and each half treated separately. In order to visualise the binding activities an aliquot from each assay was resolved on a separate 6% (w/v) non-denaturing gel, which was later fixed, dried and exposed as described above.

The two halves of the gel were removed from the glass plates and immersed into separate scrupulously clean pyrex dishes, each containing 200 ml of 10 mM Tris-HCl, pH 8.0. 2 ml of OP solution (40 mM 1,10-phenanthroline) was mixed with 2 ml of Cu²⁺ solution (9 mM CuSO₄) and made up to 40 ml with dH₂O to form the [OP]²⁻Cu²⁺ chelate: the efficient formation of which is indicated by a light blue colour. 20 ml of the chelate was added to each of the equilibrating gels and distributed evenly. The chemical nucleolytic reaction was initiated in each gel by the addition of 20 ml of MPA solution (58 mM 3-mercaptopropionic acid) which turns the combined solution faintly brownish. The
appearance of a brown precipitate indicates the presence of impurities. The reactions for
the plus and minus assays were allowed to proceed for 5 and 2 min., respectively, before
being quenched with the addition of the STOP solution (28 mM
2,9-dimethyl-1,10-phenanthroline (Neocuproine)) which changes the colour of the reaction
to a deep yellow. This was incubated for 2 min. with constant shaking before being rinsed
with four changes of dH₂O.

The individual lanes containing the four assays were excised from the rest of the
gel and split into 1 cm sections. The free and shifted radiolabelled DNA sections were
identified using Cherenkov counting. Elution from the gel sections was carried out
overnight at 55 °C in the elution buffer, and the eluate was with phenol/chloroform (1:1
v/v), to ensure the removal of any residual polyacrylamide particles. The DNA was
ethanol precipitated, dried and rehydrated in 5 μl of sequencing STOP buffer. Following 2
min. at 80 °C, the DNA ladders were resolved, alongside appropriate sequence standards,
and were processed as above.

2.6.5 Denaturing Polyacrylamide Gel Electrophoresis (SDS-PAGE) for
Protein Separation

Denaturing SDS gels were prepared by the method of Laemmli et al. (1970) with a
resolving gel of 10 % (w/v). Protein samples were mixed with an equal volume of
SDS-PAGE loading buffer and boiled for 1 min. prior to loading into the wells. The gel
was electrophoresed either at 100v for 6 hours or 8 mA overnight until the bromophenol
had reached the bottom of the gel. To visualise the proteins the gel was stained with
brilliant blue (1 % (w/v) in 40 % (v/v) propanol and 10 % (v/v) acetic acid) for 30 min.
and then de-stained over an hour with several changes of de-stain (10 % acetic acid, 5 %
methanol). The destaining process was accelerated if a polyethelene sponge was left floating in the destain adsorb the excess stain.

Increased sensitivity of staining was obtained using the silver staining method of Wray et al. (1981). The destained gel was washed at 50 °C with 50 % (v/v) methanol and water, alternating twice, to remove all the acetic acid and blue background. The proteins were then equilibrated, at 50 °C with constant agitation, in 50 % (v/v) methanol containing 0.04 % (v/v) formaldehyde. The formaldehyde was then removed by washing the gel in water followed by 50 % (v/v) methanol, again at 50 °C. This was repeated to ensure all the aldehyde had been removed. The gel was treated with the silver reagent (0.02 M NaOH, 0.21 M NH₄OH, 0.047 M AgNO₃) at 20 °C, for 30 min. with constant agitation. The reagent was removed and the gel washed repeatedly in dH₂O for 10 min. prior to the addition of the developer (1 ml of 2.5 % (w/v) citrate and 0.25 ml of 38 % (v/v) formaldehyde to 500 ml of dH₂O). Constant agitation of the developing gel normally resulted in the appearance of the silver stained protein bands in less than 15 min. The gel was fixed in 45 % (v/v) methanol and 5 % (v/v) acetic acid, and then washed in 50 % (v/v) methanol to prevent fading.

2.6.6 Post Synthetic Labelling of DNA-bound Protein with ³⁵SLR

As a more sensitive alternative to chemical staining, proteins were radiolabelled by incubating them with 1 μl of [³⁵S]-sulphur labelling reagent (³⁵SLR; Amersham) for 30 min., prior to their resolution by SDS-PAGE. This reagent forms a covalent bond with amino and sulphhydryl groups within the proteins, and so therefore, wherever possible the reaction must performed in the absence of thiol reagents such as DTT, and buffers containing a free amino group, such as Tris.
A fluorograph was prepared by treating the gel with the organic scintillator 2, 5-diphenyloxazole (PPO) as described by Laskey (1980). After the gel had been completely rehydrated in \( \text{dH}_2\text{O} \), it was soaked in dimethyl sulphoxide (DMSO) for 30 min. followed by a second immersion in fresh DMSO for an additional 30 min. This was to ensure that all the water had been removed which would have otherwise prevented the PPO from entering the gel. The gel was then immersed in 22 % (w/v) solution of PPO in DMSO for at least 1 hour. The DMSO was then removed by thorough washing in \( \text{dH}_2\text{O} \) as the presence of any residual DMSO will prevent the completion of the next step, that is, the drying under vacuum of the impregnated gel. The radiolabelled proteins were then visualised by exposing the gel to XAR Kodak autoradiographic film at -70 °C.

2.6.7 Protease Clipping

Biotinylated probes were used in a preparative EMSA reaction (16x the normal volume). The band shift complex was bound to streptavidin coated beads and separated with the magnetic particle concentrator (MPC, Dynal) as above, and then washed three times in EMSA buffer before the probe-protein complex was released by incubating the beads with 20 μl of EMSA buffer containing 0.2 M DTT. The eluate was made up to 160 μl with the addition of 0.05 M Tris, pH 8.0, and then split into two sets of 8 aliquots, which were then digested with serial dilutions of Arg C (10 ng μl\(^{-1}\)) or Glu C (10 ng μl\(^{-1}\)) for 15 min. at room temperature. The reactions were terminated by the addition of 1.5 μl of EMSA loading buffer, and the products were resolved on a non-denaturing 6% (w/v) polyacrylamide gel. The gel was fixed, dried and exposed as described above.
2.7 RNA Analysis

2.7.1 Preparation of Capped RNA for Injection

The template for the Xsna RNA was the coding sequence of the gene, cloned into the Nco/Eco RI restriction sites of the vector XβM, flanked by the SP6 promoter and the globin tail containing a poly-dA tail. The vector was cut to produce 1 µg of template DNA, and mixed with 5.0 µl of 5x Transcription buffer (Cambio), 5.0 µl of cap compound (2.5 mM GpppGppp), 2.5 µl DTT (0.1 M), 2.5 µl of low G NTP mix (10 mM UTP, ATP and CTP, and 1 mM GTP), 1.0 µl RNasin (RNAse inhibitor: Cambio) and made up to 24 µl with the addition of dH₂O. RNA polymerase was added (1.0 µl), and the mixture incubated for 20 min. at 37 °C before 1 µl of 25 mM GTP was added to the reaction. After a further incubation at 37 °C for 2 hrs., 1 µl of DNase was added and the incubation continued for another 15 min. STOP/LOAD buffer was added to an aliquot of the mixture, which was then resolved on a 6 % (w/v) PAG to check the size of the product. Once the correct size was confirmed, 100 µl of 5 mM EDTA was added, and the mixture extracted with phenol/chloroform (1:1 (v/v)), ethanol precipitated, dried under vacuum and rehydrated in dH₂O. The precipitation was repeated to remove all traces of any excess nucleotides.

2.7.2 Preparation of RNA from the Embryos

Embryos staged according to Nieuwkoop and Faber (1967), or their dissected animal caps, were dispersed in 0.5 ml embryo homogenisation buffer (RNA: 3x TEN; 1 % (w/v) SDS; 1 % (v/v) β-mercaptoethanol) as quickly as was possible. No more than 10 embryos or 30 caps were processed in 0.5 ml. The homogenate was incubated at 37 °C for 30 min. with 1 µl proteinase K (20 µg/µl:) and then extracted with phenol/chloroform (1:1 (v/v)), ethanol
precipitated, dried under vacuum and rehydrated in 100 µl of DNA digestion mix (1x One Phor All buffer (Pharmacia); 1 u DNAse I (Cambio); dH₂O). The mixture was incubated at 37 °C for 30 min. before a second phenol chloroform extraction, followed by ethanol precipitation and drying under vacuum. The RNA was then taken up in TE/S buffer (10 - 100 µl) and stored either on dry ice or at -80 °C.

2.7.3 RNAse Protection Assay

This method is a modified version of that described by Mironov (1995), and was developed by R. Morgan (personal communication).

2.7.2.1 Hybridisation

The RNAse protection probes (10 µl total volume) was added to a mixture containing 5 µl RNA (~ 2 caps or 1/2 an embryo's worth) and 5 µl 10x MOPS buffer (4 M NaCl; 0.4 M MOPS, pH 6.4; 10 mM EDTA). 30 µl of oil was laid on top of each mixture and then they were incubated at 85 °C for 60 min. After cooling as much oil as possible was removed.

2.7.2.2 RNAse Digestion

Single stranded RNA was digested by adding 80 µl of RNAse digestion buffer (10 mM Tris-HCl, pH 7; 1 mM EDTA; 10 mM MgCl₂; 300 mM NaCl) containing 0.1 µl of RNAse A (5400 kunitz units/ml in glycerol) were added to the hybridised RNA and incubating at 37 °C. After 30 min., samples were extracted with phenol/chloroform (1:1 (v/v)), ethanol precipitated and dried under vacuum. The pellet was taken up in TE/S buffer, left to rehydrate at room temperature for 20 min. and mixed with 5 µl of STOP/LOAD buffer. After 2 min. at 90 °C, the RNA molecules were resolved on a 6% denaturing polyacrylamide gel at 33 W until the bromophenol blue reached the bottom of the gel. An autoradiograph was prepared after the gel was fixed and dried, as described above.
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of Promoter Binding Proteins Using EMSA</td>
<td></td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>73</td>
</tr>
<tr>
<td>3.2 Optimum Conditions for the EMSA</td>
<td>73</td>
</tr>
<tr>
<td>3.3 The Developmental Profile of the Binding Activities</td>
<td>85</td>
</tr>
<tr>
<td>3.4 Discussion</td>
<td>88</td>
</tr>
<tr>
<td>3.5 Summary</td>
<td>89</td>
</tr>
</tbody>
</table>
3.1 Introduction

The investigation of gene regulatory mechanisms and other biological processes based on protein-nucleic acid interaction requires biochemical analysis of the binding reactions. The electrophoretic mobility shift assay (EMSA) is a powerful approach for qualitative and, to a lesser extent, quantitative analysis of protein-nucleic acid interactions. The method is based on the observation that the electrophoretic mobility of DNA through a polyacrylamide or agarose gel is reduced when protein is bound to it (Garner and Revin, 1981; Fried and Crothers, 1981). The EMSA is a quick, simple and sensitive method of identifying proteins that bind to specific sequences. It can be used to determine the nucleic acid sequence favoured by a particular DNA binding protein, and can be used to analyse subunit composition.

3.2 Optimum Conditions for the EMSA

The starting point for this investigation was to use the EMSA as an assay for proteins that bind to sequences within the region of the *Xsna* promoter that had previously been shown to be essential for the correct temporal and spatial expression of the gene (Mayor et al., 1993). In preliminary experiments it was realised that the binding activities were present in unfertilised egg extracts as well as those prepared from embryos from stages when *Xsna* is expressed (see section 3.3). Major variations in the amount of binding activity observed were encountered, and so it was more convenient to use the unfertilised eggs rather than culture embryos to these later stage. After collection, the unfertilised eggs were homogenised in four volumes of buffer in the presence of a cocktail of inhibitors. The yolk, and possibly the chromatin, were removed by centrifugation at 10,000 g for 10 min.
A wide range of variables, (table 3.1), can potentially affect the sensitivity of the assay, and were therefore systematically investigated. The assays were carried out using probes from the region of the promoter known to be responsible for the differential expression of Xsna in the mesoderm and ectoderm of the developing embryo. The probes are referred to as M and E to indicate their allegiance to these tissues (appendix A). During their preparation, care was taken to ensure that the probes had formed the correct duplex, to avoid any misinterpretation of the data (see chapter 4).

A range of sodium and potassium salt concentrations were tested, but although salt was required, neither the cation nor the salt concentration appeared critical. Acetate\(^-,\) which is thought to be a more physiologically relevant anion than Cl\(^-\), (Leirmao, S. et al., 1987), had no discernible effect. Similarly, while it was necessary to provide magnesium and spermidine, their exact concentration was also not critical. The only component whose concentration affected the assay was the non-specific competitor, lambda DNA. Lambda DNA was chosen as the non-specific competitor as it was a convenient source of specific size and unvarying properties that contained no sequence similar to that of the Xsna promoter. The principle of the method is that it should be present at the highest concentration possible to sequester non-specific DNA binding proteins without interfering with interactions between the probe and the sequence specific proteins (figure 3.1). This is confirmed by the observation that at lambda concentrations close to the maximum, competition experiments with non-radiolabelled probe indicate that the complexes formed are completely specific for the probe.

The non-ionic detergent NP40 occasionally appeared to enhance the binding activity in some assays, but this was not consistently seen. The ionic detergent CHAPS is also reported to have beneficial effects on the EMSA (Hassanain et al., 1993) but this was
Table 3.1  The effect of components of the EMSA on the binding activities.

$M_L = \text{Large mesodermal complex. } M_s = \text{Small mesodermal complex. } E = \text{Ectodermal complex.}$
## EMSA conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Standard EMSA conditions</th>
<th>Range Tested</th>
<th>Effect on the binding activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>7.5</td>
<td>6.5 - 10.0</td>
<td>No difference</td>
</tr>
<tr>
<td>Salt concentration</td>
<td>50 mM KCl</td>
<td>25-600 mM KCl</td>
<td>Salt is required, but neither the concentration nor the cation is critical</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25-600 mM NaCl</td>
<td></td>
</tr>
<tr>
<td>Anion</td>
<td>Cl⁻ (50 mM)</td>
<td>acetate⁻ (25-600 mM)</td>
<td>No difference</td>
</tr>
<tr>
<td>Mg²⁺</td>
<td>1.5 mM</td>
<td>0.0 - 5.0 mM</td>
<td>Required but concentration not critical</td>
</tr>
<tr>
<td>Spermidine</td>
<td>1.5 mM</td>
<td>0.0 - 5.0 mM</td>
<td></td>
</tr>
<tr>
<td>DTT concentration</td>
<td>0.25 mM</td>
<td>0.0 - 2 mM</td>
<td>Not necessary, but no ill effect</td>
</tr>
<tr>
<td>λ DNA concentration</td>
<td>5 µg/ml</td>
<td>0.18 - 180.0 µg/ml</td>
<td>Binding activity was observed up to 12 µg/ml</td>
</tr>
<tr>
<td>Ionic detergent - CHAPS</td>
<td>0.0 %</td>
<td>0.1 - 10.0 (w/v) %</td>
<td>2.5 - 5.0 (w/v) % depleted M₅ and E</td>
</tr>
<tr>
<td>Non-ionic detergent - NP40</td>
<td>0.0 %</td>
<td>1 (w/v) %</td>
<td>M₅ and E were slightly enhanced in some assays</td>
</tr>
<tr>
<td>1,10 ortho phenanthroline (chelating agent)</td>
<td>0 mM</td>
<td>2 mM</td>
<td>No effect, suggests that zinc is not required for binding</td>
</tr>
</tbody>
</table>
Figure 3.1  The Effect of Increasing Concentrations of Lambda DNA as Non-specific Competitor in the Electrophoretic Mobility Shift Assay.

A series of EMSA reactions containing the M probe were setup with a range of lambda DNA concentrations (0.1 to 200 nmol) as the non-specific competitor. From 10 nmol, the lambda began to effect the binding activity until at 200 nmol it was almost completely abolished. The sharp threshold was probably attributable to the lack of excess probe at the lower concentrations of lambda DNA. The same effect was seen in EMSA reactions containing probe E (data not shown).
not seen over a wide concentration range using the *X. nematodes* probes. In fact, at 2.5 - 5.0% the binding activity began to reduce. The wide range of pH in which the binding activity could still be detected was surprising as it indicates that the protein-DNA and any protein-protein interactions are unaffected by significant changes in ionic charge of the protein. The binding activities were unaffected by 1, 10, ortho-phenanthroline, a chelator of divalent cations such as Zn$^{2+}$, suggesting that this metal was not required for the protein-DNA interaction.

Two major sources of variability of the binding activity have been identified: the quality of the extracts obtained from the eggs of different frogs, and an unusual cold sensitivity of the binding activity that these extracts exhibit. The biggest problem was the cold sensitivity of the complex. This can be seen in figure 3.2, which shows an experiment where aliquots of the same cell extract were stored at 20 °C, 4 °C, -20 °C and -70 °C overnight, after which they were incubated at either 20 °C or 37 °C for 1, 2 or 3 hours to see if any inactivation was reversible. Each aliquot was tested using probe M. Freezing the extract inactivates the capacity to form a DNA-protein complex. Storage at -70 °C was less damaging than -20 °C, possibly because the freezing process was quicker. This can be avoided to some extent by incubation at 20 °C or 37 °C, but the proteins are short lived probably due to the action of proteases. Similar results were obtained with the E probe (data not shown). In this experiment it appeared that storage of the extract at 4 °C had no adverse effect on the binding activities. However, subsequent experience showed that this was not always the case, and so the extracts were routinely kept at 15 °C for periods no longer than 24 hours.

Figure 3.2 also shows the two major binding activities observed with probe M. $M_L$ was the activity normally seen when the cell extract prepared was fresh. $M_s$ was seen
Figure 3.2  The sensitivity of the binding activity to storage at low temperatures, seen with probe M.  

(See text for details).

\( M_L = \text{Large mesodermal complex.} \quad M_s = \text{Small mesodermal complex.} \)
after a period of storage probably resulting from degradation or dissociation of $M_L$, and occasionally appeared in fresh extracts from eggs that were presumably of poor quality. In this experiment $M_s$ became more apparent after the extract aliquots were incubated in the attempt to reactivate the binding activities. The binding activity seen with probe E was similar in size and intensity to $M_L$. A decrease in the size of this activity to one comparable with $M_s$ was never seen.

Another possible source of the variability could have resulted from sequestration of the binding activities within the chromatin, which may have been removed by centrifugation during extract preparation. To address this possibility, 1.0 M NaCl was added to some freshly prepared embryo extract to facilitate the extraction of proteins from the chromatin, but this procedure did not improve the yield of the binding proteins. In addition, any contaminating yolk that had not been removed from the extract prior to the NaCl treatment would have been solubilised in the salt, possibly increasing the non-specific competition for the probe DNA.

It was apparent that the action of proteases had to be minimised, and so during the course of this work the panel of inhibitors added to the fresh extracts was expanded to included the seven described in chapter 2. Even so, in a typical extract stored at 15 °C, there was a significant reduction in binding activity only 48 hours after the cell extract was prepared (figure 3.3). This figure shows the two major binding activities seen with probe $M$, and also clearly demonstrates that a similar pattern is not found with probe E. Although in this experiment the probe was not in excess, it is unlikely that the appearance of $M_s$ was solely attributable to a lack of probe saturation because in other experiments (see figure 3.2) it did appear in the presence of excess probe. However, this did not discount the possibility that $M_s$ has a lower binding affinity for this sequence than $M_L$. A larger binding activity associated with $M$, which only occasionally appeared, was also seen in the fresh cell extract. In later experiments, it was found that the stability of the extract could be improved by performing an ammonium sulphate fractionation with the extract (see chapter 6).
**Figure 3.3** EMSA's showing the change in M and E binding activities after 48 hrs.

Left-hand lane = fresh extract. Right-hand lane = the same extract 48 hours later.

\( M_L \) = Large mesodermal complex. \( M_s \) = Small mesodermal complex. \( E \) = Ectodermal complex.
3.3 The Developmental Profile of the Binding Activities

Once it was established that binding activities specific for the probes used could be detected in embryo extracts, it was important to discover at what embryonic stage they first appeared, and whether their characteristics altered around the time when they are supposed to direct \textit{Xsna} expression and to identify the embryonic stage that would be the best source of experimental material.

The distribution of binding activity throughout early development was examined in a randomised population of embryos derived from a number of frogs to ensure that samples were not biased towards the progeny of one mother. Prior to fertilisation, a sample of the unfertilised eggs was removed and stored. Further equivalent samples of the embryos were collected at the developmental stages shown in figure 3.4. The extracts from these samples were freshly prepared and tested using the EMSA with an excess of probe in order to assess any changes in concentration. The M and E probes from the putative mesodermal and ectodermal control regions were used, as well as probe P which contains both elements (appendix A).

\textit{Zygotic expression of mRNA in \textit{Xenopus} begins soon after the mid-blastula transition}. Expression of \textit{Xsna} is seen in the mesoderm from stage 9 and in the ectoderm from stage 10.5, although there is a very small amount of maternal \textit{Xsna} mRNA present in the vegetal pole prior to this (Sargent and Bennett, 1990). The promoter binding activity was detected at all stages tested, from unfertilised eggs through to stage 35 (figure 3.4). The activity was strong up to stage 13, after which it began to diminish until a time between stages 18 and 20 when it stabilised to a constant level. In addition to the main binding activities, (E and M, figure 3.4), a larger band was present during the early stages up to stage 13, that disappeared as the main activity reduced in intensity. Initially a smaller
The embryos were staged according to Nieuwkoop and Faber (1967). Each track contained the equivalent of a fifth of an embryo.
Chapter 3 - Identification of Promoter Binding Proteins Using EMSA

[Image of gel electrophoresis with multiple bands and labels such as "Free Probe," "M," and "E." There are also numerical labels indicating lanes or wells.]

M: Lanes 8, 11, 13, 14, 18, 20, 22, 27, 35
E: Lanes 8, 11, 13, 14, 18, 20, 22, 27, 35

Chapter 3 - Identification of Promoter Binding Proteins Using EMSA
activity was also seen with probe M, which may have been M_s (see section 3.2). After stage 14, this activity also reduced in intensity, while simultaneously a slightly smaller activity became more prominent. Extracts prepared on other occasions did not always contain these other bands. The pattern of activity seen using probe P was identical to that seen using probes E in figure 3.4 (data not shown).

An extract from homogenised *Xenopus* ovary was also tested using probes M and E. Binding activities similar to that seen with the mature eggs were observed (data not shown), but the precise amount per oocyte was uncertain. The binding activities probably appear during oocyte development.

The concentration of binding activities per egg can be estimated from the amount of material required to give a complete retardation of a known molar concentration of probe, assuming that one molecule equivalent of activity is bound per molecule of probe. In the most favourable of cases this suggests there are approximately 5 femol of complex per egg. The specificity of all interactions was demonstrated by competition with unlabelled probe DNA, as discussed above.

3.4 Discussion

The conditions under which a specific assay of *Xsna* promoter binding proteins can be made are not very exacting and do not offer any explanation of the enormous variation in activity found in extracts. Partially purified extracts may have more stringent requirements.

The binding activities observed with probes M and E appear to be sensitive to storage at temperatures below 4 °C. This very peculiar finding is not easy to explain. However, if the binding activity is a multi-subunit complex held together by hydrophobic
interactions, it is possible that they dissociate at low temperatures: hydrophobic bonds are considered to be weaker at low temperatures (Harper and Row, 1966). The appearance of a smaller complex derived from $M_c$ is consistent with the idea of a multi-subunit complex.

As a consequence of this discovery it became standard practice to make fresh cell extract, in the presence of a cocktail of protease inhibitors, before each experiment. During the experiments the extract was kept in a rack above some ice so that its temperature was held at approximately 10 °C. Subsequent experience suggests that the cold sensitivity of the binding activities can vary from one batch of extract to another.

The presence of a binding activity in embryos before the start of $Xsna$ transcription indicates that the appearance of this activity is not the trigger for transcription. One interpretation of this is that to become engaged in transcription, a post-translational modification of the proteins is necessary, or, it is also possible that the complex subunit composition is modified. There is very little increase in the expression of $Xsna$ mRNA after stage 14, although it persists until later stages (Sargent and Bennett, 1990). This may explain why the level of the binding activities gradually reduces to the low level observed at stage 20, possibly indicating that a basal level is all that is required for the maintenance of expression at later stages. However, the decrease in extractable binding activity may reflect sequestration of protein into the chromatin that is actively expressing $Xsna$.

### 3.5 Summary

1) The EMSA can be used to detect an activity that binds to an $Xsna$ promoter sequence in a crude extract over a wide range of conditions.

2) The amount of binding activity varies between different batches of extract, and can be inactivated at low temperatures.
3) Under certain conditions probe M can identify a lower mobility complex (M₃).

4) The developmental profile of the binding activities suggests that a post-translational modification may be necessary for these proteins to engage in transcription.
Investigating the Binding Site

4.1 Introduction 92
4.2 Using EMSA to Investigate the Binding Sequence 93
4.3 DNA Footprinting 94
4.4 Discussion 106
4.5 Summary 108
4.1 Introduction

Previous work by Mayor et al. (1993) identified a region within the \( Xsna \) promoter, -45 to -112 base pairs from the translational start site, that is responsible for the spatial and temporal expression of \( Xsna \). Deletion analysis of this region suggested two adjacent sequences could control mesodermal and ectodermal expression, possibly as the binding sites of distinct DNA transcription factors (Mayor et al., 1993). Within these two regions of the promoter there are no obvious sequence homologies with known transcription factor binding sites. It was expected that the identification of the precise limits of these binding sites, together with an investigation into the effect of base substitutions on their binding capacity, would confirm and extend the initial observations. Therefore, the main aim was to identify individual bases that, when substituted, would inhibit either of the binding activities. With this data, DNA affinity chromatography ligands could then be constructed which would bind the sequence specific proteins, or remove non-specific proteins from an extract that might otherwise be able to recognise aspects of the sequence, and thereby provide a powerful means of discriminating in favour of the specific proteins.

Throughout this work the significance of the data depended on the integrity of the probes used. Early in the investigation it was discovered that there were ssDNA binding proteins in the extract that could bind to probes that did not form a proper duplex, and which would cause a non-specific shift in the EMSA similar in mobility to the one that would be expected by a properly formed probe. To overcome this potential problem great care was taken to provide the best conditions for duplex formation. Gradual cooling from 90 °C to room temperature, whilst in the T4 kinase buffer, was sufficient for the probes based on the ectodermal region, which is GC rich, to form correctly. This was generally true for the other probes used, however, shorter mesodermal probes (set M; appendix A),
which are AT rich, required the addition of 0.1 M NaCl for satisfactory duplex formation. Even so, one probe would not anneal correctly, Ms2 (appendix A), despite repeated attempts. After the annealing reaction was completed the probes were purified on a 12 % (w/v) PAG, to remove the excess labelling reagents. Incomplete duplex formation could be identified by anomalous migration of the labelled probes. Only when it was clear that the probes had formed correctly were they used experimentally.

4.2 Using EMSA to Investigate the Binding Sequence

The specificity of the binding activities visualised as retarded bands in the EMSA was assessed by titrating the radioactive probe against increasing concentrations of specific and non-specific unlabelled probes. A specific effect will occur at a much lower concentration than the non-specific. Two non-overlapping probes derived from the mesodermal and ectodermal regions (Ms1 and Es1), were used together with a set of oligonucleotides containing four adjacent base substitutions (appendix A), designed to test their involvement in the binding activity. The substitutions were in general the alternative Watson and Crick base pair (1953), but with the purine in place of the pyrimidine. The prediction was that an alteration of an important part of the binding site would reduce its capacity to interact with the binding proteins, and that at least one of each set would show reduced binding.

Each of the probes within set E was used in a normal EMSA reaction to assess whether they contained the sequence necessary for the binding proteins to bind. These reactions were repeated in the presence of an increasing amount of specific competitor, unlabelled Es1, to estimate the relative affinities of the protein for the probes, from which their specificity profile could be assessed. Probe Es2 was bound less strongly than Es3.
and Es4, suggesting that the \(^5\text{gg}aG\) is the most significant part of the sequence. Es2 and Es4, in which the adjacent 4 bases were changed, seemed to have a greater binding activity. Probes Es2, Es3 and Es4 were not in excess in all reactions, and therefore this data should be treated as preliminary (figure 4.1).

A similar analysis of the M region was conducted using the oligonucleotides Ms1 - 4, although no useful information could be obtained using Ms2 as it could not be annealed satisfactorily. The shorter exposure of the autoradiograph (figure 4.2) indicated that Ms4 had a greatly reduced binding capacity, suggesting that the \(^5\text{atta}G\) sequence was particularly significant. The probe excess qualification mentioned above also applies to this data.

These data suggested distinct sequences within the mesodermal and ectodermal regions mediate the interaction of the binding proteins with the promoter. However, it seems unlikely that single or even double base substitutions will abolish binding completely. While the binding sites appeared to extend over a number of bases, the binding capacity seemed to be tolerant to gross substitutions.

An alternative approach to using the radioactive probes above, is to assess the binding specificity of the radiolabelled probes Ms1 and Es1 in the presence of increasing amounts of unlabelled duplexes containing the substituted bases. Uncertainty about the integrity of the unlabelled duplexes suggested this may be unreliable. However, subsequent work indicates that similar results are obtained (M. F. Bennett, personal communication).

### 4.3 DNA Footprinting

DNA footprinting is another powerful method that has been developed to identify sequence specific contacts between DNA and binding proteins. The principle is that sequence specific proteins protect the DNA to which they are bound from enzymic or
Figure 4.1 (a) The Binding Activities Observed Using Probe Set E.
The oligonucleotides for set E were based on Es1 (see figure 4.1b), and synthesised so that each one had four adjacent bases substituted. The autoradiograph was exposed to the gel for either 5 or 18 hours. The specific competitor was Es1.

(b) A Summary of Binding Activities Seen With Set E.
The underlined bases indicate the region suggested by deletion analysis to be responsible for tissue specific expression of Xsna. Ms1 and Es1 are the sequences from the putative mesodermal and ectodermal regions. The binding activity is shown relative to Es1. The 3' end is -80 bases from the transcription start site.
a) 5 Hour Exposure

18 Hour exposure

Free Probe

Key

Increasing Concentration of Specific Competitor (Es1) (first lane no competitor)

b) 5' cacaagctgcatattatagttagttggaggccccacacqca 3'

|----------------------------------|---|---|---|---|

Binding Activity

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Binding Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ms1 TAGTagctgcatattagt</td>
<td>++++</td>
</tr>
<tr>
<td>Es1 gtaggggccagca</td>
<td>++++</td>
</tr>
<tr>
<td>Es2 TTGTggccacagca</td>
<td>+</td>
</tr>
<tr>
<td>Es3 ggAATTcagca</td>
<td>+++</td>
</tr>
<tr>
<td>Es4 ggaggggATATgca</td>
<td>+</td>
</tr>
</tbody>
</table>
a) Investigating the Binding Site

- **5 Hour Exposure**
- **18 Hour Exposure**
- **Free Probe**

**Key**

\[\text{Increasing Concentration of Specific Competitor (Es1)}\] (first lane no competitor)

b) 

\[
5' \text{cacaagctgcatattagtttagtggagggcacagca} 3' \\
\text{[---M---] [---E---]}
\]

**Binding Activity**

- **Ms1 TAGTagctgcattagta**
  - ++
- **Es1 g gaggggccacagca**
  - +++
- **Es2 TTGTaggccacagca**
  - ++
- **Es3 ggagATTCacagca**
  - +++
- **Es4 ggagggcATATgca**
  - +
Figure 4.2  (a) The Binding Activities Observed Using Probe Set M.
The oligonucleotides for set M were based on Msl (see figure 4.1b), and synthesised so that each one had four adjacent bases substituted. The autoradiograph was exposed to the gel for either 5 or 18 hours. The specific competitor was Msl.

(b) A Summary of Binding Activities Seen With Set M.
The underlined bases indicate the region suggested by deletion analysis to be responsible for tissue specific expression of Xsna. Msl and Es1 are the sequences from the putative mesodermal and ectodermal regions. The binding activity is shown relative to Msl. The 3' end is -80 bases from the transcription start site. It was not possible to be confident that probe Ms2 had formed correctly, and was therefore not used (see text for details).
Chapter 4 - Investigating the Binding Site

a)  

Key

= Increasing Concentration of Specific Competitor (Ms1) (first lane no competitor)

5 Hour Exposure

18 Hour exposure

Free Probe

b)  

5' cacaagctgcatattagttagttgaggggccacagcag 3'

Binding Activity

| 5' cacaagctgcatattagttagttgaggggccacagcag 3' |

| Esl | Ms1 TAGTagctgcatattagt | Ms3 agctCGaGattagtt | Ms4 agctgcattCGGCgt |

| g gaggggccacagca | ++++ |

| TAGTagctgcatattagt | +++ |

| agctCGaGattagtt | +++ |

| agctgcattCGGCgt | + |
chemical cleavage. To define the protected sites, the probes used are radiolabelled at one end and then cleaved under conditions that give approximately one random break per molecule. This gives a range of fragments which will contain gaps if a protein protects a particular site. There are several cleavage reagents available including the endonuclease DNase I, hydroxyl radicals, dimethyl sulphate (DMS) and copper-phenanthroline, each of which has advantages and drawbacks.

DNase I cleavage has been widely used successfully many times but is not very incisive as it is relatively bulky and cannot cleave flush with the protein due to steric hindrance (Leblanc and Moss, 1994). To compound matters, the enzyme is not completely random in its action as tracts of A and T residues, or TpA dinucleotide islands, are resistant to endonucleolytic cleavage by DNase I. Since the binding sites on the promoter were already roughly known it was felt that DNase I footprinting would not resolve the binding sequences sufficiently to justify spending time optimising the system.

Strand scission by hydroxyl radicals is a chemical alternative to DNase I (Schickor and Heumann, 1994). The radicals are generated by the Fenton reaction in which iron (II) is oxidised with hydrogen peroxide. The exact manner in which the hydroxyl radicals act on DNA is not known but they are thought to remove an H-atom from the sugar moiety of the DNA backbone, and secondary reactions involving the newly created sugar radical cause the backbone to break. Hydroxyl radicals are very small and cut with almost no sequence dependence resulting in a footprint of almost perfect resolution (within 1 bp). Unfortunately the cutting reaction is incompatible with glycerol, an important component of the EMSA buffer. Glycerol acts as a radical scavenger when present at concentrations higher than 0.5 % (v/v).
4.3.1 DMS Protection Assay

Another alternative involves the use of piperidine to cleave DNA after methylation by dimethyl sulphate (DMS: Kuwabara and Sigman, 1987; Sasse-Dwight and Gralla, 1991). DMS primarily methylates guanines in the major groove and, to lesser extent, adenines in the minor groove. This methylation renders the neighbouring phosphodiester bonds susceptible to cleavage by piperidine (Shaw and Stewart, 1994). As the E region of the Xsna promoter contains a sequence of guanine residues interrupted by only one adenine, this approach was expected to identify the relevant binding site.

It is usual when using either the DNAse I or the DMS approach, for the cleavage or modification reaction to be executed on the DNA-protein complex formed in the crude extract. The disadvantage of this is, if the DNA molecules are not saturated with the binding protein there will be a "background" ladder when the cleaved molecules are resolved, which will obscure the footprint. To overcome this, the DNA-protein complexes were separated from the free DNA on a 6 % (w/v) PAG, in the manner previously described for the EMSA. As unlabelled DNA was used at this stage of the footprinting procedure, a radiolabelled EMSA reaction using the same DNA sequence was run alongside, to identify the unlabelled DNA-protein complex. The methylation reaction was then carried out while the complex was trapped within the gel slice. Not only does this remove the unbound DNA, but it is thought that the gel matrix forms a "cage" that prevents a dissociated protein from diffusing away, and so the affinity constant is effectively enhanced (Papavassiliou, 1994).

After the DMS reaction was quenched the modified DNA was eluted from the gel slice and cleaved with piperidine. The resulting fragments of DNA were amplified by multiple cycles of primer extension using Taq polymerase in a PCR machine and resolved
Figure 4.3  (a) DMS Protection Assay of the Promoter Region that Contains the M and E Elements. 
The methylation was carried out on EMSA complex purified by PAG, prior to cleavage by piperidine. The tracks on the left are the sequence of the coding strand. The results for the coding strand, after DMS treatment for 1.0 and 2.5 min., are shown. 
(See text for details).

(b) The Sequence of the Protected Region. 
The underlined bases indicate the region suggested by deletion analysis to be responsible for the expression of $X_{xna}$. M and E are the putative mesodermal and ectodermal regions. The 3' end is -80 bases from the transcription start site. $\Box$ = protected guanine residue.
Chapter 4 - Investigating the Binding Site

DMS Treatment

a) 

-80 bp 3' g a c g a c a c c g a c -80 bp g

-90 bp g g g g g t g t g t

-100 bp t t t t a c c a c c c c

-110 bp c g t c g a c a g t g c a t a t t a g t t a g t g a g g g g c c a c a g c a g c g t g c t t 3'

-120 bp 5' t c g a

DMS Treatment

1.0' 2.5'

← adenine -93

← adenine -113

b) 

-120 -110 -100 -90 -80 -70

5' ccgtcccaacaagctgcataattgtagttagtgaggggccacagcagcgcgtgcttt 3'

|--------M--------|------E------|

x-x-x-x-x-x-x-xx
on a 6 % (w/v) denaturing polyacrylamide gel to reveal that the protection of the guanine residues on the coding strand between -93 and -112 (figure 4.3). In the deletion analysis (figure 1.5), removal of this sequence resulted in the loss of mesodermal expression while the ectodermal component was maintained, although at reduced level (Mayor et al., 1993). The same experiment was attempted on several occasions with the non-coding strand, but no bands could be detected. As this method only reports on the disposition of guanine residues relative to binding proteins, this experiment could not identify other interactions. The sequence 5' of the M element also shows protection which may represent the proteins which bind to the upregulator element indicated in figure 1.5 (see section 1.6).

4.3.2 Copper-Phenanthroline Footprinting

Copper-phenanthroline (OP-Cu^2+) is an efficient chemical nuclease that cleaves the phosphodiester bonds of nucleic acids at physiological pHs and temperatures, by oxidation of the deoxyribose moiety (Papavassiliou, 1994). (OP-Cu^2+) is a small molecule and can therefore approach and cleave the DNA closer to the protein than DNase I, producing a more detailed footprint. Unlike DMS, it is almost completely random in its action although there is a minor preference for TAT and CG rather than TA. In addition, after an EMSA reaction the entire gel can be immersed in buffer and treated with the reagents directly, in order to take advantage of the cage effect described above.

The probes for this procedure were made using PCR reactions that contained a radiolabelled primer for one strand or the other, so that each reaction contained amplified DNA fragments labelled on one strand only. The primer sites were located 100 bases upstream from the sequence of interest to ensure that this region would appear clearly after resolution on a 6 % (w/v) denaturing polyacrylamide gel. Two EMSA reactions, plus and minus extract, were prepared for each probe and resolved as normal. After the gel had been immersed in buffer and treated with the cleavage reagents, the position of the free
Figure 4.4 Copper-Phenanthroline Footprinting Assay of the Promoter Region that Contains the Mesodermal and Ectodermal Elements.

(a) In order to be certain that the protein-DNA complex had correctly formed, aliquots from the four EMSA reactions prepared with $^{32}$P-probes (each probe, plus and minus extract), were resolved on a 6 % (w/v) PAG. The rest of the reactions were resolved on a separate PAG, immersed in buffer and treated with the cleavage reagents.

(b) The position of the free and complexed DNA was established using Cherenkov counting so that they could be excised, eluted and resolved alongside a sequence standard. Hypersensitive sites are visible at positions -92 and -105 bp on the coding strand, and -105 bp on the noncoding strand. The -92 bp site disappears when the proteins are bound to the DNA.
Chapter 4 - Investigating the Binding Site

a) Coding Strand Non-coding Strand

Wells -

Binding Activity -

Free Probe -

b) Coding Strand Non-coding Strand

-80 bp  -80 bp

a g c  a

g g a c

g a c

g a c

g a c

g a c

t c g a  t c g a

-90 bp  -90 bp

a g c  a

g g a c

g a c

g a c

g a c

g a c

t c g a  t c g a

-100 bp  -100 bp

a g c  a

g g a c

g a c

g a c

g a c

g a c

t c g a  t c g a

-110 bp  -110 bp

a g c  a

g g a c

g a c

g a c

g a c

g a c

t c g a  t c g a

-120 bp  -120 bp

a g c  a

g g a c

g a c

g a c

g a c

g a c

t c g a  t c g a

5' 3'
and complexed DNA was established using Cherenkov counting so that they could be excised, eluted and resolved alongside a sequence standard (figure 4.4). Using either strand, the method failed to identify protected regions, but several hypersensitive sites were apparent. The DNA without bound proteins contained two distinct hypersensitive sites on the coding strand, within the region of interest: one at approximately position -92 bp (\(^{gaggggccac}\)) and the other at -105 bp (\(^{gcattag}\)). The more 3' site disappears when the proteins were bound. There was only one site on the non-coding strand, corresponding exactly to position -105 bp.

4.4 Discussion

The identification of the M and E regions of the promoter using deletion analysis and histological scoring of the phenotypes of \(\beta\)-Gal reporter constructs rested heavily on just three deletions (Mayor et al., 1993). Removal of the sequence -112 to -96 bp greatly reduced mesodermal expression, while further reduction to -93 bp removed all mesodermal expression and reduced the level of the ectodermal component. The loss of the sequence -92 to -89 bp (\(^{gggc}\)) removed all expression. While these deletions identified regions with qualitatively different transcriptional potential, it seemed unlikely that the regions could be refined by further analysis by deletions, as it would depend on the scoring of very large numbers of injected embryos. The strategy behind this work was to investigate the M and E regions of the promoter more incisively by defining the binding site \textit{in vitro}, and then to examine the effect of base substitutions in these sites on the phenotype of reporter constructs.

The presence of proteins in extracts capable of binding to M and E provided corroboration for the idea that these elements were critical for \(Xsna\) transcription in
mesoderm and ectoderm, respectively. The analysis of the specificity of binding to these sequences using sets of oligonucleotides in which blocks of four adjacent bases were substituted, indicated that certain regions of each sequence were more important for binding than others (significant binding region in bold: M = $^g$catattaagt$^3$; E = $^5$ggaggggccac$^3$). It has not escaped notice that the sequence highlighted from the E region corresponds to the deletion described above (-92 to -89 bp) which resulted in the complete abatement of $Xsna$ expression. These two altered sequences are separated by 13 - 16 bp, which is greater than one helical turn, indicating quite strongly that the differential expression of $Xsna$ is co-ordinated through two separate trans-components.

It was expected that altering adjacent blocks of four bases in the oligonucleotide sets would produce at least one inactive probe for each region. However, the fact that the alteration of these bases reduced the binding affinity but did not abolish it completely, suggests that the proteins responsible for the binding activity make several contacts with the DNA. To discover the precise points of interaction, further investigations will have to be conducted involving different combinations of single base substitutions across the 12 bases in each set that were associated with the reduced binding activities. The effect of competition was not an all-or-nothing affair, but does suggest these regions would produce an effect on the phenotype of reporter constructs.

The DMS protection assay revealed a region on the coding strand between -93 and -113 bp, in which the guanine residues were protected from methylation by binding proteins. This included four residues within the M region ("gctgcataattagttag"), and two in the E region ("ggaggggccac"). The other guanines in the E region are not protected, but this may be because they reside in the minor groove. A low level of binding activity in the extract would limit the amount of DNA present in this assay, reducing the
intensity of the resolved pattern. Therefore, to ensure that a pattern could be detected, the DNA fragments were amplified by multiple cycles of primer extension using Taq polymerase in a PCR machine. However, it is possible that certain fragments could have been favoured by this reaction and in the process, the importance of part of the sequence could have been exaggerated at the expense of other sections.

Of the hypersensitive sites revealed by the copper-phenanthroline footprinting assay, the one within the M region at position -105 bp (\(^\text{gcatattagt}\)) was present on both strands. The other site which appears to be a guanine -92 bp (\(^\text{ggaggggccac}\)), was only seen on the coding strand when no protein was bound to the DNA. Although this method has not yet yielded a clear picture of the binding sequence, the fact that the hypersensitive sites reside within the M and E regions may suggest they have some significance in the transcriptional control of \(Xsna\).

4.5 Summary

1) An analysis of the sequence preferences of the protein factors that bind to the M and E regions suggest that -102 to -105 bp (\(^\text{atattagt}\)) and -88 to -91 bp (\(^\text{ggaggggccac}\)), respectively, are the most critical for binding.

2) DMS protection studies indicate that all of the guanine residues within the M region are protected, as are the first two in the E region. This assay did not detect a role for the other guanine residues within the E region, but does not prove that they are unimportant.

3) The copper-phenanthroline protection studies did not identify any particular binding sequences, but three hypersensitive sites were apparent, one of which disappeared in the presence of DNA binding proteins.
4) Taken together, the three sets of data indicate the sequences in which it is most likely that base substitutions would produce an effect on the phenotype of reporter constructs.
Survey of Potential Purification Methods

5.1 Introduction 111
5.2 Ammonium Sulphate Fractionation 111
5.3 Hydrophobic Interaction Chromatography 112
5.4 Ion Exchange Chromatography 113
5.5 DNA Affinity Chromatography 118
5.6 Discussion 123
5.7 Summary 124
5.1 Introduction

The rationale behind most purification strategies is to separate the active form of a particular protein, from other proteins using its distinctive properties of size, hydrophobicity and charge. The final resolution of the target protein is then achieved using a substrate affinity chromatographic step. This involves the specific and reversible adsorption of the protein onto a substrate analogue which is covalently attached to an insoluble support matrix.

The advent of sequence specific DNA affinity chromatography (Kadonaga and Tijan, 1986) and the availability of the EMSA to rapidly analyse chromatography fractions, have made the purification of DNA binding proteins a realistic goal when studying the regulation of transcription. A survey was made of potential purification steps that might be useful as preliminary steps before the final affinity purification. It is desirable to remove classes of contaminants such as nucleic acids, nucleases and possibly acidic proteins that may otherwise co-purify with the target protein, or attack the DNA affinity matrix.

Since most transcription factors are present in the cell in relatively small quantities it is important that the percentage loss at each step is minimal before affinity chromatography is employed. In addition for this investigation, the apparent instability and cold sensitivity of the proteins involved in the binding activities, imposed a tight time constraint on the overall purification procedure.

5.2 Ammonium Sulphate Fractionation

Ammonium sulphate fractionation of crude extracts provides a simple and useful first step in protein purification that can separate material on the basis of their different solubilities.
Many proteins can be stored at high concentrations in high ammonium sulphate solutions with no loss in stability.

Fractionation of the clarified unfertilised egg or embryo cell extract was achieved by increasing the ammonium sulphate concentration to 1.7 M (50 % (w/v) saturation) with the addition of the solid salt. It has proven to be impossible to isolate the activities in one discrete ammonium sulphate cut, however, at this level of saturation a large amount of the cellular proteins are precipitated, but the binding activities associated with the \textit{Xsna} promoter elements are mostly left in the supernatant. It was possible to use this supernatant in the EMSA without further treatment to remove the salt. The 1 in 10 dilution that occurred during the set-up of the assay was sufficient to reduce the salt concentration to an innocuous level. However, for procedures such as the ion exchange chromatography (see section 5.4) the proteins responsible for the activities were precipitated by increasing the ammonium sulphate concentration to 100 % (w/v) saturation. The precipitate was then resuspended for use in an equivalent volume of homogenisation buffer.

5.3 Hydrophobic Interaction Chromatography

The principle of hydrophobic interaction chromatography is that proteins with hydrophobic surfaces tend to adsorb to a hydrophobic matrix in the presence of a high concentration of salts that favour salting out. They are then eluted by decreasing the salt concentration, often with a corresponding increase in the concentration of a non-ionic detergent. To test the value of hydrophobic chromatography an aliquot of the supernatant from a 50 % (w/v) ammonium sulphate fractionation was passed through a phenyl sepharose column so that
the proteins in solution could be adsorbed onto the matrix. The loaded column was eluted with decreasing concentrations of ammonium sulphate in the presence of NP40.

Binding activities were detected with both probe M and E in the 12.5 - 6.25 % (w/v) ammonium sulphate, 0.1 - 0.2 % (w/v) NP40 fractions, (figure 5.1). In a small scale trial approximately 75 % of the cellular proteins, as judged by protein staining on an SDS PAG (data not shown), had also been eliminated. Although this indicated that the proteins are not significantly hydrophobic as detergent elution was not required to release the activity, this seemed a potentially useful purification step. However, when the process was scaled up from a 1 ml to a 30 ml column no activity could be detected in any fraction. The reason for this is unclear, but it is suspected that the complex was inactivated during the increased experimental time.

5.4 Ion Exchange Chromatography

For these experiments the precipitate from a 50 - 100 % (w/v) ammonium sulphate cut was resuspended in the buffers in which the matrix was equilibrated. Phosphocellulose is widely used in the purification of DNA binding proteins since most are basic, and therefore contaminating acidic proteins and nucleic acids do not bind. Using a phosphocellulose column equilibrated to either pH 8.5 or 9.0 it was found that the binding activities seen using probes M and E in the DNA mobility shift assay bound to the matrix and could, in the best cases, be eluted separately (figure 5.2a and b). At pH 9.0, E eluted in the 0.25 M NaCl fraction (figure 5.2a), whereas only faint activity could be seen with M (not shown). Conversely, at pH 8.5 binding activity M eluted in the 1.0 M NaCl fraction (figure 5.2b), but no binding activity with E was seen (not shown).
Figure 5.1  Hydrophobic Interaction Chromatography.
The binding activity seen in the fractions from a phenyl sepharose column equilibrated in a 50 \% (w/v) saturated solution of ammonium sulphate at pH 7.5 using the EMS assay.
### Chapter 5 - Survey of Potential Purification Methods

#### Elution

<table>
<thead>
<tr>
<th>(NH₄)₂SO₄ (%)</th>
<th>Flow through (Unbound)</th>
<th>Elution</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>12.5</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>6.25</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td></td>
</tr>
</tbody>
</table>

**Free probe E**

50 % (w/v) (NH₄)₂SO₄ extract

**Free probe M**

50 % (w/v) (NH₄)₂SO₄ extract

- Free Probe

- E

- NP40 (%)
Figure 5.2  Ion Exchange Chromatography.

The binding activity seen in the fractions from various IEC columns, using the EMSA:

(a) pH 9.0 phosphocellulose column, assayed with probe E,
(b) pH 8.5 phosphocellulose column, assayed with probe M,
(c) pH 7.5 DEAE-sepharose column, assayed with probe M,
(d) pH 7.5 DEAE-sepharose column, assayed with probe E.
Chapter 5 - Survey of Potential Purification Methods

(a) pH 9.0

(b) pH 8.5

(c) pH 7.5

(d) pH 7.5
It was expected that if the extract was passed through a DEAE-sepharose column (pH 7.5), the acidic proteins and nucleic acids would be retained by the matrix and the binding activities would come through in the void volume. Surprisingly the binding activities were retained by the matrix and eluted with 0.25 M NaCl fraction (figure 5.2c and d). This paradoxical result suggests the binding activities have both basic and acidic properties that enable them to bind to anionic and cationic exchangers. However, it is possible that although the proteins are positively charged, they are retained through their interaction with negatively charged nucleic acids. It was noted that the binding activities in the 0.25 M NaCl fractions showed a lower mobility than expected. This may correspond to the larger activity that is occasionally seen (for example, figure 3.4), but it could also be an artefact since the proteins did not interact with the matrix as expected.

Considerable difficulty has been found in reproducing these results with phosphocellulose, DEAE and heparin (data not shown). The main problem may be heterogeneity of the binding activity in the extracts (that is, the extent to which it is bound to other materials, such as chromatin), although it is possibly that the charge of the matrix destabilises protein-DNA or protein-protein interactions within the binding complex.

5.5 DNA Affinity Chromatography

DNA covalently attached to insoluble matrices of cyanogen bromide activated agarose (Sepharose) has been used for many years to purify various enzymes such as *Escherichia coli* DNA polymerase I (Arndt-Jovin et al., 1973). This method was developed further to take advantage of sequence specificity of transcription factors such as SP1, in order to effect their purification (Kadonaga and Tijan, 1986). DNA Affinity chromatography, in which binding proteins are bound to a specific sequence of DNA and subsequently eluted with high salt concentrations, became the most powerful method of purifying DNA binding proteins. More recent approaches employed tandem affinity columns, where the first column would contain beads conjugated to the DNA motif of interest with a base
substitution to eliminate sequence specific binding, and the second would contain the wild type motif (Ostrowski and Bomsztyk, 1993). The elimination of non-specific binding proteins could be improved by adding extra base substituted DNA-affinity modules, and the final purification could be enhanced by performing two rounds through the final column.

An alternative to the attachment of DNA to cyanogen activated matrices, was the use of biotinylated oligomers that could be bound to an avidin coated agarose matrix (Kasher et al., 1986). This was improved with the replacement of avidin with streptavidin which has the same biotin binding affinity, but less non-specific binding (Pahler et al., 1987). Both the streptavidin coated agarose matrix (Hagenbuchle and Wellauer, 1992), and streptavidin coated magnetic beads (Gabrielson and Huet, 1993) allowed the biotinylated DNA motifs to be mixed directly to partially purified extract before being passed through the affinity column, or mixed with the magnetic beads. This presented the possibility that the binding reaction could be performed in standard EMSA reaction conditions in which the protein-DNA interaction can be optimised (Peck et al., 1994).

Most of these methods relied on elution of the binding proteins with salt, which risks introducing contamination from non-specific proteins bound to the matrix or beads. A development to circumvent this potential problem has been the idea of using cleavable bonds to release DNA-protein complexes bound to streptavidin matrices, for example, with restriction enzymes (Peck et al., 1994), photo-cleavable bonds (Olejnik et al., 1995) or disulphide bonds (see chapter 6).

When using DNA affinity based techniques it is important that nucleases present in the cell extract should not destroy the substrate. Removal of magnesium by chelating agents is often sufficient to inactivate nucleases, although usually a preliminary purification
step is required. The apparent instability of the complexes described here necessitates using a rapid preliminary step that removes the majority of the unwanted bulk and nucleases from the extract, and then DNA affinity chromatographic steps.

The first approach to affinity chromatography was to equilibrate the precipitate from a 50 - 100 % (w/v) ammonium sulphate cut with concatamerised DNA attached through a biotin label to streptavidin coated magnetic beads (appendix A). The beads were mixed with the partially purified extract for 30 min. at room temperature. On two occasions using this protocol, faint activity was detected in the material eluted from the DNA using NaCl. The result was discouraging as there was clearly a major loss of material during the experiment. Subsequent experiments indicated that the complex is probably sensitive to high NaCl concentrations (see chapter 6).

On one of these occasions, even though the DNA used in the purification was a concatamer of the M element, activity was detected using probe E (figure 5.3). This raises the possibility that the two binding activities could be associated in one complex. The activity detected using M in the EMSA could be competed out by increasing amounts of unlabeled M, whilst equivalent amounts of unlabelled E had no effect, demonstrating that the binding activity was specific for the mesodermal element. The binding activity seen with E was also found to be specific.

In the experiment shown in figure 5.3, the smaller activity $M_s$, seen with probe M, was the only activity present in the crude extract, but this disappeared after the ammonium sulphate cut to be replaced by $M_L$. Assuming that $M_s$ is a subunit of $M_L$, this may indicate that high salt conditions increase the stability of the main complex.
Figure 5.3  Affinity Chromatography of Promoter Binding Activity on Concatamerised M Probe.

The binding activity detected, using the EMSA, in material eluted from concatamerised M DNA bound to magnetic beads.
Chapter 5 - Survey of Potential Purification Methods

Free probe M
Extract 50 % (w/v) (NH₄)₂SO₄ extract
Extrate 50 % (w/v) (NH₄)₂SO₄ extract
Free probe E
Extract 50 % (w/v) (NH₄)₂SO₄ extract
Extrate 50 % (w/v) (NH₄)₂SO₄ extract
After treatment (Unbound)

ML -
MS -
Free Probe -

E
5.6 Discussion

A large proportion of the Xsna promoter binding proteins are soluble after a 50 % (w/v) ammonium sulphate cut, which removes a large amount of inactive proteins. This is therefore a relatively quick and easy first step in the purification protocol. There is no need to remove the salt by dialysis as ammonium sulphate is not detrimental to the EMSA providing the concentration does not exceed 0.2 M.

A logical progression was to use hydrophobic interaction chromatography as an ammonium sulphate cut could be applied to the matrix. Initial results were encouraging as it was possible to isolate the binding activities from approximately 75 % of the other proteins. The fact that both activities appeared in the same fraction seemed to be an advantage as if they could be kept together until the final affinity purification step, it would halve the amount of labour and time involved for the complete purification. However, all attempts to scale-up this procedure to use a realistic volume of starting material were unsuccessful. It is likely that the inevitable increase in experimental time had a detrimental effect on the stability of the proteins involved.

Ion exchange chromatography is a standard early step in many purification schemes but in this case identification of the optimum conditions has proved to be very difficult owing to the losses in activity encountered. It is possible that if the activity is a multi-protein complex it may be dissociated by the charge of the matrix. This would appear to contradict the observation that the binding activities can be detected over a wide pH range by EMSA, however, it is possible that after removal from the bulk of the extract, the complex becomes more vulnerable to variations in its micro-environment.

Although DNA affinity chromatography was attempted a number of times, only on two occasions was it possible to detect binding activity in salt eluates of the DNA. On the
other occasions the complex was apparently adsorbed by the DNA, as judged using the EMSA, but could not be detected in any of the eluted fractions. As the procedure takes a significant time, the proteins may have been destroyed by proteases present in the crude cell extract. Alternatively, if a multi-protein complex is involved, it is possible that the proteins become dissociated by the elevated salt concentrations used to elute the DNA. Attempts to detect some binding activity in the eluate after the salt concentration had been reduced, either by dilution or by dialysis, were unsuccessful. Further experiments described below (chapter 6), indicate that exposure to high salt concentrations irreversibly damages the protein-DNA interactions.

5.7 Summary

1) Ammonium sulphate fractionation has clear advantages as the first step in the purification protocol, as long as the salt does not interfere with later steps.

2) The instability and apparent cold sensitivity of the protein complex severely curtails the use of hydrophobic interaction and ion exchange chromatography.

3) Although DNA affinity chromatography provides the best chance of a successful purification, the approach used in this chapter was clearly restricted by the deleterious effect of salt based elution from the DNA.

4) On one occasion when protein was isolated using a concatemer of the mesodermal probe, activity was detected in the EMSA with both the M and E probes, suggesting that the two binding activities maybe associated. Another view would be that there is no specific protein, however, UV crosslinking data presented in chapter 7 suggests otherwise.
Towards the Final Protein Purification Protocol

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>126</td>
</tr>
<tr>
<td>6.2</td>
<td>DNA Affinity Purification</td>
<td>127</td>
</tr>
<tr>
<td>6.3</td>
<td>Optimisation of Experimental Conditions</td>
<td>132</td>
</tr>
<tr>
<td>6.4</td>
<td>Radioactive Labelling of the Isolated Proteins</td>
<td>144</td>
</tr>
<tr>
<td>6.5</td>
<td>Discussion</td>
<td>150</td>
</tr>
<tr>
<td>6.6</td>
<td>Summary</td>
<td>152</td>
</tr>
</tbody>
</table>
6.1 Introduction

Whatever the reasons are behind the inconsistency of the purification procedures that have been examined, it was clear that pursuing a traditional approach towards isolating Xsma promoter binding proteins was fraught with problems. Although the binding activities were very robust in the wide range of conditions tested using the EMSA, it seemed that this stability was lost during the process of purification. This was most striking during the majority of DNA affinity chromatography attempts when the binding activities were removed from the sample but could not be detected in the flow through or any matrix eluates. After considering these data together with the batch dependent variation in activity and the time constraint imposed by the stability of the protein, a new variation of the DNA affinity chromatography was developed based on a procedure described by Peck et al. (1994).

6.2 DNA Affinity Chromatography

The strategy devised to overcome the difficulties described above was to try and purify the binding proteins attached to their specific sequence. This could be done using an oligonucleotide attached to a biotin residue by means of a disulphide bridge. EMSA complexes could be formed under appropriate conditions, separated from the bulk of the proteins by means of streptavidin coated magnetic beads and purified on SDS-PAGE. Ultimately it was envisaged obtaining sufficient protein to give a Coomasie stained band that could be sequenced.

An oligonucleotide containing the binding site of interest was manufactured to contain a primary amine and a 12 carbon spacer arm at the 5' end. This was annealed to a complementary strand, radiolabelled, and then derivatised using sulfosuccinimidyl-2-
(biotinamido) ethyl-1,3-dithiopropioate (NHS-ss-Biotin) which forms a stable amide bond with the amine. The probe then contained, a biotin moiety at the 5' end, a disulphide bond, the spacer arm and a radiolabelled binding site (see appendix A). This meant that once the proteins had bound to the DNA, they could be purified on the oligonucleotide by virtue of the interaction of their biotin moieties with streptavidin coated magnetic beads.

Thereafter, the protein-DNA complex could be removed from the beads using DTT to reduce the disulphide bond. This has a special value in that DTT removes only proteins associated with the probe, and not proteins non-specifically attached to the streptavidin. The method on which this variation is based, used the restriction enzyme Eco RI to remove the protein-DNA complex from the beads for use in an in vitro transcription assay (Peck et al., 1994).

This new method has several advantages over the one used in chapter 5. To begin with the initial reaction is a scaled-up EMSA using a cleavable biotinylated probe, in place of the standard EMSA probe, which can bind the proteins under standard conditions. The DNA-protein complex can then be bound to streptavidin coated magnetic beads and washed in EMSA buffer to remove non-specific contamination. Addition of 0.2 M DTT should then release the DNA-protein complex which can then be further purified on a 6 % (w/v) PAG.

This new method was tested successfully using two probes that contained either the putative mesodermal or ectodermal control element (Bio-M and Bio-E), and a probe that contained both of these elements (Bio-P: appendix A). The specificity of the interactions occurring during EMSA was demonstrated by performing the procedure in the presence of specific or non-specific DNA (figure 6.1b and 6.2). In figure 6.1a the binding activity of M seemed to be less specific than E or P, however, subsequent experiments have satisfied us that the binding protein is more specific than suggested here (M. Sargent, personal communication).
Figure 6.1  Isolation of Specific Protein-DNA Complexes Using Streptavidin Coated Magnetic Beads. The isolation of the binding activities associated with (a) the M region, and (b) the E region. The specificity of the interactions was challenged by the effect of the specific and non-specific competition. (See appendix A for the sequence of the probes).
Chapter 6 - Towards the Final Protein Purification

a) 

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>Bio-M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Extract + Bio-M
Extract After Treatment
Wash
M + Wash
Elute

Specific Competitor
Non-specific Competitor

Free Probe M →
Free Bio-M →

b) 

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>Bio-E</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Extract + Bio-E
Extract After Treatment
Wash
E + Wash
Elute

Specific Competitor
Non-specific Competitor

Free Probe E →
Free Bio-E →
Figure 6.2 Isolation of Specific Protein-DNA Complexes Using Bio-P

Isolation of the binding activities associated with the region that contains both the mesodermal and ectodermal elements. The specificity of the interactions was demonstrated by the effect of the specific and non-specific competition.

(See appendix A for the sequence of the probe).
<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>Bio-P</th>
<th>Extract + Bio-P</th>
<th>Extract After Treatment</th>
<th>Wash</th>
<th>P + Wash</th>
<th>Eluate</th>
<th>Specific Competitor</th>
<th>Non-specific Competitor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>x10</td>
<td>x100</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>x50</td>
<td>x100</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
<td>x100</td>
<td>x100</td>
</tr>
</tbody>
</table>

Free Bio-P →
Free Probe P →
6.3 Optimisation of Experimental Conditions

Once it was clear that the method was promising, experiments were performed to optimise each aspect of the procedure to ensure that the final protocol would yield the maximum amount of protein for subsequent peptide sequencing.

6.3.1 Biotinylation of the Probes

In order to limit the loss of synthesised probe it is imperative that the biotinylation reaction occurs at its maximum rate. The NHS-ss-Biotin reagent was prepared by dissolving the solid in 100 mM borate buffer, pH 8.5. When the reagent is present at concentrations above 6 mg ml⁻¹ it forms a gel, but this can be rectified with the further addition of buffer. Three solutions of the reagent (6 mg ml⁻¹, 2 mg ml⁻¹ and 0.6 mg ml⁻¹) were prepared, and mixed 3:1 (v/v) with a 2 pmol stock of TFIIIA probe. After 16 hours at room temperature the efficiency of the reaction was assessed by measuring the amount of probe that could bind to the streptavidin coated beads, using Cherenkov counting. It was expected that the biotinylation would improve with an increase in the amount of reagent present. However, the amount of probe bound to the beads in each of the above concentrations was 39, 35 and 48%, respectively. Similar results were found in many other experiments and were not improved by the further addition of the reagent. Although this demonstrated that the concentration of the reagent was not limiting, it also showed that at this stage more than 50% of the radiolabelled DNA was unable to bind to the beads and would be competing with the biotinylated probe for the proteins.

6.3.2 Optimising the Binding of the Probes to the Beads

It has been reported that the binding of biotinylated DNA to streptavidin coated beads is favoured by increased salt concentrations (Huang et al., 1994; Dynal, 1995). Since an increase in salt concentration may de-stabilise the protein-DNA complex, an experiment
was performed to determine the maximum concentration of \( \left[ \text{NH}_4 \right]_2\text{SO}_4 \) that could be used without affecting the binding activity. The biotinylated probe Bio-M was added to EMSA reactions containing a range of \( \left[ \text{NH}_4 \right]_2\text{SO}_4 \) concentrations (0.05 to 0.4 M). The probe was recovered using the magnetic beads, as described above, and the percentage of the total input probe bound to the beads was determined using Cherenkov counting. The protein-DNA complexes were released by the addition of DTT, and resolved on a 6 % (w/v) PAG (figure 6.3a). When there was no extract present, 48 % of the probe bound to the beads, reflecting the maximum achieved in section 6.3.1. When extract was added to the EMSA reaction, the amount of bound probe dropped to 26 %, suggesting that some component of the extract, such as indigenous biotin or avidin, was interfering with the interaction. Over the range of salt tested in this experiment the amount of probe bound to the beads doubled from 26 % to 52 %, reflecting the reported increase in the efficiency of binding. At the same time, the yield of protein-DNA complex released from the beads at 0.1 M salt was equal to that seen in the control, which contained no \( \left[ \text{NH}_4 \right]_2\text{SO}_4 \), above 0.1 M the yield was reduced until at 0.4 M there was none at all. Figure 6.3b indicates the optimum \( \left[ \text{NH}_4 \right]_2\text{SO}_4 \) concentration for binding the probe to the beads was 0.1 M: the protein-DNA interaction was unaffected, while 48 % of the total probe was bound.

The binding capacity of the streptavidin coated beads is reported to be dependent on the length of the DNA fragment, with twice as many copies of a 500 bp fragment binding than a 1000 bp fragment: the overall mass of DNA bound remains approximately constant (Dynal, 1995). However, in this case the reason why all the probe (22 - 37 bp) does not bind to the beads is unknown and is a potential drawback of the method. Factors that must be considered are, (a) the extent of derivatisation of the 5' end with the dodecylamine residue; and (b) the intrinsic limitation of the NHS-ss-Biotinylation reaction.
Figure 6.3  The Effect of (NH₄)₂SO₄ on the Capacity of Streptavidin Coated Magnetic Beads to Bind Bio-M. (See text for details).
Chapter 6 - Towards the Final Protein Purification

a) Binding Activity -

\[(\text{NH}_4)_2\text{SO}_4 (\text{M})\]

Bio-M

Free Probe -

b) Amount of Bio-M Bound to the Streptavidin Coated Magnetic Beads (%)

Relative Binding Activity (U)
It seems unlikely that the capacity of the beads is limiting. In other experiments it has not been possible to increase the amount bound from equivalent aliquots of a biotinylated oligonucleotide, by increasing the concentration of the beads. The possibility of concatamerising oligonucleotides to provide more binding sites per biotin residue has been considered, but it has been shown that there is a marked deterioration in the efficiency of binding with increasing length (Dynal, 1995; Huang et al., 1996). However, even a five fold gain in capacity would be useful.

### 6.3.3 Optimising the Starting Material

An unfertilised egg extract was brought up to an \( \left( \text{NH}_4 \right)_2 \text{SO}_4 \) concentration of 1.8 M (50 % saturation) by the addition of an equal volume of a saturated salt solution. This was left on ice for 15 min. and then centrifuged at 10,000 g for 10 min. to separate the precipitated protein from the supernatant. The \( \left( \text{NH}_4 \right)_2 \text{SO}_4 \) concentration of the supernatant was then increased to 3.6 M (~100 % saturation) to precipitate the remaining protein. The pellet was resuspended in extract buffer and the cocktail of protease inhibitors, to the extract's original volume. Any insoluble particles were spun down after five min, and the supernatant adjusted to a neutral pH with the addition of 1 M ammonia. The supernatants were split into two, and the binding activity in one was concentrated using the biotinylated ectodermal probe, Bio-\( \text{P} \) in the manner described above. The precipitates were resuspended in extract buffer and, along with the rest of the supernatants, were tested using the Bio-\( \text{P} \) probe in a normal EMSA. These were resolved on a 6 % (w/v) PAG alongside the DTT eluates (figure 6.4a).

The extract prepared for this experiment had relatively poor activity, but this was successfully bound to the beads using the biotinylated probe. Although the precipitate from the 1.8 M (50 % (w/v)) cut contained substantial activity, the supernatant did contain
a discrete activity. However, this was not concentrated by using the beads. The supernatant from the 3.6 M (100 % (w/v)) cut contained no activity, as it had precipitated. The activity in the precipitate was bound to the beads and concentrated by at least five fold.

The tubes containing the various fractions were frozen by immersion in a mixture of ethanol and dry ice, and then stored at -70 °C overnight. The fractions were thawed at room temperature for one hour and then re-tested using EMSA (figure 6.4b). Surprisingly, given earlier experiences, the binding activities were unaffected by the freeze-thaw cycle, and some even appeared stronger (* labelled *). However, the same final concentration of the precipitate from the 100 % (w/v) cut was achieved.

A smaller binding activity which was present after the original (NH₄)₂SO₄ precipitations (figure 6.4a, labelled #), also appeared in the thawed extract (fig. 6.4b, #). This maybe a consequence of precipitation as particles had come out of solution during the freeze-thaw cycle.

This experiment demonstrates the inconsistencies that have plagued the preparation of active extracts. The original extract contained relatively little binding activity but did not inactivate at -70 °C, whereas in other experiments higher levels of binding activity have been completely lost. In some experiments various concentrations of the cryoprotectant glycerol, and other methods of freezing were investigated without obvious benefit. The presence of ammonium sulphate appears to have some beneficial effect. However, it is clear that the new DNA affinity chromatographic procedure concentrates any binding proteins that are present, and that the 100 % (w/v) precipitate can stored for use in
Figure 6.4  Ammonium Sulphate Fractionation of UFE Extract.
The binding activity present in the 1.8 and 3.6 M ammonium sulphate cuts, (a) when tested immediately, and (b) after "snap-freezing" and storage at -70 °C. (See text for details).
The uneven mobility of the free probe is caused by the differing ionic strength of the salt fractions.

Key:  Ex. = UFE Extract
       b = Binding activity recovered using the streptavidin coated magnetic beads.
       ppt = Binding activity in the resuspended precipitate.
       sup = Binding activity in the supernatant.
Chapter 6 - Towards the Final Protein Purification

**Binding Activity**

<table>
<thead>
<tr>
<th>Bio-EP Ex. b</th>
<th>ppt</th>
<th>sup</th>
<th>b</th>
<th>0.0 M</th>
<th>0.0 M</th>
<th>0.0 M</th>
<th>0.0 M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.6 M</td>
<td>3.6 M</td>
<td>3.6 M</td>
<td>3.6 M</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ammonium sulphate treatment</td>
<td>Ammonium sulphate treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Free Probe**

b)

<table>
<thead>
<tr>
<th>Bio-EP Ex. b</th>
<th>ppt</th>
<th>sup</th>
<th>b</th>
<th>0.0 M</th>
<th>0.0 M</th>
<th>0.0 M</th>
<th>0.0 M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.6 M</td>
<td>3.6 M</td>
<td>3.6 M</td>
<td>3.6 M</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ammonium sulphate treatment</td>
<td>Ammonium sulphate treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Free Probe**
subsequent experiments. In subsequent experiments a 40 - 90 % (w/v) \( \text{NH}_2\text{SO}_4 \) cut has been used to produce the starting material.

### 6.3.4 The Removal of Contaminating Proteins

The beads were routinely pre-washed in 0.1 % (w/v) Triton and 0.25 M NaCl to remove all traces of the storage buffer, and to reduce the opportunities for non-specific interactions with the beads. Once the DNA-protein complexes from the EMSA reactions had been bound to the beads, it was necessary to remove any non-specific proteins associated with them by washing with the highest concentration of salt compatible with the stability of the complex. The exact reduction in binding activity after washing in 0.3 M salt can not be estimated since the probe was not in excess, but very little of the remaining activity was lost when the salt concentration was increased to 0.9 M (figure 6.5). The amount of probe bound to the beads, which was monitored by Cherenkov counting, remained constant throughout the experiment. Further investigations indicated that washing in 0.1 M salt, with or without the addition of 10 \( \mu \text{g ml}^{-1} \) of non-specific (lambda) DNA, left slightly more binding activity than at 0.3 M NaCl. It therefore became standard practice to wash the beads, and attached complexes, in the lower salt concentration together with non-specific DNA.

### 6.3.5 Factors Affecting the Cleavage of the Disulphide Bond and Non-Specific Binding to the Magnetic Beads

The disulphide bond in the biotinylation reagent can be broken very easily by low concentrations of DTT, but curiously the bound DNA was not easily released from the streptavidin coated beads. While the conditions for the release of the biotinylated probes from the beads had to be optimised, it was clear that probes could adhere to the beads through sites other than the biotin residue. This "non-specific" binding could be reduced
by, (a) washing the beads in a blocking buffer; (b) phenol extraction of the probe; and, (c) the elimination of spermidine from the EMSA.

If the gel purified probe DNA was extracted with phenol/chloroform (1:1 (v/v)) prior to its biotinylation, non-specific binding was reduced, suggesting that some of the probe was associated with contaminants that either interacted directly with the streptavidin beads, or interfered with the reduction of the disulphide bond by DTT. The non-specific binding could be reduced by a factor of four by omitting spermidine from the reaction buffer. This did not affect the isolation of the proteins which was surprising as it has been reported that spermidine improves protein-DNA interactions in assays such as the EMSA (Panagiotidis et al., 1995), and the initial characterisation of the EMSA conditions at the beginning of this investigation had indicated that spermidine was beneficial (see table 3.1).

Cleavage of the disulphide bonds in the bound protein-DNA complexes is affected by, (a) the temperature, (b) the concentration of the reducing agent, (c) the pH, and (d) steric factors, relating to the interaction of biotin with streptavidin, that may protect the disulphide bond from reduction.

Reduction of the disulphide bond by DTT will work better at 37 °C, but 22 °C was preferred in order to minimise possible proteolytic activity. The concentration required of 0.2 M is many orders of magnitude greater than that required for breaking disulphide bonds in solution (Cleland, 1964), presumably indicating steric hindrance to the reagent reaching the bond. A high pH (pH8) was essential for the action of DTT. An alternative reducing agent, bis(mercapto) ethylsulphone, is reported to cleave disulphide bonds more effectively because of its smaller size and hydrophobicity, and so in theory should remove the apparently more firmly attached material (Lamoureux and Whiteside, 1993). It is insoluble in 0.1 M borate and has to be dissolved in DMSO. However, it was found that
Figure 6.5  Optimum Conditions for Washing the Streptavidin Coated Magnetic Beads.
NaCl concentrations in washes on the yield of protein-DNA complex released by DTT.
(See text for details).

Figure 6.6  An Attempt to Elute Binding Activity from the Streptavidin Coated Magnetic Beads.
Aliquots of a Bio-P complex formed under standard conditions, was bound onto magnetic beads. Sample 1 was eluted with DTT (Lane 1). Sample 2 eluted with 2 M NaCl at room temperature. The beads were eluted with DTT (Lane 2a). The 2 M eluate was diluted 20 fold, fresh Bio-P probe and beads added, and the process was repeated (Lane 2b). Sample 3 was treated as for sample 2, but on ice (Lanes 3a and 3b).
Chapter 6 - Towards the Final Protein Purification

<table>
<thead>
<tr>
<th>NaCl (M)</th>
<th>Bio-EP</th>
<th>0.0</th>
<th>0.3</th>
<th>0.6</th>
<th>0.9</th>
</tr>
</thead>
</table>

Binding Activity -

Free Probe -

<table>
<thead>
<tr>
<th>Bio-P</th>
<th>1</th>
<th>2a</th>
<th>3a</th>
<th>2b</th>
<th>3b</th>
</tr>
</thead>
</table>

Binding Activity -

Free probe -
it is little more effective than DTT at 22 °C and pH 8.0, as it is only soluble up to 10 mM.

6.3.6 Elution of the Proteins with Salt

An experiment was performed to establish the feasibility of eluting the proteins from the DNA attached to the beads using a high concentration of salt. The probe Bio-P was used to isolate the activity from the stored fraction using the beads, as previously described (figure 6.6). The amount of activity isolated from the extract following the normal affinity method and eluting with DTT is shown in lane 1. After the treatment with 2.0 M salt, at room temperature or on ice, there is very little protein left on the DNA (lanes 2a and 3a). However, although the salt treatment had obviously stripped the proteins from the DNA, no activity could be detected in the diluted eluates (lanes 2b and 3b). The experiment indicates strongly that the protein complex is irreversibly disrupted by salt.

6.4 Radioactive Labelling of the Isolated Proteins

The proteins isolated using the above procedure were radioactively labelled by adding \(^{35}\text{S}\)-sulphur labelling reagent (\(^{35}\text{SLR}:\ \text{Amersham}) to the eluate, and incubating the reaction at room temperature for 30 min. This reagent forms a covalent bond with amino and sulphydryl groups within the proteins, but unfortunately the manufacturer also reports that it reacts with thiol reagents such as DTT, and buffers containing a free amino group such as Tris. The labelling reaction can proceed in the presence of DTT if the labelling reagent is used in excess, although inhibition is still evident (figure 6.7a). This problem was resolved (see below) by precipitating the DTT released protein with acetone, in the presence of dextran as a carrier. The manufacturer suggests that the reaction should be carried out in a 0.1 M borate buffer (pH 8.5), and terminated with the addition of 100 µl of
Figure 6.7 (a) The Effect of DTT on the Radiolabelling of Proteins by $[^3]S$-sulphur labelling reagent.
The proteins were isolated from UFE extract using Bio-P (see appendix A), and radiolabelled
using 1 μl of $[^3]S$-SLR. (See text for details).

Aliquots of BSA and Cytochrome C (5 mg ml$^{-1}$) were radiolabelled using 1 μl of $[^3]S$-SLR.
The aliquots were made up to 100 μl with the addition of dH$_2$O, and then 1 and 5 μl samples,
(i) and (ii) respectively, were resolved on a 10 % (w/v) SDS-PAG. (See text for details).
a) kDa Markers +DTT -DTT

200 -
97 -
66 -
46 -
30 -
21 -
14 -
12 -

b) - Toluene + Toluene
(i) (ii) (i) (ii)

BSA -
CytC -
0.2 M glycine. In this investigation a NaHCO₃ buffer was used because it is less insoluble in the acetone used to precipitate the eluted proteins (see below), and the reaction was terminated using 0.2 M ethanolamine. The proteins were resolved with SDS-PAGE.

³⁵SLR is supplied in a toluene solution which is immiscible with water. However, it was discovered that this is of no consequence as the labelling reaction proceeds in the presence of the solvent, as demonstrated in figure 6.7b in which aliquots of cytochrome C and BSA were labelled equally well in the presence and absence of the toluene.

To obtain material that could be labelled using ³⁵SLR, a Bio-M complex was formed under standard conditions and then isolated using the magnetic beads. To demonstrate that the proteins were specific for the mesodermal probe, the isolation was also performed in the presence of a non-specific oligonucleotide, and another oligonucleotide containing a section of 5s RNA that includes the TFIIIA binding site. An inspection of the sequence indicated there was a substantial similarity to the region of the Xsna promoter under investigation, which would be expected to act as a specific competitor for the proteins involved in the binding activities (figure 6.8a). Therefore, the TFIIIA oligonucleotide could be used to show the specificity of the isolated proteins.

The material isolated using the beads was released with the addition of 0.2 M DTT, and precipitated in two volumes of acetone and 1 % (w/v) dextran (Pharmacia) at -70 °C overnight to remove the reducing agent, before the proteins were labelled with 2 µl ³⁵SLR. As a control, one aliquot containing 1 µg of BSA in 0.2 M DTT and another containing just 1 µg BSA alone, were also precipitated and labelled. The labelled proteins were then resolved on a 10 % (w/v) SDS-PAGE (figure 6.8b). The proteins isolated were sequence specific as 0.2 µg of TFIIIA oligonucleotide suppressed binding of protein to the probe, but were unaffected by the same amount of non-specific oligonucleotide. Precipitating the
Figure 6.8 Demonstration of a Restricted Group of Proteins Bound to Bio-P.

(a) Specific Competition Using the TFIIIA Oligonucleotide.

The isolation of the protein complexes associated with the mesodermal and ectodermal regions was performed using Bio-M and Bio-E, in the presence of decreasing concentrations of either, (i) an oligonucleotide containing random non-specific sequence, or (ii) an oligonucleotide containing the TFIIIA binding site from the 5s RNA promoter.

Concentrations used = 500, 50 and 5 fmol. The TFIIIA oligonucleotide competes specifically with the probes for the complexes, and was therefore regarded as the "specific competitor".

E element: \text{ggaggggccacgatc}
TFIIIA Oligo: \text{cgggcctggttagtacctggatggagaccgcc}
M element: \text{agctgcatattagt}

(b) The Eluate from Bio-P Radiolabelled by $[^35]S$-SLR.

The proteins were also isolated using Bio-P in the presence of the TFIIIA and non-specific oligonucleotides in order to demonstrate their specificity for the probe. (See text for details).
Chapter 6 - Towards the Final Protein Purification

149

a) + Unfertilised Egg Extract

Bio-M Bio-E

Binding Activity

Free Probe

Key

= TFIIIA Oligo. = Non-specific Oligo.

b) Bio-P BSA BSA (DTT) + TFIIIA + Non-specific Oligo. + Extract Extract

Free Probe

kDa

- 97

- 66

- 46

- 30
DTT eluate with acetone to remove the reducing agent, was evidently successful as BSA in the presence of DTT was more strongly labelled than the BSA which was free of DTT. There are several weak bands of various sizes labelled from the material isolated using Bio-P, but there four strong bands visible: there are three close bands of about 97 kDa and one of approximately 50 kDa. The BSA which was precipitated from the DTT solution exhibits a larger molecular weight than the BSA precipitated in the absence of DTT presumably as a result of derivatisation. The reasons for this are unclear but it suggests that the molecular weights determined by this approach are likely to be over-estimated.

TFIIIA is a protein of 38 - 40 kDa in its undegraded form, and therefore it is unlikely that Bio-M and Bio-E have mistakenly isolated TFIIIA.

6.5 Discussion

The DNA affinity isolation method described above involves reduction of a disulphide bond positioned between the biotin moiety and the protein binding site, to release the protein-DNA complex from the matrix. It is similar to other methods in which either a restriction enzyme (Peck et al., 1994), or a photo-cleavable bond (Olejnik et al., 1995) are used to release the protein-DNA complex from a matrix.

Having established that this approach was realistic it was essential to optimise the reaction conditions so that it could be scaled up efficiently to obtain enough material for peptide sequence analysis. Disappointingly the biotinylaion of the radiolabelled probes is inefficient in that less than 50 % of the DNA molecules are modified even when a vast excess of the reagent is present. This suggests that the oligonucleotides may need to be concatamerised to increase the number of binding sites per biotin residue, although the optimum length of the concatamers will have to be established as it has been shown that long biotinylated DNA molecules (> 1 Kb) have a substantially reduced capacity to bind to streptavidin coated beads (Dynal, 1995; Huang et al., 1996).
Non-specific interactions between the biotinylated probes and the beads has been reduced by pre-washing the beads in a blocking buffer, purifying the DNA before biotinylation and by eliminating spermidine from the reaction buffer. The efficiency with which the biotinylated probes bind to the beads is halved in the presence of the extract, but this can be reversed by increasing the salt concentration. It has also been reported that mass transfer considerations limit the rate of binding to streptavidin (Huang et al., 1996), and so ideally the smallest beads available would probably be best. It is claimed that streptavidin binds to the surface of polystyrene latex beads in a mono-layer, allowing a less sterically hindered interaction with the biotin moieties (Huang et al., 1996). However, these beads are not magnetic, and it is unclear whether their density is sufficient to allow their isolation by centrifugation in moderately dense solutions.

Approximately 50% of the protein-DNA complexes can be isolated using the streptavidin coated magnetic beads, and only 60% of these (30% of the original) can be separated from the beads for subsequent procedures. About 1.0 pmol of protein is needed to obtain an N-terminal sequence, and so almost 10.0 pmol of protein-DNA complexes has to be present in the starting material. It has already been calculated that the amount of binding activity per egg is approximately 5 fmol. As previously described, this will vary between different batches of unfertilised eggs but in the region of $2 \times 10^6$ eggs would be required to make enough extract for the full scale purification. Some increase in the efficiency of purification may be made by using concatamers of the probe. Problems will arise if a large proportion of the unfertilised eggs contain a low amount of activity, but if it is possible to use concatamerised DNA this target should be realistic, particularly if it remains feasible to freeze the prepared extract. It will be possible to store the isolated
proteins until enough has been stockpiled to proceed with the SDS-PAG and subsequent sequencing.

When the proteins were removed from the DNA using a high concentration of salt, it was impossible to detect binding activity in the eluate, even when the salt concentration was diluted to an innocuous level. Exactly the same observation was made using the first DNA affinity method (see chapter 5). As this new approach provides the opportunity to remove the DNA from the beads by reducing the disulphide bond, it was possible to confirm that after being removed from the nucleic acid, the binding proteins had been rendered inactive by the salt. This suggests that a multi-subunit structure has been disrupted that cannot re-associate at low subunit concentrations.

### 6.6 Summary of Results

1) A novel approach to DNA affinity purification has been devised.

2) The method has been optimised as a preliminary to scaling up the procedure for peptide sequencing.

3) High salt concentrations appear to damage the protein-DNA complex, and may explain the failure of the salt elution based purification procedures used previously.

4) It seems possible to isolate enough protein to obtain some N-terminal sequence.
Biochemical Characterisation of the *Xsna* Promoter Binding Activities

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>154</td>
</tr>
<tr>
<td>7.2</td>
<td>Determination of the Size of the Complexes Involved</td>
<td>154</td>
</tr>
<tr>
<td>7.3</td>
<td>Analysis of Individual Peptides</td>
<td>155</td>
</tr>
<tr>
<td>7.4</td>
<td>A Comparison of the M and E Complexes Using a Protease Clipping Assay</td>
<td>166</td>
</tr>
<tr>
<td>7.5</td>
<td>Discussion</td>
<td>169</td>
</tr>
<tr>
<td>7.6</td>
<td>Summary</td>
<td>169</td>
</tr>
</tbody>
</table>
7.1 Introduction

Once the presence of sequence specific DNA binding activities in a crude extract was established, the factors could be characterised with respect to molecular mass, subunit composition, possible interactions with non-DNA binding factors and specificity (see chapter 4). Preliminary experiments using a concatamerised mesodermal element in an affinity selection (see chapter 5) indicated that the M and E binding activities may co-purify, possibly indicating that there was an interaction between the two binding activities.

7.2 Determination of the MW of the Complexes Involved

The size of the proteins responsible for the binding activities can be estimated using gel filtration under conditions which should not disrupt their structure. Gel filtration media contain pores into which relatively small molecules can diffuse and which exclude slightly larger molecules. This property makes it possible to fractionate a population of proteins on the basis of their size when they pass through a column of the matrix as a band. Using almost spherical globular proteins, size is proportional to molecular weight and their elution volume is an inverse logarithmic function of size. Therefore, the molecular weight of a protein complex responsible for a detectable binding activity can be estimated from its elution volume compared to that of standard proteins.

A suitable gel matrix is one in which the elution volume of the factor is midway through its fractionation range. Sephacryl S-300 HR (Pharmacia; fractionation range = 10-1500 kDa) was used to determine the approximate size of the complex. From this information Sephacryl S-400 HR (Pharmacia; fractionation range 20 - 8000 kDa) was chosen for a finer analysis. A slow flow rate provides the best resolution but can be
counter productive if the material is unstable. A flow rate of approximately 6 ml hr\(^{-1}\) was used without excessive loss of activity. The recommended sample size of 1 % of the bed volume (Pharmacia) was used to maximise resolution. A 50 ml column was selected as a peak of binding activity was detectable within only three fractions (figure 7.1a). This assay was repeated several times, and the main binding activity, observed using EMSA with probe P, was consistently about 370 kDa (figure 7.1b). Occasionally a smaller binding activity of approximately 70 kDa was observed, which may be analogous to M\(_S\) (see chapter 3). The large apparent size of the major binding activity, strongly suggests that it is a multi-subunit complex. It is possible that the action of a phosphotase in fractions 30 - 34 (seen by the degradation of the probe) is masking a second peak. However, more recent data would suggest this is not the case (M. Sargent, personal communication).

### 7.3 Analysis of Individual Peptides

Irradiation of DNA using 260 nm ultra violet light causes the formation of purine and pyrimidine free radicals that are capable of forming covalent bonds with other molecules that are physically close. If the radiolabelled DNA is irradiated while in contact with a peptide the free radicals are able to react irreversibly with the amino acids, effectively radiolabelling the peptide. Therefore, after resolution by SDS-PAGE, any peptide in direct contact with the DNA can be identified and its molecular weight estimated. Radiolabelled DNA with thymidine residues substituted by bromodeoxyuridine (BrdU) is more efficient at forming covalent bonds with attached proteins than unsubstituted DNA (Hutchinson, 1973; Lin and Riggs, 1974). As BrdU absorbs UV at a higher wavelength (\(\lambda = 300-310\) nm) the effects of irradiation are specific to BrdU and do not affect the chromophores of the protein. Other nucleotide analogues, such azidodeoxyuridine (Evans and Haley, 1987) and 5-iodouracil (Willis et al., 1993), are reported to be substantially more efficient than BrdU, but are not readily available.
Figure 7.1 (a) The Elution from a Sephacryl S400 Gel Filtration Column.
Experimental details: 50 ml column, equilibrated in extract buffer. Flow rate = 6 ml/hr; 1 ml fractions. Fractions tested using probe P (see appendix A).

\( P_L \) = Large binding activity. \( P_S \) = Small binding activity.

(b) Molecular Weight Determination.
The binding activity was quantified using a densitometer and are plotted beneath the calibration curve. The peak of \( P_L \) activity corresponds to 370 kDa.
(Error +/- one fraction = +140 kDa / -110 kDa)
Chapter 7 - Biochemical Characterisation of the Xsna Promoter Binding Activities

Free Probe

fraction Numbers

Extract 50% (NH₄)₂SO₄ extract

Pₐ

Pₘ

Thyroglobulin

370 KDa

BSA

Cyt. C

Molecular Mass (KDa)

Activity (u)

Fraction Number

12 14 16 18 20 22 24 26 28 30 32 34 36 38 40 42 44

10000

1000

100

10

1

100

10

1
The traditional method of performing this experiment involves the irradiation of an EMSA reaction containing the substituted radiolabelled probe, followed by resolution of the samples on SDS-PAGE. This was attempted on several occasions but was ineffective, probably because of dispersion or absorption of radiation by components of the extract.

To overcome this problem an alternative approach was taken in which the EMSA reactions, performed using BrdU substituted Bio-M and Bio-E probes (appendix A), were resolved on a 6 % (w/v) polyacrylamide gel prior to irradiation (Barnett et al., 1995). The gel was then placed above some ice and irradiated for 30 min. The lanes of the gel containing the free probe and the protein-DNA complexes were excised and placed in the stacking gel of a 10 % (w/v) SDS-PAG, and resolved alongside molecular markers (figure 7.2). Both probes yielded a similar pattern of three close bands: their sizes were approximately 97 kDa, 90 kDa and 85 kDa. The larger band may represent material retarded at the boundary of the spacer and resolving gels. While these results were less than perfect, they indicated that the difficulties with the direct approach could be overcome with a partially purified extract.

A far clearer result was achieved by exploiting the reducible biotinylated probes and streptavidin coated magnetic beads in conjunction with the traditional UV irradiation approach. Protein-DNA complexes were isolated from unfertilised egg and stage 14 embryo extract using BrdU substituted Bio-M and Bio-E (appendix A), and the beads as described in chapter 6. After the disulphide bond had been reduced by DTT, the eluates were irradiated for 30 min. above ice, and then resolved on a 10 % (w/v) SDS-PAG alongside molecular markers (figure 7.3). One band of about 97 kDa was seen with both probes in the unfertilised egg extract (figure 7.3a), with the ectodermal band appearing slightly larger. There were three smaller mesodermal bands of about 62 kDa, 58 kDa and
Figure 7.2  The Two Dimensional Resolution of UV Crosslinked Protein-DNA complexes. EMSA reactions containing Bio-M and Bio-E were resolved in the first dimension on a 6 % (w/v) PAG alongside free probe. These lanes were excised, irradiated with UV light (λ = 300-310 nm) and placed into the spacer gel of a 10 % (w/v) SDS-PAG, for resolution in the second dimension. (See text for details).
Figure 7.3 UV Crosslinking after Isolation Using the Paramagnetic Beads.

The binding activities associated with probes Bio-M and Bio-E were isolated using the streptavidin coated magnetic beads from (a) unfertilised egg extract, and (b) stage 14 embryo extract. The eluates were irradiated with UV light ($\lambda = 300-310$ nm) and resolved on a 10% (w/v) SDS-PAG alongside some molecular markers. (See text for details).
39 kDa clearly visible which were not apparent with the ectodermal probe. A similar pattern was observed in the stage 14 embryo extract (figure 7.3b) with the ectodermal band again appearing to be, by a small margin, the larger of the two, however the size of both bands was apparently slightly smaller at around 90 kDa. It is not certain if this is a significant difference. There were no smaller bands visible but there was a faint band above the main mesodermal band, similar in size to the ectodermal band.

Partially purified protein-DNA complexes evidently provide substantial improved material compared with crude lysate for UV irradiation studies.

To assess the efficiency of the crosslinking reaction induced by UV irradiation, a protein-DNA complex from unfertilised egg extract was prepared using the streptavidin coated beads, in the normal way. After the DTT eluate was irradiated it was split into two and 0.1 % (w/v) SDS was added to one half. The samples were then resolved on a 6 % (w/v) PAG. The SDS removed all the non-crosslinked proteins from the UV treated protein-DNA complexes, and allowed a comparison with the total possible amount of binding activity (figure 7.4). In this extract, M binding activity was much weaker than E, but the level of both activities was unaffected by the UV irradiation. SDS treatment of the non-crosslinked protein-DNA complex resulted in the complete ablation of the activity. After similar treatment with the crosslinked ectodermal complex there was binding activity of reduced size with vastly reduced intensity. There was a very faint band with the mesodermal probe, that was probably equivalent to the reduction seen with the ectodermal probe. These experiments showed that UV induced crosslinking of protein to nucleic acids is inefficient, although it is clearly useful to make qualitative observations such as the size of the bound protein. In at least one instance a high yield of protein-DNA conjugate
Bio-M and Bio-E complexes were isolated using the streptavidin coated magnetic beads and eluted with DTT. Aliquots from each sample were UV irradiated. SDS (0.1 % (w/v)) was then added to half of each sample to remove the non-crosslinked proteins from the DNA. All the samples were then resolved on a 6 % (w/v) PAG.

Key:  
+ $E_+$ = Plus extract.  
- = Minus extract.  
+ UV = After UV irradiation.  
+ SDS = After SDS treatment.  
BA* = Crosslinked binding activity after SDS treatment.
Chapter 7 - Biochemical Characterisation of the Xsna Promoter Binding Activities

Bio-M Bio-E
+ Ex - + Ex - (1a) (1b) (2a) (2b) + UV (1a) (2a) (3a) (4a) + SDS + SDS

Binding Activity -

Free Probe -

- BA*
formed as a result of UV irradiation has been purified on SDS-PAG using azidodeoxyuridine substituted DNA (Roux et al., 1989).

7.4 A Comparison of the M and E Complexes Using a Protease Clipping Assay

The treatment of a protein, or a protein complex, with a dilution series of a specific protease, will result in a characteristic fragmentation profile, called a protease clipping assay. The endoproteases Arg C and Glu C, which attack the peptide bond on the carboxy side of either arginine or glutamic acid residues respectively, were used to characterise the protein complexes associated with the M and E binding activities. The probes Bio-M and Bio-E were used with the paramagnetic beads to isolate their binding activities, as described in chapter 6. The protein-DNA complexes were released from the beads by treatment with 20 μl of DTT, after which the eluates were made up to a volume of 160 μl with the addition of 0.05 M Tris, pH 8.0. A dilution series of each protease was prepared, by taking 1 μl of stock (10 ng μl⁻¹) and successively diluting 1:3, and the proteolytic cleavage was initiated by adding 1 μl of each dilution to a 10 μl aliquot of the eluate. The reaction proceeded at room temperature for 15 min., after which the samples were resolved on a 6 % (w/v) PAG (figure 7.5). The Arg C degradation of the M complex was almost identical to that of the E complex: the main binding activity was reduced in intensity as the protease concentration increased, while simultaneously, a slightly smaller activity appeared (figure 7.5a). The Glu C profiles were similar to those of Arg C in that, with the increase in protease concentration, the main activity became weaker to be replaced with a smaller activity (figure 7.5b). At the highest concentration of Glu C there was a further decrease in the mobility of the protein-DNA complex.
Figure 7.5 Protease Clipping Assays.

The degradative profiles after treatment with (a) Arg C, and (b) Glu C.

(See text for details).

Key: - = Minus extract.
     + = Plus extract, without protease.
Chapter 7 - Biochemical Characterisation of the Xsna Promoter Binding Activities

(a) Bio-M Bio-E

Binding Activities

Free Probe -

(b) Bio-M Bio-E

Binding Activities

Free Probe -

Key: § = Arg C  — = Glu C
The most striking aspect of the fragmentation patterns was that for each enzyme the pattern associated with the two probes was almost identical. This suggests that the organisation of the proteins within the M and E protein complexes is very similar.

### 7.6 Discussion

The high apparent molecular weight of the P complex on gel filtration (~370 kDa) relative to the size of the proteins identified by UV crosslinking (~90 kDa), suggest that the binding complexes are multimeric. This may explain the sensitivity of the binding activities to freezing and to salt elution. It may be possible using a protein-protein chemical crosslinking agent such as bis[sulphosuccinimidyl] suberate in conjunction with UV crosslinking of the protein to the DNA, to demonstrate contact between polypeptides within the complex. Samples that have been treated in this manner would contain larger components than those observed after UV crosslinking alone, from which the size of individual components could be calculated.

The apparent similarity in protease fragmentation products of the M and E complexes using two different proteases is quite surprising. It would be expected that these two endoproteases would produce distinctly different patterns (for example, Snape et al., 1991; Roux et al., 1989). However, other data suggest there may be a relationship between the binding activities (see section 5.5).

### 7.7 Summary

1) A large multi-subunit complex of ~370 kDa is responsible for the binding activities observed using EMSA.

2) Both the mesodermal and ectodermal elements are in direct contact with a peptide
of about 97 kDa: the mesodermal peptide is the smaller of the two.

3) In stage 14 embryos, these peptides are apparently 90 kDa.

4) The protease clipping profiles for the M complex are almost identical to that of the E complex, possibly indicating a surprising similarity in organisation.
Chapter 8 - The Regulation of Xsna

The Regulation of Xsna

8.1 Introduction 172
8.2 Xsna Induction is Regulated by PKC\(\alpha\) 174
8.3 Discussion 174
8.4 Summary 179
8.1 Introduction

The ectodermal expression pattern of \( Xsna \) that appears at stage 11 - 11.5 is the first indication that the prospective neural crest is distinguishable from the prospective neural plate (Essex et al., 1993). This expression, which begins as a low arc on the dorsal aspect, outlines the neural plate during subsequent extension of the latter (termed the neural plate border; Essex et al., 1993). By the open neural plate stage, \( Xsna \) is down-regulated in the anterior fold prior to its incorporation into the forebrain, but remains strongly expressed on the lateral edges which later become the neural folds (Essex et al., 1993).

For some time it has been suspected that the patterning of the nervous system in \( Xenopus \) involves signals emanating from the underlying mesoderm and through the plane between the organiser and the ectoderm (Spemann and Mangold, 1924), and several molecular candidates for the inductive signals have emerged (Ferreiro and Harris, 1994). More recently it has been suggested that interactions between the dorsal and ventral ectodermal domains (chapter 1), lead to induction of the prospective neural crest (Mayor et al., 1995; Mancilla and Mayor, 1996)).

Using the specific neural crest marker \( Xslu \) (Mayor et al., 1995), a parologue of \( Xsna \), an expression cloning study was conducted in which it was discovered that the translation initiation factor eIF4AI and eIF4AII could specifically induce the genes of the neural plate border (Morgan and Sargent, submitted). Over-expression of eIF4AI, or the murine homologues mmeIF4AI and mmeIF4AII, up-regulated the neural crest markers \( Xslu \), \( Xsna \) and \( Pax-3 \) (but not \( twist \)), the anterior fold markers \( XANF \) and \( hairy 2A \), and the cement gland marker \( XGC1 \). There was no up-regulation of the neural plate markers \( NCAM \), \( Tubulin-\beta \), \( Hox-B9 \) and \( Xif-6 \), or the mesodermal markers \( Xbra \), \( Xhox-3 \) and \( myo-D \).
RNAse protection analysis showed that eIF4AII was expressed in all tissues at all embryonic stages and that eIF4AII mRNA increased sharply after stage 11, with more than 90% being confined to the dorsal ectoderm (Morgan and Sargent, submitted). eIF4AII was up-regulated by a variety of neural induction regimes but the strength of the signal determined whether neural plate or neural plate border genes were expressed. The minimum amount of each inducer which up-regulated eIF4AII, also up-regulated Xsln and suppressed epidermal markers, but as the amount of inducer was increased, NCAM was expressed and Xsln repressed (Morgan and Sargent, submitted).

The expression of the *Xenopus* protein kinase Ca (PKCa), which has been implicated in the acquisition of competence for neural induction (Otte et al., 1988, 1990, 1991; Otte and Moon, 1992), is transcriptionally up-regulated by mmeIF4AII. Ectopic expression of either eIF4AII or PKCa mRNA sensitised the ectoderm to neural induction by noggin: micro-injection of PKCa reduced the amount of noggin needed to induce NCAM and suppress Xsln to 1% (Morgan and Sargent, submitted). The bis-indolylmaleimide GF109203X is a specific inhibitor of the calcium dependent isoforms of PKC (α, β and γ; Toullec et al., 1991). Ectopic expression of eIF4AII in the presence of GF109203X resulted in the suppression of Xsln, but not EF1α and the inhibitor had no effect on axial mesoderm formation. These data suggest that induction of Xsln is dependent on PKC activity, and that eIF4AII and PKCa are intermediaries in neural crest induction. Since *Xsna* expression is upregulated by overexpression of eIF4AII and PKCa, and is inhibited by GF109203X, it seems likely that the activity of the ectodermal element of the *Xsna* promoter is dependent directly or indirectly on phosphorylation.

Another aspect of the role of phosphorylation came to light in experiments described below in which injections of *Xsna* mRNA were found to induce the endogenous
Xsna gene in animal cap assays. This autoregulatory loop was enhanced in the presence of GF109203X.

8.2 Xsna Induction is Regulated by PKCα

The effect of injected capped Xsna mRNA on the synthesis of the endogenous Xsna gene was studied using RNAse protection. An experiment was performed in which embryos were injected with either 1 or 3 ng of Xsna mRNA. At stage 8-9, animal caps were cut and cultivated to stage 13 in NAM, in both the presence and absence of GF109203X. At stage 13 the RNA was extracted and analysed for Xsna and EF1α using an RNAse protection assay. EF1α is a constitutive zygotic mRNA used as a loading control. The Xsna RNA probe was derived from the 3’ end of the gene which was not present in the injected mRNA, so that it would only detect the endogenous Xsna RNA. Figure 8.1 showed that the injected Xsna induces the synthesis of the endogenous gene. However, in the presence of GF109203X there was a substantially greater synthesis of Xsna. As GF109203X inhibits the calcium activated PKC isoforms, this implied that the absence of phosphorylation permitted greater induction of the endogenous gene. One explanation of this is that in the absence of ongoing PKC activity, phosphorylated proteins that regulate Xsna may be dephosphorylated by phosphatases.

8.3 Discussion

While the observations described here are clear cut, their interpretation is provisional as they must be integrated with a substantial amount of recent data relating to the effect of translation initiation factor eIF4AII. That work (Morgan and Sargent, submitted) suggests that Xsna is upregulated as a result of the activity of PKC. Upregulation of Xsna by
Figure 8.1  An RNAse Protection Assay Demonstrating that Xsna can Regulate its Own Transcription. Embryos were injected with either 1 or 3 ng of Xsna mRNA. At stage 8 - 9, animal caps were cut and cultivated to stage 13 in NAM, in both the presence and absence of the specific inhibitor of calcium activated PKC isoforms, GF109203X. The Xsna RNA probe was derived from the 3' end of the gene which was not present in the injected mRNA, so that it would only detect the endogenous RNA.

Key:  GF = GF109203X
      EF1α = Loading Control
Figure 8.2 The Order of Molecular Events Leading to Neural Crest Induction:

(a) as suggested by the recent work of Morgan, R and Sargent, M. G. (submitted),

(b) incorporating the Observation that Xsna can regulate its own Transcription. As Xsna is probably a transcription factor it could be supposed that the basis of the loop may be an unknown gene whose product activates Xsna only when dephosphorylated.
Chapter 8 - The Regulation of Xsna

(a)

\[ eIF4AII \rightarrow PKC\alpha \rightarrow Xsna \]

(b)

\[ eIF4AII \rightarrow PKC\alpha \rightarrow Xsna \]

Activated by Phosphorylation

Blocked by GF109203X

Unknown Transcription Unit

Xsna Promoter
eIF4AII is inhibited by the PKC inhibitor GF109203X. This suggests a "pathway" leading to neural crest induction, shown in figure 8.2a, with an unknown phosphorylation event required for the transcriptional upregulation (*) of \textit{Xsna}. The data described here indicates an additional element in this pathway in which \textit{Xsna} mRNA upregulates itself, and furthermore, it does so by a process that is enhanced by PKC inhibition (figure 8.2b). \textit{Xsna} expression in activin treated caps cultured until stage 14, shows very little inhibition when co-treated with GF109203X (Sargent, M. G., personal communication), suggesting that the mesodermal component is activated by a pathway that does not involve PKC.

It has been shown that \textit{Xsna} induces \textit{Xslu} in animal caps in a fashion that is consistent with an ectodermal role (Morgan and Sargent, submitted), but it is not possible to rule out an activity that relates to its mesodermal function. These observations do not provide a physical model of the activity of the \textit{Xsna} promoter, but do suggest that it may be regulated in a complex way by phosphorylation.

### 8.4 Summary

1) PKC\(\alpha\) is involved in the regulation of \textit{Xsna} transcription.

2) \textit{Xsna} mRNA can induce the endogenous \textit{Xsna} gene by an autoregulatory loop.

3) The phosphorylation states of \textit{Xsna} regulators appear to be important for the transcriptional control of the gene.
## General Discussion

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1</td>
<td>Binding Activities Associated with the \textit{Xsna} Promoter</td>
<td>181</td>
</tr>
<tr>
<td>9.2</td>
<td>Transcriptional Control of \textit{Xsna}</td>
<td>183</td>
</tr>
<tr>
<td>9.3</td>
<td>Purification of \textit{Xsna} Promoter Binding Proteins</td>
<td>185</td>
</tr>
<tr>
<td>9.4</td>
<td>Similarities Between the M and E Complexes</td>
<td>187</td>
</tr>
<tr>
<td>9.5</td>
<td>Speculation about the Regulation of \textit{Xsna}</td>
<td>188</td>
</tr>
</tbody>
</table>
9.1 Binding Activities Associated with the Xsna Promoter

The Xenopus gene Xsna is expressed during the development of two distinct embryonic cell populations: the mesoderm and the prospective neural crest. The mesodermal (M) expression is first detectable at stage 9, and continues until the histological differentiation of derivatives such as notochord, somites and pronephros (Essex et al., 1993). Ectodermal (E) expression begins at stage 11 in the neural plate border, which is the prospective neural crest (Essex et al., 1993; Mayor et al., 1993). Deletion analysis of its promoter identified two contiguous elements, approximately 100 base pairs upstream from the transcriptional start site, which controlled this differential pattern of expression (Mayor et al., 1993). The removal of the more distal element eliminated almost all mesodermal Xsna expression, while the E component remained, whereas the removal of the proximal element resulted in the complete ablation of all Xsna expression.

I have shown that radiolabelled probes containing the sequences from the M and E elements can be used in the electrophoretic mobility shift assay (EMSA), to detect specific binding activities in extracts prepared from unfertilised Xenopus eggs and embryos. The EMSA has proved to be a versatile assay capable of detecting the activities over a wide range of conditions, and it has been used, alone and in conjunction with other techniques, to characterise the proteins involved and the nucleic acid sequence with which they interact.

A more detailed analysis of the M and E elements was obtained by examining the binding of sets of radiolabelled probes with altered sequences, in EMSA reactions. This data indicated that the regions of the sequence most critical for the protein-DNA interactions were identified as (highlighted bold and underlined):

\[ ^\text{agctgcat} \text{attagttagtgagggg} \text{cacagca} \]
Neither of these sequences correspond to the precise recognition sites of known transcription factors, although the 5' M element does show some similarity to the binding sequence of the HNF family of proteins (Faisst and Meyer, 1992), and in particular the *Xenopus* HNF family member *Pintallavis*, with 8 out 12 bases being conserved (Ruiz i Altaba and Jessell, 1992). However, experiments using the M and E radiolabelled probes in EMSA reactions containing *in vitro* translated Pintallavis, showed no binding activity (data not shown). It is possible that a hitherto unidentified HNF family member is involved in the interaction, but it is equally likely to involve a protein that recognises a similar but specific sequence. It has not escaped notice that the complementary sequence to $^{5'}\text{atta}^{3'}$ is present in the binding sites of many homeodomain proteins (Hayashi and Scott, 1990). It is believed that the sequences outside the core tetramer $^{5'}\text{taat}^{3'}$ are responsible for distinguishing between the various family members (Ekker et al., 1991).

Neither the DMS protection nor the copper-phenanthroline footprinting studies provided the clear representation of the binding sites that was anticipated. This may have been attributable to the variation in activity encountered between batches of extract. It is possible that the amount of protein present in the binding reactions was too close to the lower limit required for a precisely defined footprint using these methods, although it has been reported that if binding activity is detectable in an EMSA reaction, then both these protocols should be feasible (Shaw and Stewart, 1994; Papavassiliou, 1994). However, both of these approaches did confirm that the elements suggested by the deletion analysis are involved in protein-DNA interactions.

The mobility of the individual M, E and P EMSA complexes were very similar. It might be expected that the latter would have lower mobility representing the sum of M and
E, but this is not the case, suggesting that the two activities are not separate modules whose masses are additive, and might indicate that they might create similar structures.

The size of the activity P from unfertilised egg extract was very large (~ 370 kDa), suggesting the presence of a number of subunits. UV crosslinking experiments using unfertilised egg extract, indicated that a peptide of approximately 97 kDa, or a number of peptides totalling this size, make the closest contact with the DNA probe. The possibility that more than one polypeptide of identical size are involved in this cross-link, cannot be eliminated. Protease clipping assays using two endoproteases with different cleavage specificities, showed that the M and E complexes shared very similar sensitivity to protease, that is, the major fragments derived from the complexes were very similar in size. This seems to indicate an unexpected similarity of structure between the two binding activities. It is usually possible to show distinct fragmentation patterns in unrelated DNA binding proteins (Snape et al., 1991; Roux et al., 1989).

9.2 Transcriptional Control of Xsna

A developmental profile of the M, E and P binding activities showed that the proteins responsible are available to interact with the promoter in ovaries, unfertilised eggs and embryos up to stage 35. From stage 13 the level of activity gradually decreases until it stabilises at a constant low level, between stages 18 and 20. The availability of the proteins prior to the start of Xsna expression indicates that the trigger for transcription is unlikely to be their synthesis. A more plausible view is that a post-translational modification of the pre-formed proteins has to occur before they engage in transcription. This could be a phosphorylation, as with CREB (Hoeffler, 1988), a dephosphorylation as with AP1 (Boyle et al., 1991), or a change in the subunit composition of the complex.
For example NFkB becomes active when IkB dissociates from a cytoplasmic complex as a result of phosphorylation (Liou and Baltimore, 1993). \( X \)snaf promoter binding proteins may also change from a cytoplasmic to a nuclear form at the start of transcription. An entirely different possibility is that the chromatin structure is altered to allow unmodified proteins to approach the nucleic acid (Wolfe, 1994b).

When the UV crosslinking experiment was repeated using stage 14 embryo extract, it appeared that the peptide, or peptides, containing the DNA binding domain were smaller than in unfertilised eggs. The significance of this observation is unknown, but it is consistent with changes in phosphorylation which have a marked effect on the apparent molecular weight of a protein on SDS gel (Li and Aderem, 1992). The loss of a subunit from the main complex has already been proposed as a possible mechanism through which the expression of \( X \)snaf might be controlled. At stage 14 most of new \( X \)snaf expression has finished and the detectable promoter binding activity is falling to its eventual low level, and so this apparent decrease in size raises several possibilities; the reduction may be concerned with the activation of transcription at stage 9; the specificity of the complex for the M or E element at stage 11; or the reduction in expression at stage 13.

Under certain conditions an activity with a smaller mobility shift to the main activity, is observed using the M and P radiolabelled probes in EMSA reactions (Mg and Pg), but not using the E probe even within the same experiment. This may be a degradation product, or it may indicate a difference in their the subunit composition.

Until recently it has been difficult to devise model systems in which the ectodermal and mesodermal induction of \( X \)snaf could be distinguished. However, induction of \( X \)snaf by eIF4AII is an ectodermal manifestation of its expression (Morgan and Sargent, submitted), while induction of \( X \)snaf before stage 11 in animal caps treated with activin (Sargent and
Bennett, 1991) would be mesodermal (Essex et al., 1993). By examining the UV crosslinking products obtained with extracts from tissues engaged in specific E or M Xsna synthesis, the changes in UV crosslinking products might be associated with particular developmental events.

### 9.3  Purification of Xsna Promoter Binding Proteins

Generally the binding complexes have been shown to be vulnerable to inactivation after freezing, even in the presence of a cryoprotectant such as glycerol, while other batches of extract have been inactivated by storage even at 4 °C. This instability at low temperatures may reflect the involvement of hydrophobic bonds in maintaining the integrity of the binding complex (Mahler and Cordes, 1966). Long term storage of the extracts was not possible at temperatures above 4 °C, even in the presence of several protease inhibitors, until further investigations revealed that an ammonium sulphate fraction containing the binding activity could be successfully frozen without detrimental effect on complex stability. Indeed, ammonium sulphate fractionation proved to be effective in removing a large amount of the cellular proteins, and has therefore become the first step in the final purification protocol.

The cold sensitivity limited progress of early purification attempts. Hydrophobic interaction chromatography would have been a logical progression after the ammonium sulphate fractionation, but early successes could not be repeated at a realistic scale. This was probably because the complex was inactivated during the increased experimental time. Ion exchange chromatography is a standard early step in many purification protocols, however it proved difficult to optimise the experimental conditions, as on many occasions binding activity was lost. It is possible that the complex was de-stabilised by the charge of
the matrices, but it now seems more likely that a major contributing factor was the exposure to high salt during the elution steps. This was almost certainly the reason why the first approach to DNA affinity purification was unsuccessful, as later experiments indicated.

To overcome this problem, a new variation of the DNA affinity technique was developed in which a radiolabelled probe containing the specific DNA binding motif is biotinylated at one end, so that it can be isolated from an EMSA reaction using streptavidin coated magnetic beads. The innovation was to include a disulphide bond between the biotin moiety and the binding site so that the protein-DNA complex could be released from the beads by DTT avoiding the need to elute with salt. This new approach eliminated from the purification those proteins that bind non-specifically to the streptavidin matrix, and probably also maintained the stability of the complex by preserving its interaction with the DNA. Comparable methods that make use of cleavable bonds to release DNA-protein complexes bound to streptavidin matrices have been described, for example, using a restriction enzyme (Peck et al., 1994) and photo-cleavable bonds (Olejnik et al., 1995).

An extensive investigation has been conducted to optimise all aspects of this technique to facilitate a successful scale-up using sufficient material for peptide sequencing. The important factors include the concentration of the non-specific DNA required to ensure the maximum sequestration of non-specific protein; the conditions for binding the biotinylated DNA to the beads, and their subsequent release; and the preparation of the starting material. The most disturbing problem was the inability to obtain total derivatisation of the probe DNA, as the non-biotinylated DNA will compete
for the promoter binding proteins. The explanation for this may be that the addition of the
dodecylamine residue during the synthesis of the oligonucleotide was inefficient.

The amount of activity present per egg or embryo varies considerably between
batches, but in the most favourable cases approximately 5 fmol per egg or embryo can be
detected. It has been calculated that approximately $2 \times 10^4$ eggs or embryos will be
required as starting material for the purification of 1 pmol of binding protein. By
combining the improvements that have been made with the new affinity technique, it
should be possible to stockpile enough protein to proceed with the N-terminal sequencing.
The final purification step prior to the sequencing will be SDS-PAG.

### 9.4 Similarities Between the M and E Complexes

Several aspects of this investigation point to a close relationship between the M and E
complexes. On one occasion after using a concatemer of the M element in an early DNA
affinity purification, it was discovered that activity for the M and E probes could be
detected in the NaCl eluate. In the other purification methods used, apart from ion
exchange chromatography, the binding activities have consistently co-purified indicating
that they share many biochemical properties. The mobility of the activities detected using
the M and E probes are the same size. The variation in the amount of activity present in
the extracts, and the cold sensitivity of the complexes, affect both M and E in the same
manner: neither activity has been detected in the absence of the other. Perhaps the most
suggestive data is the observation that the protease clipping assays showed similar
fragmentation patterns for both activities.

However, there is some data which indicates there are two distinct binding
activities. They bind to discrete and independent sequence motifs within the M and E
elements. The activity which interacts with probe M can, under certain conditions, fragment to give a characteristic smaller mobility shift, $M_n$, which has never been seen using probe E. The size of the peptides that interact directly with the two DNA binding motifs although close, are distinctly different. The activities do not co-purify after ion exchange chromatography, but this may not be relevant as there were major losses of activity.

This apparent paradox can be resolved if the binding activities are considered to be two versions of the same complex. Each would contain the same subunits apart from possibly one factor which would be solely responsible for the observed differences, including the sequence specificity responsible for the tissue specific expression of the gene.

### 9.5 Speculation about the Regulation of $X_{snA}$

The role of the $X_{snA}$ promoter binding complex in $X_{snA}$ induction suggested by these studies raises many more questions than it answers. There is clearly a mechanism that directs $X_{snA}$ expression in two different germ layers at different times. Earlier data suggested that two different but contiguous regions of the promoter controlled this expression, but the evidence presented here, although not conclusive, points to there being only one complex involved. This suggests that at the appropriate time post-translational modifications and/or a change in subunits may effect the specificity of the transcriptional activation. The presence of the complex in embryos from the earliest moments of embryogenesis in a form that can bind oligonucleotides, indicates that it is not the capacity to bind DNA per se that is regulated. The $X_{snA}$ promoter binding protein complex, if indeed it can be regarded as one unit, seems to be a mechanism that can respond to two different and mutually exclusive signal inputs that cause $X_{snA}$ transcription (that is,
mesoderm induction and neural crest induction). The involvement of PKC dependent phosphorylation in the E and not the M mode of regulation is one indication of how this might occur. The question of how the complex interacts with DNA is also uncertain. Deletion of the M sequence does not affect qualitative expression from the E sequence, while deletion of E ablates all Xsna expression (Mayor et al., 1993) as though this sequence plays a part in both modes of transcription. There is at least one example of a heterodimer protein whose binding sequence is the sum of the sequences recognised by its subunits (Roux et al., 1989).
Appendix A - Summary of the Probes

Summary of the Probes 191
Chapter 3 191
Chapter 4 191
Chapter 5 192
Chapter 6 192
Chapter 7 193
Summary of the Probes

This is a summary of the probes used for the experiments described in the previous chapters. They are all based on the following section of the Xsna promoter sequence which includes the putative mesodermal (M) and ectodermal (E) control regions:

\[
\begin{array}{cccccccc}
-120 & -110 & -100 & -90 & -80 & -70 \\
\hline
5' & ccgtcccaagctgcatattagttagttaggaggggccacagcgctgcctt & 3' \\
\end{array}
\]

Chapter 3

M \quad cacaagctgcataattagttagttggaACTAGTTCc

E \quad tttagtgaggggccacagcagTGC

P \quad CTAGagctgcataattagttagtgaggggcca

Chapter 4

Set M:

Ms1 \quad TAGTagctgcataattagt

Ms2 \quad CTGCgcataattagt

Ms3 \quad agctCGaGattagt

Ms4 \quad agctgcatCGGCgt

Set E:

Es1 \quad ggaggggcacagca

Es2 \quad TTGTgggccacagca

Es3 \quad ggagATTcacagca

Es4 \quad ggaggggCATATgca

DMS Footprinting probe:

A 235 bp fragment, from +74 to -160, amplified by PCR (see Materials and Methods).

Copper-Phenanthroline Footprinting probe:

Amplified by PCR. The primer sites were situated 100 bases upstream on both strands (see Materials and Methods).
Chapter 5

DNA affinity chromatography

A concatemer was made from the following probe:

cTagctgcatattagtA

Chapter 6

Biotinylated Probes

Capital letters = base substitutions. **Bold bases** = bases added by DNA polymerase.

Bio-M    TaGTagctgcatattagttagttagt
          AtCAtcgacgtataatcaatca

Bio-E    tggaggggccacacagcgctcTgGA
         acctccccgggtgctgctgcAcCT

Bio-P    cTAGagctgcatattagttagtggaggggccacGATC
          gATCtcgacgtataatcaatcactacctccccgggtgCTAG

These probes had a primary amine attached through a 12 carbon spacer arm to the 5' end of the upper strand. NHS-ss-Biotin was then able to form a stable amide bond with the amine. For example, Bio-M was formed thus:

[Biotin]-ss-NHS ------> NH₂-[spacer arm]-5'TaGTagctgcatattagttagttagt³
Appendix A - Summary of the Probes

Chapter 7

M cacaagctgcata tagttagtggaACTAGTTbC
E ttagtgga ggccccacagca gcTGC
P CTAGagctgcata tagttagtgga ggggccacGATC

Bio-M TaGTagctgcata taggttagtgga ACTAGTTbC
AtCA tgcacgta atca tcaacctccccggtggtcgc AcCT

Bio-E tggagg ggcacagc agcgTgGA acctccccggtggtcgc AcCT

Bio-P cTA Gagctgcata tagttagtgga ggggccacGATC gATC tgcacgta atca tcaacctccccggtggtcgc AcCT

Bio-MP cTA Gagctgcata tagttagtgga ggggccacGATC gATC tgcacgta atca tcaacctccccggtggtcgc AcCT

Bio-EP cTA Gagctgcata tagttagtgga ggggccacGATC gATC tgcacgta atca tcaacctccccggtggtcgc AcCT

Biotinylated Probes

As for chapter 6 above, these probes had a primary amine attached through a 12 carbon spacer arm to the 5' end of the upper strand, which could then be biotinylated using NHS-ss-Biotin.

Capital letters = base substitutions.

Bold bases = bases added by DNA polymerase (for UV crosslinking experiments, thymidine was substituted by bromodeoxyuridine).
References


