

XENOPUS BOREALIS TADPOLE GLOBIN

GENE EXPRESSION

JANE H. GOULD

Doctor of Philosophy  
University of Edinburgh

1983



## DECLARATION

I declare that this thesis was composed by myself and that the work presented here is my own, unless otherwise stated.

## CONTENTS

	PAGE
DECLARATION	I
CONTENTS	II
ACKNOWLEDGEMENTS	VII
ABBREVIATIONS	VIII
ABSTRACT	X
CHAPTER ONE - INTRODUCTION	1
1.1 Haemoglobin synthesis in amphibians	
1.1.a Haemoglobin synthesis in <u>Rana</u>	3
1.1.b Haemoglobin synthesis in <u>Xenopus</u>	6
1.1.c Mechanism of haemoglobin transition	10
1.1.d Effects of thyroid hormones on haemoglobin synthesis in amphibia	15
1.1.e Effects of anaemia on haemoglobin synthesis	17
1.2 Studies on <u>Xenopus laevis</u> globin gene expression	
1.2.a Globins	18
1.2.b Isolation of <u>X. laevis</u> globin cDNA clones	19
1.2.c <u>X. laevis</u> globin gene organisation	21
1.3 Objectives of this thesis	26
CHAPTER TWO - MATERIALS AND METHODS	28
2.1 Buffers and stock solutions	29
2.2 Microbiological media	34
2.3 General Methods	
2.3.a Gel electrophoresis	36
2.3.b Precipitation of nucleic acids	43
2.3.c Phenol extraction	44

	PAGE
2.3.d TCA and CTAB precipitation	45
2.3.e Sephadex G-75 column fractionation of incorporated and unincorporated nucleotides	45
2.3.f Deionising formamide	46
2.3.g Preparation of carrier calf thymus DNA	47
2.4 Animals	47
2.5 Globin synthesis in isolated red blood cells	48
2.6 Preparation of <u>X. borealis</u> tadpole blood cell 9S mRNA	49
2.7 Translation of <u>X. borealis</u> tadpole 9S mRNA	50
2.8 Preparation of cDNA clones from <u>X. borealis</u> tadpole 9S poly A plus RNA	
2.8.a cDNA synthesis	52
2.8.b Tailing cDNA/RNA hybrids with dCTP	52
2.8.c 7-30% sucrose gradient of tailed cDNA/RNA	53
2.8.d Preparation of pAT 153 for cloning	53
2.8.e Annealing reactions	54
2.8.f Transformation of <u>E. coli</u> HB101 with annealed pAT 153 - cDNA	56
2.8.g Screening HB101 transformants for ampicillin sensitivity	56
2.9 Colony hybridisation	56
2.10 Determination of insert sizes in <u>X. borealis</u> tadpole globin cDNA clones	59
2.11 Preparation of plasmid DNA for detailed analysis	60
2.12 Restriction analysis of clones with inserts greater than 600 nucleotides	61
2.13 Hybridisation-selection of RNAs homologous to cloned sequences	61
2.14 Hybridisation of RNA bound to nitrocellulose	63



	PAGE
2.14a Binding RNA to filters	63
2.14b Preparation of probes for hybridisations	64
2.14c Hybridisations	65
2.15 DNA sequencing	66
2.15a Stock solutions	66
2.15b Subcloning into M13mp8	70
2.15c Preparation of single stranded template DNA	72
2.15d One track screening of templates	73
2.15e Sequencing	74
 CHAPTER THREE - ANALYSIS OF GLOBINS AND 9S mRNA ISOLATED	 76
FROM <u>XENOPUS BOREALIS</u> TADPOLE RBCs	
3.1 Introduction	77
3.2 Tadpole blood contains many immature RBCs	79
3.3 Tadpole globins	80
3.4 Globin synthesis in isolated RBCs	81
3.5 RNA preparation	84
3.6 <u>In vitro</u> translation of <u>X. borealis</u> tadpole blood cell 9S mRNA	86
3.7 Discussion	86
 CHAPTER FOUR - PREPARATION OF <u>XENOPUS BOREALIS</u> TADPOLE	 91
GLOBIN cDNA CLONES	
4.1 Introduction	92
4.2 Preparation and identification of clones containing sequences complementary to <u>X. borealis</u> tadpole blood cell 9S poly A plus RNA	93

	PAGE
4.3 Determination of insert sizes in <u>X. borealis</u> tadpole globin cDNA clones	95
4.4 Restriction analysis of clones with inserts greater than 600 nucleotides	96
4.5 Discussion	97
 CHAPTER FIVE - CORRELATION OF <u>XENOPUS BOREALIS</u> TADPOLE $\alpha$ - AND $\beta$ -GLOBIN cDNA SEQUENCES WITH PROTEIN AND RNA BANDS ON GELS	 102
5.1 Introduction	103
5.2 Hybridisation-selection of homologous RNAs	103
5.3 Hybridisation of RNA bound to nitrocellulose	106
5.4 Discussion	108
 CHAPTER SIX - DETERMINATION OF <u>XENOPUS BOREALIS</u> TADPOLE $\alpha$ - AND $\beta$ -GLOBIN cDNA SEQUENCES	 114
6.1 Introduction	115
6.2 Results	116
6.3 Discussion	120
 CHAPTER SEVEN - DISCUSSION	 126
7.1 <u>Xenopus borealis</u> tadpole globins	127
7.2 Onset of <u>X. borealis</u> adult globin synthesis	129
7.3 Comparison of Xbt9S poly A plus and poly A minus RNA	132
7.4 Preparation and analysis of <u>X. borealis</u> tadpole globin cDNA clones	133
7.5 Future experiments	136

	PAGE
REFERENCES	138
Appendix I Calculation of amount of cDNA synthesised	146
Appendix IIA Calculation of number of residues added to 3' termini in cDNA dC-tailing reaction	148
IIB Calculation of number of residues added to 3' termini in pAT dG-tailing reaction	149
Appendix III Calculation of amounts of tailed cDNA/RNA in sucrose gradient fractions	151
Appendix IV Determination of restriction fragment sizes	152
Appendix V Linear regression programme for Texas Instruments SR56 calculator	154
Appendix VI Estimation of RNA Cot and reaction extent for hybrid-selection experiments	156

## ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Peter Ford, my adviser Dr. Mary Bownes and all those who have also helped me during the last three years.

I was supported by a Medical Research Council studentship from 1979 to 1982.

## ABBREVIATIONS

AMPS	ammonium persulphate
Ap	ampicillin
BSA	bovine serum albumin
butyl PBD	2-(4'-tert-butylphenyl)-5-(4"-biphenyl)-1,3,4-oxadiazole
°C	degrees centigrade
cDNA	complementary deoxyribonucleic acid
Ci	Curie(s)
cm	centimeter(s)
cpm	counts per minute
CTAB	cetyltrimethylammonium bromide
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
DTT	dithiothreitol
EDTA	diaminoethanetetra-acetic acid
EGTA	ethyleneglycol bis-(aminoethyl)-tetra-acetic acid
h	hour(s)
<sup>3</sup> H	tritium (β emitting isotope of hydrogen)
Hepes	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
[ <sup>3</sup> H]poly U	tritium labelled homopolymer of uracil
Kb	kilobase
M	molar
mA	milliampere(s)
mg	milligram(s)
ml	millilitre(s)
mm	millimeter(s)
mM	millimolar
mmol	millimole
mRNA	messenger ribonucleic acid
MS222	ethyl-M-aminobenzoate
Na <sub>2</sub> EDTA	diaminoethanetetra-acetic acid-disodium salt
Na-MOPS	morpholinopropanesulfonic acid
NaOAc	sodium acetate
ng	nanogram(s)
<sup>32</sup> P	radioactive (β emitting) isotope of phosphorous
PEG 6000	polyethylene glycol 6000
%	per cent
pH	log <sub>10</sub> (hydrogen ion concentration)
PIPES	Piperazine-N,N'-bis[2-ethane sulphonic Acid]
poly A	adenosine homopolymer
PPO	2,5-diphenyloxazole
RBC	red blood cell
RNA	ribonucleic acid
RNase	ribonuclease
<sup>35</sup> S	radioactive (β emitting) isotope of sulphur
SDS	sodium dodecyl sulphate (sodium lauryl sulphate)
Sm	streptomycin

ss	single strand
Tc	tetracycline
TCA	trichloroacetic acid
TEMED	NNN'N'-tetra-methyl-1,2-diamino-ethane
Tris	tris(hydroxymethyl)aminomethane
Triton-X 100	iso-octylphenoxypolyethoxyethanol
$\mu$ Ci	microcurie(s)
$\mu$ g	microgram(s)
$\mu$ l	microlitre(s)
$\mu$ M	micromolar
v/v	volume per volume
w/v	weight per volume
Xba 9S	<u>Xenopus borealis</u> adult anaemic blood cell 9S
Xbt 9S	<u>Xenopus borealis</u> tadpole blood cell 9S

## ABSTRACT

Globin genes are a family of genes which are coordinately regulated during development; different sets of globin genes are sequentially expressed during development. For this reason, and since globin mRNA constitutes a large proportion of the mRNA in RBCs, analysis of globin gene regulation is a convenient system in which to study control of gene expression during development.

In this thesis globin gene expression in Xenopus borealis tadpoles is investigated by the preparation and analysis of globin cDNA clones. Larval globin genes are expressed in Xenopus tadpoles and a switch to adult globin synthesis occurs during metamorphosis. Tadpole globins are shown to consist of three major and five minor species, with no evidence of individual polymorphism, which are completely separable from adult globins on triton-acid-urea polyacrylamide gels. The transition from tadpole to adult globin synthesis begins about stage 57, as demonstrated by labelling studies of RBCs in culture.

9S mRNA sequences which are known to be normally expressed in tadpole RBCs at stages 57-59 of development are isolated since it is unnecessary to induce anaemia to enable RNA extraction from tadpole RBCs and the RNA is therefore prepared from physiologically normal animals. This 9S mRNA codes for all globin species normally present in tadpole RBCs, shown by in vitro translation in rabbit reticulocyte lysate.

Restriction enzyme analysis sorted cDNA clones prepared from stage 57-59 9S poly A plus RNA into four classes, identified as two  $\alpha$ - and two  $\beta$ -like globin coding classes by comparison with X. laevis tadpole globin cDNA clones (Widmer et al, 1981). DNA sequencing confirms this for the two largest groups (one  $\alpha$ , one  $\beta$ ).

Most globin mRNA is polyadenylated as shown by dot and gel blot hybridisations of tadpole blood cell 9S poly A plus and poly A minus RNA with cloned  $\alpha$ - and  $\beta$ -globin probes.

9S poly A plus RNA is shown to be heterogeneous on formaldehyde-agarose gels, the two major bands containing  $\alpha$ - ( $848 \pm 66$  nucleotides) and  $\beta$ - ( $680 \pm 56$  nucleotides) globin coding mRNA, demonstrated by hybridisation to cloned  $\alpha$ - and  $\beta$ -globin probes. Tadpole 9S mRNA contains a low level of adult globin message which comigrates with tadpole  $\beta$ -globin mRNA on formaldehyde-agarose gels, as shown by hybridisation of a cloned adult globin probe to an RNA gel blot.

Sequences coding for all three major bands on triton-acid-urea polyacrylamide gels have been cloned; this was demonstrated by hybridisation-selection of RNA homologous to cloned DNA sequences bound to nitrocellulose, and translation of this RNA in rabbit reticulocyte lysate.



CHAPTER ONE

INTRODUCTION

Haemoglobins are found throughout the vertebrate phylum and appear to constitute the only respiratory protein in blood (Maclean and Jurd, 1972). In most vertebrates haemoglobins are tetrameric proteins of molecular weight 61000 to 72000, consisting of four globin chains each of which is associated with an iron-containing porphyrin haem group. There are some exceptions to the tetrameric structure: Lampetra have monomeric haemoglobin (molecular weight 17000); Eptatretus burgeri (hagfish) have four monomeric haemoglobins, some of which (but not all) can associate to form tetramers; Myxine glutinosa has haemoglobins of molecular weights 17000 and 34000 indicating the presence of monomers and dimers (for references see Maclean and Jurd, 1972).

Tetrameric haemoglobins are present in all remaining vertebrate groups and appear to be present in virtually all species (Maclean and Jurd, 1972). A few animals do not apparently contain haemoglobin e.g. larvae of the eel Leptocephalus contain no haemoglobin until they develop into elvers and antarctic ice fish have no erythrocytes or haemoglobin in their blood (Ruud, 1954). Isolated specimens of Myxine (Rybak, 1960) and X. laevis (Ewer 1959, de Graaf 1957) lacking red blood cells have been reported, but these are apparently mutants.

In most vertebrates different haemoglobins are synthesised during development (Maclean and Jurd, 1972) and this is thought to reflect changes in globin synthesis. In addition to this multiple haemoglobins are generally present at all stages of development. It has been proposed that this is a functional advantage to the animal (Kellet, 1971). In amphibia adaptive properties have been attributed to the ontogenic forms of haemoglobin (Riggs, 1951) since the mode of

respiration and oxygen availability changes dramatically during metamorphosis. However these adaptive advantages are not apparent for other vertebrate groups. It has been suggested for placental mammals that foetal and adult haemoglobins are modified for use on either side of the placental barrier, however evolution of the different developmental forms of haemoglobin preceded evolution of placental gas exchanges (Maclean and Jurd, 1972).

Maclean and Jurd (1972) suggest that multiple haemoglobins confer neither selective advantage nor disadvantage on organisms, but instead they have arisen through random evolutionary change of multiple globin loci in the absence of selective pressure. It has also been proposed (Gluckson-Waelsch, 1960) that multiple haemoglobins could act as a buffer against mutational change, since a deleterious effect in one haemoglobin may be compensated by the production of another.

## 1.1

### HAEMOGLOBIN SYNTHESIS IN AMPHIBIANS

#### 1.1.a Haemoglobin synthesis in Rana

McCutcheon (1936) first reported the occurrence of larval and adult forms of haemoglobin, differing in their oxygen affinities, in the bullfrog Rana catesbeiana. Tadpole haemoglobin has subsequently been shown to differ from adult in its electrophoretic or chromatographic mobility (Herner and Frieden, 1961; Baglioni and Sparks, 1963; Hamada et al 1964; Moss and Ingram, 1968a), polypeptide subunits (Baglioni and Sparks, 1963; Moss and Ingram,

1968a; Aggarwal and Riggs, 1969), and antigenic determinants (Maniatis et al, 1969).

Rana catesbeiana tadpole haemoglobins consist of four major components, numbered I-IV, each of which is made up of two  $\alpha$ -like and two  $\beta$ -like globin subunits (Watt and Riggs, 1975). There are three distinct  $\alpha$ -globin chains and three major and three minor  $\beta$ -globin chains. Haemoglobin components II and III have an  $\alpha$ -globin chain in common and components III and IV have the same major  $\beta$ -globin chain (Riggs personal communication to Broyles, 1981). The four haemoglobins are functionally distinct, components I and II having higher oxygen affinities than components III and IV. Approximately 90% of the haemoglobin of very young tadpoles (Taylor and Kollros (TK) stages III and IV, Taylor and Kollross, 1946, see figure 1, Broyles 1981 for comparison of amphibian staging systems) is type I and II whereas older, larger tadpoles (about TK stage X, hind limb paddle stage) have about 65% of their haemoglobin as types III and IV (Watt and Riggs, 1975).

Haemoglobins of adult Rana catesbeiana consist of four types, 2 major (designated B and C) and two minor (A and D) (Aggarwal and Riggs, 1969). Component C has a higher oxygen affinity than B and Saucier (1976) observed that it is the first adult haemoglobin to appear during metamorphosis and comprises essentially all the haemoglobin of recently metamorphosed frogs. McCutcheon (1936) showed that haemoglobin oxygen affinity of large sexually mature frogs is lower than that of young frogs; thus Broyles (1981) states that there is a transition in haemoglobin types during the development of Rana catesbeiana, which is correlated with changes in

globin gene expression, from those of high oxygen affinity to haemoglobins with a lower oxygen affinity. Frieden (1968) observed that larval and adult haemoglobins are well adapted to their respective environments: those of tadpoles having a high oxygen affinity suitable for an aquatic lifestyle where oxygen tensions may be low, and those of adults having a lower oxygen affinity facilitating oxygen release into the animals' tissues to support increased muscular activity.

The transition from tadpole to adult haemoglobin production in Rana has been extensively studied (for references see Broyles 1981). In Rana catesbeiana it usually occurs rapidly; electrophoretic analysis of haemoglobins on polyacrylamide gels shows that tadpole haemoglobins start to disappear and adult haemoglobins to appear during tail regression (TK stage XXII) and the transition is almost complete by TK stage XXIV (roughly equivalent to Xenopus stage 65, Nieuwkoop and Faber, 1956)(Just and Atkinson, 1972). Many other studies support this data, (for references see Broyles 1981), and it is generally agreed that adult haemoglobin is first detectable in circulating RBCs shortly after emergence of the forelimbs (i.e. soon after TK stage XX). By TK stages XXII-XXIII about half the haemoglobin is adult type and the transition is complete by stages XXIV-XXV (tail resorbed).

Just and Atkinson (1972) investigated the synthesis of tadpole and adult haemoglobins by injecting animals with labelled amino acids and analysing RBC products on polyacrylamide gels. Selective synthesis of adult haemoglobins begins about TK stage XXII: in six animals tested adult haemoglobin represented 20% of total haemoglobin

at this stage but it contained, on average, 40% of the radioactivity. By stage XXIII there was complete suppression of larval haemoglobin synthesis and 70% of the incorporated radioactivity was associated with adult haemoglobin. High production of adult haemoglobin continued in one week old frogs.

In contrast to the results described above, Benbassat (1970) reported that the transition from tadpole to adult haemoglobin production in Rana catesbeiana begins at metamorphic climax (TK stage XX), but is not complete until after tail resorption. Analysis on polyacrylamide gels of haemoglobins from RBCs of newly metamorphosed frogs showed that they were predominantly tadpole haemoglobins. Transition in these animals was not complete until 4-10 weeks after metamorphosis. The same delay was also observed in Rana pipiens (Benbassat, 1970). Summer tadpoles were used in this work, but Benbassat (1970) did not state whether the experiments were carried out in early or late summer. Broyles and Dorn (unpublished observations, reported in Broyles 1981) found that the haemoglobin transition takes place at later metamorphic stages as the season progresses from late spring to early summer.

#### 1.1.b Haemoglobin synthesis in Xenopus

Haemoglobins of X. laevis are also heterogeneous: Maclean and Jurd (1971a) reported that stage 45 (Nieuwkoop and Faber, 1956) X. laevis tadpole RBCs contain two haemoglobin species, (HbF<sub>1</sub> and HbF<sub>2</sub>) present in approximately equal amounts, which are replaced in adults by two different haemoglobins (HbA<sub>1</sub> and HbA<sub>2</sub>), one of which predominates. Just et al (1977) analysed haemoglobins from

stage 57 tadpoles and found that they consisted of 5 to 9 different species, the number varying between individuals. This variation was absent in haemoglobins of isogenic tadpoles derived from female hybrids of X. laevis/X. gilli, suggesting it was due to genetic polymorphism. Further analysis on gels of varying polyacrylamide concentrations demonstrated that the multiple haemoglobins were proteins of the same molecular weight but different charge (Just et al, 1977). It is unclear why Maclean and Jurd (1971a) detected no polymorphism, although the different developmental stages used could account for the discrepancy.

Maclean and Jurd (1971a) observed a change in the proportions of HbF<sub>1</sub> and HbF<sub>2</sub> during development: by stage 50 HbF<sub>1</sub> comprises about 75% of total haemoglobin. This is analogous to the change in Rana catesbeiana tadpole haemoglobins discussed in the previous section (Watt and Riggs, 1975). In both cases there is a shift towards the faster electrophoretic forms of tadpole haemoglobins during tadpole development.

Adult X. laevis RBCs contain about four types of haemoglobin. Maclean and Jurd (1971a) found one major (HbA<sub>1</sub>) and one minor (HbA<sub>2</sub>) component, accounting for 92% and 5% of the total haemoglobin respectively, with the rest being divided between two very minor components which might be residual tadpole haemoglobins. Just et al (1977) could detect 3-4 types of haemoglobin, on polyacrylamide gels, in blood of normal and isogenic Xenopus.

The transition from tadpole to adult haemoglobin synthesis has been extensively investigated in X. laevis. Early investigations

suggested that haemoglobin transition takes place after metamorphosis (Jurd and Maclean, 1969, 1970, 1974; Maclean and Jurd, 1971a). However by injecting animals with radioactive amino acids and analysing incorporation into RBC products by gel electrophoresis Just et al (1977) showed adult haemoglobin synthesis to be first detectable at about stage 57: up to stage 57 approximately 82% of haemoglobin synthesis was found to be of larval type, this decreases to 32% by stages 65-66 and to 10% in three week old frogs. Haemoglobin synthesis in RBCs of frogs 7-8 weeks after metamorphosis is 97% adult type. Comparison of the ratios of %cpm/%OD410nm for adult and tadpole haemoglobins before and after haemoglobin transition shows it remains one and Just et al (1977) conclude that small amounts of adult haemoglobin are synthesised in larval stages and after metamorphosis there is still a low level of tadpole haemoglobin synthesis.

In the experiments of Just et al (1977) described above it was not possible to distinguish unequivocally between tadpole and adult haemoglobins. Based on the observations of Maclean and Jurd (1971a) that tadpole and adult haemoglobins are immunologically distinct, Just et al (1980) therefore investigated the accumulation of adult X. laevis haemoglobin by immunological techniques. Antisera were prepared against tadpole and adult haemoglobins and their specificity demonstrated - neither antiserum gave any cross reaction with the heterologous Xenopus haemoglobin or with human haemoglobin. Haemolysates from animals at various developmental stages were tested for the presence of tadpole and adult haemoglobins by immunoelectrophoresis. No adult haemoglobin could be detected at stage 57; stage 62 haemolysates all reacted with larval antiserum and half also



gave a weak reaction with adult antiserum; stage 66 haemoglobin reacted with both larval and adult antibodies, although the stronger reaction was with larval antiserum. Later stages analysed showed a gradual decrease in amount of tadpole haemoglobin and simultaneous increase in adult haemoglobin as the frogs grew older. By 8 weeks after metamorphosis only one of the animals investigated had RBCs containing detectable quantities of tadpole haemoglobin.

Just et al (1980) quantitated, by the Mancini radial immunodiffusion test (Fey et al, 1976), the amounts of larval and adult haemoglobins present during development. At stage 60 3.4% of the haemoglobin was adult type, this increased to 8.0% by stage 64 and between 13 to 65% (mean 35%) at stage 66. Two weeks after metamorphosis adult haemoglobin comprised 90% of the total and by 6 weeks after metamorphosis it had reached its final level of 98.8%.

Just et al (1980) suggest that the differences in timing of the haemoglobin transition observed by Maclean and Jurd (Jurd and Maclean, 1969, 1970; Maclean and Jurd, 1971a) and themselves might be due to variation in feeding and maintenance conditions which they have found to affect the transition.

The immunological studies of Just et al (1980) confirm Maclean and Jurd's earlier reports (Jurd and Maclean, 1969, 1970; Maclean and Jurd, 1971a) and their own electrophoretic evidence (Just et al, 1977) that Xenopus larval haemoglobins are more heterogeneous than adult haemoglobins. In Ouchterlony tests (Ouchterlony, 1958) tadpole RBCs reacted with homologous antiserum to give two precipitation lines, whereas adult antiserum reacted against adult RBCs gave only

one line of precipitation. This variation is most likely due to genetic polymorphism since isogenic animals from a clone of X. laevis/X. gilli show no variation in haemoglobin electrophoretic pattern on polyacrylamide gels (Just et al, 1977) or in immunoelectrophoretic pattern (Just et al, 1980). Thus X. laevis appears to be similar to Rana catesbeiana, the larval haemoglobins of which are more complex and variable than those of the adult (Maniatis and Ingram, 1971b).

#### 1.1.c Mechanism of haemoglobin transition

Larval haemoglobin disappears rapidly from circulating RBCs of metamorphosing anurans, in Xenopus about 80% is lost in four weeks (Just et al, 1980) and in Rana catesbeiana the transition is complete in about ten days (Dorn and Broyles, 1982). Coupled with the report of Forman and Just (1976), that the lifespan of amphibian RBCs is about 100 days, this suggests that larval RBCs are selectively eliminated during the transition. Kistler and Weber (1975) have observed intense phagocytosis of red cells in the liver of X. laevis tadpoles undergoing induced metamorphosis, however the mechanism by which circulating RBCs could be eliminated is not known. There is some doubt as regards the accuracy of Forman and Just's estimate of the lifespan of RBCs (Broyles, 1981): RBCs were labelled by injecting [<sup>3</sup>H]thymidine into tadpoles recovering from phenylhydrazine induced anaemia and the decrease in number of nuclei containing label measured by autoradiography. Three criticisms can be made of this work (Broyles, 1981): (1) No evidence was given that the label could not be conserved and recycled in the tadpoles; (2) The turnover of RBCs in anaemic animals is likely to be abnormal;

(3) No early time points (before 40 days after injection of label) were given in the data.

If loss of tadpole haemoglobin is achieved by eliminating RBCs in which it is contained it seems probable that adult and tadpole haemoglobins are synthesised in different sets of RBCs. There are conflicting reports, which will be described below, regarding this.

Differences in morphology of larval and adult RBCs have been observed in Rana pipiens (Hollyfield, 1966a; Benbassat, 1970), Rana catesbeiana (McCutcheon, 1936) and X. laevis (Jurd and Maclean, 1970). The mature RBC of amphibians retains its nucleus in both tadpoles and adults. Benbassat (1970) observed that Rana pipiens tadpole erythroid cells consist of two main types; most are large and elliptical, some are small and round. Nuclei of both types are round and a cytoplasmic reticular substance could be detected by staining cells. Adult RBCs are more uniform in shape and size, their nuclei are elliptical and no reticular substance could be detected in the cytoplasm. These results suggest that adult RBCs are more mature cells and less active in the synthesis of nucleic acids and proteins as compared to larval cells. Benbassat (1970) confirmed this by showing that RBCs derived from premetamorphic tadpoles of Rana catesbeiana and Rana pipiens have an increased capacity to incorporate labelled amino acids, thymidine or uridine as compared to young frogs or adults.

Thus morphological and biochemical evidence shows that there are clear differences between tadpole and adult blood cells. Many attempts therefore have been made to find out whether both tadpole

and adult haemoglobins are present in a single cell or whether the transition from larval to adult haemoglobin synthesis is achieved by elimination of one cell line and introduction of another (reviewed in Broyles, 1981).

Electrophoresis of haemoglobin from single cells (Rosenberg, 1970) and double immunofluorescence labelling of RBCs with anti-tadpole and anti-adult haemoglobin-specific antibodies (Maniatis and Ingram, 1971b, c) indicated that tadpole and adult haemoglobins are found in separate cells of Rana catesbeiana at metamorphic climax. However, using similar immunofluorescence techniques Benbassat (1974) found that up to 16% of circulating RBCs of metamorphosing Rana catesbeiana tadpoles contained both tadpole and adult haemoglobins. Jurd and Maclean (1970) found up to 25% of X. laevis RBCs contained both types of haemoglobin at the peak of metamorphic transition, and this proportion decreased as metamorphosis proceeded.

There are several possible explanations for these conflicting results (Broyles, 1981): (1) the dual labelling method used by Maniatis and Ingram (1971b, c) might fail to detect adult haemoglobin in RBCs which contain a high proportion of tadpole haemoglobin, and vice versa, since one fluorescent dye might mask the other. Rhodamine and fluorescein were the fluorescent labels used, fluorescein has a much stronger fluorescence than rhodamine which might prevent the rhodamine fluorescence being detected in the presence of fluorescein; (2) the results of Rosenberg (1970) agree with those of Maniatis and Ingram, (1971b, c), however analysis of haemoglobin from single RBCs is not as sensitive as immunofluorescent labelling and small amounts of haemoglobin may not have been

detected. Also Rosenberg's separation of adult and tadpole haemoglobin was based on the dimerisation of adult haemoglobin which Moss and Ingram (1968b) have reported does not occur at low haemoglobin concentrations.

More recently Dorn and Broyles (1982) have analysed haemoglobins contained within erythrocytes of metamorphosing Rana catesbeiana tadpoles. Artificial mixtures of adult and tadpole erythrocytes can be completely separated on Percoll density gradients. During metamorphosis tadpole erythrocytes increase in density and by TK stage XXIV have disappeared from the circulation. Beginning at TK stage XXI a population of erythroid cells of immature morphology appears, increasing in number as metamorphosis proceeds. When tadpole erythrocytes are no longer present (TK stage XXIV) these cells mature and increase in density. Cells of these different morphologies are completely separable at all stages of metamorphosis. Animals at various developmental stages were injected with [<sup>3</sup>H]-leucine and [<sup>35</sup>S]-methionine and incorporation into haemoglobin determined for each of the erythroid cell populations separated by Percoll density gradient centrifugation. At no stage could adult and tadpole haemoglobin be detected in the same cell populations. Although it is possible that a small proportion of cells could contain both types of haemoglobin, or that all cells could contain a very small amount of the alternative haemoglobin, these results demonstrate that tadpole and adult haemoglobins are essentially expressed in separate cell populations.

Related to the appearance of new erythroid cell types during metamorphosis is the possibility that a change in erythropoietic site

is responsible for haemoglobin transition. Broyles and Deutsch (1975) suggested that erythropoietic microenvironments might be important in determining which type of haemoglobin is synthesised. There is evidence to suggest that tadpole and adult erythropoiesis occur in different organs (reviewed by Foxon, 1964). In Rana catesbeiana tadpoles conflicting reports have indicated that kidneys (Hollyfield, 1966b) and liver (Maniatis and Ingram, 1971a,c) are the main erythropoietic sites. In adults the spleen, liver and bone marrow are reported to be active (Jordon and Spiedel, 1923). However Maniatis and Ingram (1971a,c) observed no change in erythropoietic sites of tadpoles and young frogs, the liver being active in both.

Broyles and Frieden (1973) appear to have resolved the discrepancies mentioned above: short term organ cultures of Rana catesbeiana tadpole liver and kidney with radioactive precursors of haem and globin showed that both organs incorporated label into haemoglobins. They also demonstrated that the kidney and liver synthesise different proportions of each type of tadpole haemoglobin, and Broyles and Deutsch (1975) showed that each organ contains different larval haemoglobins with no detectable overlap after one week of culture. It was shown (Broyles and Deutsch, 1975) that the different tadpole haemoglobins disappear from the circulation at different rates during phenylhydrazine induced anaemia. These results suggest that the different tadpole haemoglobins of Rana catesbeiana are contained within separate cell lines produced either by the kidney or by the liver. As mentioned earlier erythroid cells of different morphologies can be seen in blood smears from Rana catesbeiana tadpoles (Benbassat, 1970).

The work of Just et al (1977) suggests that in X. laevis a switch in cell types is also responsible for haemoglobin transition: analysis of haemoglobin electrophoretic patterns from normal and isogenic animals shows that the presence or absence respectively of haemoglobin polymorphism persists throughout the transition. This suggests that the larval components are controlled as an entity, tadpole erythrocytes gradually disappearing as adult erythrocytes are produced.

The expression of different haemoglobins within single cells or within separate cell populations has important implications regarding determination of haemopoietic cells and control of haemoglobin synthesis. If the switch from tadpole to adult haemoglobin synthesis is achieved by elimination of one cell line and introduction of another the decision regarding which haemoglobin type is to be synthesised could be made at an early stage in haemopoietic cell differentiation. If, on the other hand, both tadpole and adult haemoglobins are present in the same cells it is possible that the choice is made later, after the cell is determined to become an erythroid cell. No experiments directly addressing this problem have yet been done.

#### 1.1.d Effects of thyroid hormones on haemoglobin synthesis in amphibia

Since the transition in haemoglobin synthesis occurs during metamorphosis it was proposed that thyroid hormones may be involved in the switch. The early experiments investigating this are reviewed in Broyles (1981). Maclean and Turner (1976) reared X. laevis

tadpoles from stage 52 in a solution of propylthiouracil, which stopped morphological development at stage 54 although the animals were still able to grow in size. Analysis of haemoglobin from such developmentally retarded animals showed that tadpole haemoglobin was no longer detectable and apparently normal adult haemoglobin was present. In a similar experiment Just et al (1977) raised tadpoles from hatching in 0.05% thiourea, arresting morphological development at stage 53 whilst still allowing growth. After 21 weeks haemoglobin transition had started and it was almost complete by 24 weeks.

Maclean and Turner (1976) concluded from their results described above and the observation that neotenus adult Ambystoma mexicanum has adult not larval haemoglobins (Maclean and Jurd, 1971b), that haemoglobin transition is not dependent on thyroid hormones. However, this may not necessarily be the case: Kollros (1961) demonstrated that various metamorphic events have differing thresholds of sensitivity to thyroid hormones; Tompkins and Townsend (1977) induced precocious haemoglobin transition in Ambystoma mexicanum, by administration of thyroid hormones, without influencing morphological transformation; Xenopus laevis tadpoles arrested at stage 57 with thiourea and then induced to metamorphose by thyroxine treatment undergo precocious and complete haemoglobin transition (Just et al, 1977). It has therefore been proposed that haemoglobin transition in Xenopus is responsive to thyroid hormones, but at a level which is insufficient to induce morphological changes (Just et al, 1977).



### 1.1.e Effects of anaemia on haemoglobin synthesis

Thomas and Maclean (1975) showed that injection of phenylhydrazine into adult X. laevis destroys mature erythrocytes, resulting in a wave of erythropoiesis and release of many immature erythroid cells into the circulation. Several groups have analysed haemoglobin present in erythroid cells of animals recovering from phenylhydrazine administration. Maclean and Jurd (1971a) observed a modification in the relative proportions of the different adult haemoglobins and also partial reversion to tadpole haemoglobin synthesis. Battaglia and Melli (1977) analysed globins from normal and anaemic X. laevis adults and observed a change in their proportions which presumably reflects the haemoglobin changes noticed by Maclean and Jurd (1971a). However, they did not detect any tadpole globins in blood of anaemic animals. Hentschel et al (1979) examined globins of frogs during the early stages of recovery from phenylhydrazine-induced anaemia, since Maclean and Jurd (1971a) found that tadpole-like haemoglobins were most prominent at this time. They were unable to detect any major tadpole globins, but they observed a novel globin species not normally present in circulating adult RBCs. Their gel system did not allow clear distinction between minor adult and tadpole globin species and so it remains possible that this novel globin is a minor tadpole component.

1.2.a Globins

Immunological analysis of X. laevis tadpole and adult haemoglobins showed that antisera prepared against either one did not cross react with the other (Maclean and Jurd, 1971a; Just et al, 1980). Antisera against human haemoglobins have been shown to cross react only if the haemoglobins share a common polypeptide (Reichlin et al 1965) and Just et al (1980) therefore concluded that X. laevis adult and tadpole haemoglobins have no globin chains in common. Solution-hybridisation experiments (Perlman et al, 1977) confirm this and show that adult globin sequences do not hybridise to tadpole globin sequences. Stratton and Frieden (1967) demonstrated that Rana catesbeiana adult and tadpole haemoglobins also share no globin chains.

Electrophoresis of haemoglobin in denaturing conditions through acid-urea polyacrylamide gels (Panyim and Chalkey, 1969) separates the constituent globin chains. Using this gel system Hentschel et al (1979) detected 2 major and 2 minor globin species in pooled haemolysates from tadpoles and 2 major and 4 minor globin species in blood from individual adults. No polymorphism was detected in adult globins and since pooled samples were analysed for tadpoles it was not possible to ascertain whether the observed haemoglobin polymorphism (Just et al, 1977) is due to differences in globin chains.

The work of Hentschel et al (1979) demonstrated that the major adult and tadpole globin components are different, but there was some

uncertainty as to whether the minor components overlapped. Addition of Triton-X 100 to the gel improves resolution and Hosbach et al (1982) were able to show that X. laevis adult and tadpole globins are completely distinct polypeptides.

### 1.2.b Isolation of X. laevis globin cDNA clones

cDNA clones coding for X. laevis adult and tadpole  $\alpha$ - and  $\beta$ -globins have been isolated (Humphries et al, 1978; Hentschel et al 1979; Kay et al, 1980; Richardson et al, 1980; Widmer et al, 1981). Clones were constructed from 9S mRNA isolated from blood of anaemic adults and tadpoles. Hentschel et al (1979) identified by hybrid-arrested translation (Paterson et al, 1977) clones containing incomplete copies of one major and one minor adult globin sequence. Subsequently the same group isolated almost full length copies of a major  $\alpha$ - and a major  $\beta$ -globin sequence (Kay et al, 1980). Partial sequence analysis identified the clones as  $\alpha$ - and  $\beta$ -globin coding sequences and hybrid-arrested translation showed the faster migrating major globin on polyacrylamide gels to be  $\alpha$ - and the slower one  $\beta$ -globin.

Widmer et al (1981) isolated four types of tadpole and four types of adult globin cDNA clones and demonstrated by solution hybridisation that they fall into four groups each of which can be subdivided into two related classes. By comparison with restriction maps of X. laevis adult  $\alpha$ - and  $\beta$ -globin cDNA sequences, reported by Kay et al (1980) and Richardson et al (1980), Widmer et al (1981) were able to identify their adult cDNA sequences: two correspond to the major  $\alpha_1$ - and  $\beta_1$ -globin sequences isolated by Kay et al

(1980) and the remaining sequence in each group is a minor globin sequence. Partial DNA sequence analysis and comparison of amino acids conserved in most other globins (Dayhoff, 1976) enabled identification of tadpole  $\alpha$ - and  $\beta$ -globin sequences.

These experiments indicate that there are at least eight globin genes in X. laevis, four tadpole and four adult.

Widmer et al (1981) analysed the sequence relationship between the cDNA sequences in each group by comparison of the melting temperatures of homo- and heteroduplexes. They concluded that tadpole  $\alpha$ - and  $\beta$ - sequences have diverged by about 14% and 13% respectively, whereas the adult sequences have only diverged by about 6% and 8%. This has implications regarding the evolution of globin genes: Bisbee et al (1977) proposed that the Xenopus genome was duplicated approximately 30 million years ago giving rise to duplicated sets of genes that have subsequently diverged. If this is the case, and further evidence is provided from examination of the albumin and vitellogenin genes (Westley et al, 1981; Wahli et al, 1979), one would expect all duplicated genes to have diverged to the same extent. This is clearly not so for tadpole and adult globin genes and Hosbach et al (1982) offer two explanations: (1) the tadpole globin genes were duplicated first by a gene duplication and again when the whole genome duplicated. This implies that the tadpole genes have been separated longer than the adult genes, which were only duplicated by the genome duplication, and therefore the tadpole genes would be expected to be more widely diverged. This model assumes that the rates of divergence of tadpole and adult genes are similar and that some of the tadpole genes produced in this way have been inactivated, or are expressed at a very low level, since

only two  $\alpha$ - and two  $\beta$ -globin cDNA sequences have been detected (Widmer et al, 1981); (2) tadpole and adult globin genes were duplicated at the same time, but different selective pressures have caused the tadpole genes to diverge more rapidly than the adult genes. Alternatively separate gene conversion events could have occurred in tadpole and adult genes. If gene conversion in the adult genes occurred more recently than in the tadpole genes one would expect the tadpole genes to be less related than the adult genes since the tadpole genes would have been diverging for a greater length of time. This could be investigated by sequencing the globin genes and surrounding genomic DNA sequences.

#### 1.2.c X. laevis globin gene organisation

X. laevis globin gene organisation has been investigated by probing total genomic DNA or X. laevis genomic DNA clone banks with  $\alpha$ - and  $\beta$ -globin cDNA sequences (Kay et al, 1980; Jeffreys et al, 1980; Patient et al, 1980; Hosbach et al, 1982).

EcoRI digested X. laevis blood cell DNA was separated by agarose gel electrophoresis and hybridised to either a major  $\alpha$ - or a major  $\beta$ -globin cDNA probe (Kay et al, 1980; Jeffreys et al, 1980). Two fragments of identical size could be detected with both probes: in each case one fragment formed stable hybrids which only melted out when post-hybridisation washes were very stringent, these were therefore assumed to contain the major  $\alpha$ - or  $\beta$ -globin genes ( $\alpha_1$ ,  $\beta_1$  genes); the second fragment in both cases formed less well matched hybrids and these globin genes were designated  $\alpha_2$  and  $\beta_2$ . At lower stringencies both probes hybridised to additional bands

which were thought to be other globin sequences more diverged from the major  $\alpha$ - and  $\beta$ -globin probes.

Since both  $\alpha$ - and  $\beta$ -globin cDNA hybridised to bands of identical size it was thought that the  $\alpha$ - and  $\beta$ -globin genes might be linked. This was supported by restriction analysis: a Bam HI site to the 5' side of the  $\alpha_1$  globin gene is polymorphic (Jeffreys et al, 1980), some frogs were homozygous for the presence of this site and others were heterozygous. In all cases the same number of fragments were detected with each probe, indicating that the  $\alpha_1$ - and  $\beta_1$ -globin DNA fragments were the same and therefore that the  $\alpha_1$ - and  $\beta_1$ -globin genes are linked.

Linkage of X. laevis  $\alpha$ - and  $\beta$ -globin genes was confirmed by isolation of individual recombinant clones containing both  $\alpha$ - and  $\beta$ -globin sequences from a X. laevis genomic DNA library (Patient et al, 1980). DNA sequencing and comparison with  $\alpha$ - and  $\beta$ -globin cDNA sequences (Kay et al, 1980) demonstrated that these clones contain major  $\alpha$ - and  $\beta$ -globin genes. The  $\alpha$ - and  $\beta$ -globin genes are separated by about 8Kb (Jeffreys et al, 1980; Patient et al, 1980) a similar distance to that separating globin genes within mammalian globin gene clusters.

Introns in the globin genes were mapped by electron microscopy of R loops and comparison of restriction fragment sizes from cDNA and genomic clones (Patient et al, 1980). Both  $\alpha$ - and  $\beta$ -globin genes have two introns, 150 and 310, and 185 and 930 base pairs long respectively. Alignment of DNA sequences from cDNA and genomic clones allows intron/exon junctions to be delineated (although there

may be some error due to sequence repetition at the boundaries, Patient et al, 1980). Comparison with globin gene sequences from mouse and rabbit showed that the introns were located at homologous positions within the genes.

Linkage of X. laevis  $\alpha$ - and  $\beta$ -globin genes was unexpected as in all other species studied  $\alpha$ - and  $\beta$ -like globin genes are clustered on separate chromosomes. Human globin gene organisation has been extensively studied (reviewed by Weatherall and Clegg, 1979; Maniatis et al, 1980; Efstratiadis et al, 1980): the  $\alpha$ -like globin genes are located on chromosome 16 and the  $\beta$ -like globin genes on chromosome 11. The genes are arranged in the order of their expression during development i.e.

5' -  $\zeta 2$  -  $\zeta 1$  - ( $\psi\alpha 1$ ) -  $\alpha 2$  -  $\alpha 1$  - 3' and  
5' - ( $\psi\beta 2$ ) -  $\epsilon$  -  $G_{\gamma}$  -  $A_{\gamma}$  - ( $\psi\beta 1$ ) -  $\delta$  -  $\beta$  - 3',

and it was therefore proposed that the order is important for regulated expression during development. This kind of arrangement has also been shown in mouse  $\alpha$ - and  $\beta$ -globin genes and rabbit  $\beta$ -globin genes (Jahn et al, 1980; Leder et al, 1981; Lacy et al, 1979). However, analysis of chicken  $\beta$ -globin genes revealed a different arrangement (Roninson and Ingram, 1982): the two embryonic globin genes  $\rho$  and  $\epsilon$  are located at the 5' and 3' ends of the chicken  $\beta$ -like globin gene cluster respectively. Roninson and Ingram (1982) suggest that a relatively recent gene conversion event in noncoding DNA 5' to the  $\epsilon$  gene is responsible for coordinate expression of  $\rho$  and  $\epsilon$  globin genes. Prior to this event they propose that the  $\epsilon$  globin gene was expressed in adults and the

physical arrangement of genes in the  $\beta$ -globin gene cluster coincided with the order of their expression, as in the chicken  $\alpha$ -globin gene cluster and mammalian globin gene clusters.

Jeffreys et al (1980) proposed the following model of globin gene evolution to account for the linkage of X. laevis  $\alpha$ - and  $\beta$ -globin genes: amino acids conservation in mammalian  $\alpha$ - and  $\beta$ -globins indicates that globin genes arose by duplication of a primordial gene 570 million years ago (Dayhoff, 1972). The linked X. laevis  $\alpha$ - and  $\beta$ -globin genes represent an early step in the evolution of separate  $\alpha$ - and  $\beta$ -globin gene clusters as found in birds and mammals. An event, such as chromosome duplication or a translocation between genes, then gave rise to two clusters of genes. According to the model this occurred approximately 300-350 million years ago in reptiles, prior to the emergence of mammals, but after divergence of the reptilian and amphibian lines of evolution, assuming that a single event was responsible. Subsequently duplication and divergence of genes and inactivation of the  $\alpha$ -globin gene in one cluster and  $\beta$ -globin gene in the other gave rise to separate  $\alpha$ - and  $\beta$ -globin gene clusters. Several testable predictions arise from this model (Jeffreys et al, 1980): (1) there should be inactivated  $\beta$  and  $\alpha$  globin genes at the 3' and 5' ends of the human  $\alpha$ - and  $\beta$ -globin gene clusters respectively; (2) some reptiles should show unlinked  $\alpha$ - and  $\beta$ -globin genes; (3) all fish should have linked  $\alpha$ - and  $\beta$ -globin genes, assuming they have not been separated by an independent event.

Hosbach et al (1982) have isolated clones from X. laevis DNA libraries containing linked tadpole and adult globin genes and



conclude that the genes are arranged as follows:  $\alpha_1$  tadpole - 5Kb -  $\alpha_1$  adult - 8Kb -  $\beta_1$  adult - 22Kb -  $\beta_2$  tadpole. Recently J. Williams and colleagues have also isolated genomic clones containing both the major  $\alpha_1$  and minor  $\alpha_2$  tadpole globin sequences linked 5' to the  $\alpha_1$  and  $\alpha_2$  adult globin genes respectively (J.G. Williams - personal communication). They have observed that the two clusters do not have identical organisation, the major tadpole  $\alpha$ -globin gene is separated from the major adult  $\alpha$ -globin gene by 5Kb whereas there are 11Kb between the minor tadpole and adult  $\alpha$ -globin sequences. They suggest that an insertion or duplication event in the minor globin gene cluster is responsible for this. It has not been possible to isolate clones containing adult and tadpole  $\beta$ -globin sequences from the clone banks prepared by Jeffreys et al (1980) and J. Williams (personal communication) assumes that this is because of the large distances separating the two genes (22Kb, Hosbach et al, 1982).

The linkage of Xenopus  $\alpha$ - and  $\beta$ -globin genes is the only example of different subunits of a multimeric protein being coded for by closely linked genes (Jeffreys et al, 1980). Bodmer (1979) observed that genes encoding different subunits of a single protein are generally unlinked e.g. human  $\alpha$  and  $\beta$  globins, immunoglobulin heavy and light chains amongst others. He suggested that it might be advantageous to the animal to separate such genes, however it is clearly not vital for the functioning of the protein, as linkage of  $\alpha$ - and  $\beta$ -globins in X. laevis demonstrates.

Globin genes are a family of related genes which are coordinately expressed during development. As discussed in the preceding sections multiple globin genes are expressed at any particular stage of development and different genes are expressed throughout development. It is relatively easy to isolate globin mRNA, since it constitutes a large percentage of the mRNA in RBCs, and therefore the study of globin gene expression is convenient for analysis of gene regulation. Much is now known about globin gene expression in mammalian and avian systems (for references see Section 1.2.c) and more recently X. laevis adult globin gene expression and organisation have been analysed (see Section 1.2 for references). There is little published information regarding X. laevis tadpole globin gene expression (Widmer et al, 1981; Hosbach et al, 1982).

The work described in this thesis is intended to provide some information about globin gene expression in X. borealis tadpoles. X. borealis was chosen in preference to X. laevis since the tadpoles are much larger than those of X. laevis. This means that more globin mRNA can be isolated from an individual, thus enabling sufficient material for analysis to be more readily isolated. RNA was prepared from blood cells of physiologically normal tadpoles which had not been treated with any chemicals to prevent metamorphosis or induce anaemia (see Sections 1.1.d and 1.1.e) and therefore the pattern of globin gene expression observed and the cDNA clones isolated are representative of normal development. This avoids the problems of deciding whether some globin species detected are induced by, for instance, phenylhydrazine treatment, as is known to be the case in

several species (see Chapter 3 for references).

Since X. borealis had not been previously utilised for such experiments it was first necessary to characterise the globins. Chapter 3 shows tadpole globins consist of three major and five minor species which are completely separable from X. borealis adult globins on triton-acid-urea polyacrylamide gels. Adult globin gene expression is estimated (Chapter 3) to start about stage 57 (Nieuwkoop and Faber, 1956) and cDNA clones were therefore prepared from 9S RNA isolated from stage 57-59 tadpoles (Chapter 4) in order to minimise contamination with adult globin sequences. 9S RNA isolated from these stages was shown, by in vitro translation, to code for all components seen on electrophoresis of tadpole RBC lysates. cDNA clones containing tadpole globin sequences were identified by colony hybridisation (Chapter 4). DNA sequencing (Chapter 6) identified  $\alpha$ - and  $\beta$ -globin cDNA sequences and permitted identification of  $\alpha$ - and  $\beta$ -globins on polyacrylamide gels by hybridisation-selection of homologous RNA (Chapter 5). Hybridisation of  $\alpha$ - and  $\beta$  tadpole globin cDNA sequences to RNA dot blots or RNA gel blots demonstrated that most globin mRNA is polyadenylated and that the two major RNA bands seen on formaldehyde-agarose gels code for  $\alpha$ - or  $\beta$ -globins (Chapter 5).

CHAPTER TWO

MATERIALS AND METHODS

All stock solutions are sterilised with 0.01% diethylpyrocarbonate or autoclaved. AnalaR chemicals are used where possible. Double distilled water is always used unless otherwise stated. Radiochemicals are obtained from Amersham International Ltd.

Barth X (Barth and Barth, 1959)

<u>Solution A</u>	51.50 g	NaCl
	0.75 g	KCl
	2.04 g	MgSO <sub>4</sub> ·7H <sub>2</sub> O
	0.62 g	Ca(NO <sub>3</sub> ) <sub>2</sub> ·H <sub>2</sub> O
	0.60 g	CaCl <sub>2</sub> ·2H <sub>2</sub> O

Made to 1l and divided into 100ml aliquots.

<u>Solution B</u>	2 g	NaHCO <sub>3</sub>
-------------------	-----	--------------------

Made to 500ml and divided into 50ml aliquots.

<u>Solution C</u>	24.22 g	tris base
	20 mg	phenol red

pH adjusted to 7.6 with HCl, made to 500ml and divided into 50ml aliquots.

Stock solutions are autoclaved and stored at room temperature. 1 bottle of each solution is mixed, the pH adjusted to about 7.6 and

made to 1l for 1 x Barth X. 10 µg/ml each streptomycin, ampicillin and gentamycin are added to 1 x Barth X.

Lysis Buffer

1 x is	0.1 M	NaCl
	5 mM	MgCl <sub>2</sub>
	20 mM	tris-HCl pH 7.5

10 x stock made and stored at 4°C.

NETS

1 x is	100 mM	NaCl
	10 mM	EDTA
	10 mM	tris-HCl pH 7
	0.2%	SDS

5 x stock made and stored at room temperature.

NTE

Made as for NETS without SDS.

SET

20 x SET is	3 M	NaCl
	30 mM	EDTA
	0.3 M	tris-HCl pH 7.5

TE

1 x is	10 mM	tris-HCl pH 8.0
	1 mM	EDTA

10 x stock made and stored at room temperature.

### Oligo dT Column Buffers

Binding buffer is 0.4M NaCl in NETS.

Elution buffer is NETS.

### Mix-C for cDNA Synthesis

(modified from Efsratiadis et al, 1976)

<u>μl</u>			
50	1	M	tris-HCl pH 8.3 measured at 37°C
40	0.5	M	DTT
30	0.2	M	MgCl <sub>2</sub>
20	3	M	NaCl
50	20	mM	dATP
50	20	mM	dGTP
50	20	mM	TTP
50	100	μg/ml (pT) <sub>10</sub>	
<u>460</u>		H <sub>2</sub> O	
800			

stored frozen at -20°C.

### Restriction Enzyme Buffers

10 x Bam HI	0.2	M	tris-HCl pH 8
	70	mM	MgCl <sub>2</sub>
	20	mM	2-mercaptoethanol
	1	M	NaCl
10 x Eco RI	1	M	tris-HCl pH 7.5
	0.1	M	MgCl <sub>2</sub>

	0.5 M	NaCl
	1 mg/ml	gelatin
10 x HindIII	0.2 M	tris-HCl pH 7.4
	70 mM	MgCl <sub>2</sub>
	0.6 M	NaCl
10 x PstI	0.2 M	tris-HCl pH 8
	0.5 M	NH <sub>4</sub> SO <sub>4</sub>
	0.1 M	MgCl <sub>2</sub>
10 x PvuII	60 mM	tris-HCl pH 7.5
	60 mM	MgCl <sub>2</sub>
	0.6 M	NaCl
	60 mM	2-mercaptoethanol
	1 mg/ml	gelatin
10 x Taq I	0.1 M	tris-HCl pH 8.4
	60 mM	MgCl <sub>2</sub>
	1 M	NaCl
	60 mM	2-mercaptoethanol
10 x universal	0.33 M	tris acetate pH 7.9
	0.10 M	Mg acetate
	0.66 M	K acetate
	0.10 M	DTT
	1 mg/ml	gelatin



5mM Amino Acid Mix

<u>Amino Acid</u>	<u>mg</u>
L-alanine	4.4
L-arginine	10.5
L-asparagine	6.6
L-aspartic acid	6.6
L-cysteine	6.0
glycine	3.8
L-glutamic acid	7.4
L-glutamine	7.3
L-histidine	9.6
L-isoleucine	6.6
L-leucine	6.6
L-lysine	9.1
L-methionine	7.5
L-phenylalanine	8.3
L-proline	5.8
L-serine	5.2
L-threonine	6.0
L-tryptophan	10.2
L-tyrosine	9.1
L-valine	5.9

Made to 10ml plus 10mM DTT

Amino acid(s) used for labelling are left out.

DNase I mix

100ng/ml	DNase I
50 mM	tris-HCl pH 8

5 mM MgCl<sub>2</sub>

Pol I mix-C

50 mM KPO<sub>4</sub> pH 7.4

7 mM MgCl<sub>2</sub>

1 mM DTT

33 mM dATP

33 mM dGTP

33 mM TTP

2.2

Microbiological Media

L-broth

Difco Bacto Tryptone 10g

Difco yeast extract 5g

NaCl 10g

Distilled H<sub>2</sub>O to make 1l

pH 7.2 with NaOH

L-broth agar (1.5%)

As for L-broth, plus 15g Difco agar.

pH is adjusted before agar is added.

BBL-bottom agar

NaCl 5g

BBL trypticase 10g

Difco agar 10g

Distilled H<sub>2</sub>O to make 1l

BBL-top agar

NaCl	5g
BBL trypticase	10g
Difco agar	6.5g
Distilled H <sub>2</sub> O to make 1l	

Minimal + glucose agar

300ml distilled H<sub>2</sub>O  
6g Difco Bacto agar

Autoclaved

Then add 100ml Spitzgen salts (see below)  
4ml 20% (w/v) glucose  
0.2ml 1mg/ml vitamin B1.

Spitzgen minimal salts

(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	100g
K <sub>2</sub> HPO <sub>4</sub>	700g
KH <sub>2</sub> PO <sub>4</sub>	300g
trisodiumcitrate	50g
MgSO <sub>4</sub>	10g
distilled H <sub>2</sub> O to make 10 litres	

M13 phage buffer

20mM	tris-HCl pH 7.9
20mM	NaCl
1mM	EDTA

## Antibiotic containing plates

Tc plates are L-agar plates supplemented with 25 µg/ml tetracycline.

Tc/Ap plates are L-agar plates supplemented with 25 µg/ml Tc, 50 µg/ml ampicillin.

Tc/Sm plates are L-agar plates supplemented with 25 µg/ml Tc, 200 µg/ml streptomycin.

50mg/ml tetracycline is stored at -20°C.

200mg/ml streptomycin is stored at 4°C.

2mg/ml ampicillin is stored at 4°C and made fresh every four weeks.

## 2.3

### General Methods

#### 2.3.a GEL ELECTROPHORESIS

##### Triton acid urea polyacrylamide gels

Globins are separated on 10% polyacrylamide gels containing 1% acetic acid, 3% triton and 7M urea (Alfageme et al, 1974). 30% acrylamide (BDH specially purified for biochemistry); 0.2% N,N'-methylene-bis-acrylamide stock solution is stored at 4°C. 70ml of gel is prepared as follows:

23.3ml acrylamide stock

29.4g urea

0.7ml acetic acid

Made to 66.5ml and degassed under vacuum.

2.1ml triton-X 100 - mixed by swirling at 37°C

117 µl TEMED

1.3ml 10% AMPS (made fresh every week)

Gels are poured between 25cm x 20cm glass plates separated by 1mm thick perspex spacers, with a comb already in place. Sometimes it is necessary to move the comb to remove air bubbles trapped below the teeth. Polymerisation commences after about 10-15 minutes. Gels are left for at least 1h before use.

Wells are filled with 7M urea in 7% acetic acid and gels prerun for 16-24h at 10mA, with reversed polarity (anode at the bottom), to remove ammonium persulphate. The electrode running buffer is 7% acetic acid in distilled water. Tank buffers are changed, wells rinsed with 7% acetic acid and loaded with about 20  $\mu$ l 0.2M 2-mercaptoethylamine which is run into the gel for 15 minutes at 10mA to scavenge for cyanate ions. Tank buffers are again changed, gels rinsed and electrophoresed at 10mA for a further 30 minutes. Buffers are again renewed and samples loaded using drawn out capillaries attached to a syringe.

Up to 5  $\mu$ l of sample is taken up in 10-15  $\mu$ l of 8M urea, 2% triton-X 100, 5% acetic acid, 5% 2-mercaptoethanol. When 10  $\mu$ l of sample has to be used the loading buffer is 10  $\mu$ l of 8M urea, 4% triton-X 100, 10% acetic acid, 10% 2-mercaptoethanol and 4.8mg urea is added to each tube. Samples are placed in a boiling water bath for 3 minutes, to denature the haemoglobin into its constituent globin monomers, just before loading the gel. Electrophoresis is for about 16h at 7mA constant current, with reversed polarity.

Gels are fixed for 30 minutes and 1h in two changes of a solution containing 50% methanol, 10% acetic acid. Staining is for 2h in the same solution plus 0.1% Coomassie blue R250 and gels are then washed for 10 minutes in fix, destained in several changes of 5% methanol, 7% acetic acid and finally in 50% methanol, which removes the last traces of background colour and shrinks the gels.

Gels are then photographed prior to being prepared for fluorography, as described by Laskey and Mills (1975). Gels are soaked for 2h in each of two changes of DMSO and then for a further 2h in a 22% solution of PPO in DMSO. These solutions are used five times each and then discarded (PPO is kept for recovery from the DMSO). PPO is precipitated in the gel by soaking in distilled water for 15-30 minutes. This time is not exceeded because the gels expand considerably in water and become too large to handle. Gels are dried on to Whatman 3MM paper under vacuum at about 80°C.

Preflashed film is placed on the gels and exposed at -70°C for the desired time. Films are developed in May and Baker Polycon Variable Contrast Xray developer for 5 minutes, fixed in May and Baker Perfix, High Speed Xray Fixer for 5 minutes and washed in cold, running water for 30 minutes.

#### Horizontal Agarose Slab Gels for DNA

10 x gel buffer is 0.4 M tris base  
20 mM EDTA  
0.2 M acetic acid  
pH ~ 8.3

Sample buffer is 50% (v/v) glycerol  
1% (w/v) ficoll  
0.1% (w/v) SDS  
25 mM EDTA  
bromophenol blue

Gel contains            0.8% to 1.4% (w/v) agarose  
                             1/10th volume 10 x gel buffer  
                             9/10th volume H<sub>2</sub>O

Agarose is melted by refluxing and poured when the solution has cooled to about 60-70°C.

Mini gels are made on 10cm x 14cm glass plates which have sticky tape around the edges, sealed with agarose, to make a tray. 60ml agarose solution plus 10µl 20mg/ml ethidium bromide is used per gel. For 0.2µg DNA samples a comb with teeth about 2mm wide is used. For larger samples (up to 2µg) wells are about 5mm wide. Mini gels are run at 250-300V for approximately 2h or until the dye has travelled about 5cm. The gel surface must always be covered with 1 x gel buffer to prevent overheating and care has to be taken that the wells do not dry out. Gels can be photographed immediately, without destaining, under UV illumination using Ilford FP4 film and a red filter. Exposed film is washed in Johnson Photowet, developed for 5 minutes in Ilford Microphen developer, stopped in 3% acetic acid and fixed for 5 minutes in Ilford Hypam fixer.

Larger agarose gels are made on perspex trays measuring 13cm x 23cm x 0.3cm. 130ml gel solution is used per gel. For rapid analysis of many samples a comb, with 5mm wide teeth, is positioned along the long side of the tray, thus 22 samples can be electrophoresed at once. For detailed restriction analysis and estimation of sizes of bands 12 samples are run lengthwise through the gel.

Gels are run, covered with plastic film to prevent wells drying up, at 70V constant voltage, the short ones for about 8h, the long ones overnight (about 16h). They are stained for 30 minutes in

1µg/ml ethidium bromide in distilled water, destained in distilled water (for 30 minutes or longer) and photographed. If necessary gels can be stored overnight, wrapped in plastic film, in the refrigerator before being photographed.

Acrylamide gels for DNA restriction fragments

40% acrylamide (BDH specially purified for biochemistry) and 2% N,N'-methylene-bis-acrylamide stock solutions are kept at 4°C. Gel and sample buffers are those used for agarose gels. 70ml of gel (see recipe below) is cast between glass plates as described on page 37.

Gel contains:

14ml	acrylamide
9.45ml	bisacrylamide
3.5 ml	glycerol
7 ml	10 x gel buffer
36 ml	H <sub>2</sub> O

This solution is degassed under vacuum and polymerised by addition of 350µl 10% AMPS, and 35µl TEMED. Gels are left for at least 1h prior to use.

Electrophoresis is overnight at 60V and is continued until the dye reaches about 1cm from the bottom of the gel. The final position of the dye is varied depending on the expected fragment sizes.

Gels are stained and photographed as already described on page 39.



Formaldehyde - agarose gels, for RNA (Lehrach et al, 1977)

10 x gel buffer is    0.2 M   Na-MOPS  
                          50 mM   NaOAc  
                          10 mM   EDTA

Sample buffer is    50%    formamide  
                          2.2 M   formaldehyde  
                          1 x gel buffer

10 x dyes is        0.2%   bromophenol blue in 20% ficoll

Gel contains:      1.5%   agarose  
                          1 x gel buffer  
                          2.2 M   formaldehyde (1/6th dilution of 37%  
  stock solution)

Agarose is melted in H<sub>2</sub>O, cooled to 60°C and buffer and formaldehyde are added in a fume cupboard. Gels are poured in perspex trays as already described for DNA agarose gels on page 39. It is best to leave gels on the perspex trays during electrophoresis as wells tend to break if they are moved.

Samples are dried down, under vacuum, taken up in 18µl sample buffer and heated to 55°C for 15 minutes. 2µl of 10 x dyes is added to each sample before layering under 1 x gel buffer into the wells. Gels are run at 70V constant voltage for about 14h. It is important to use running buffer only once or twice as the gels run much more slowly if older buffer is used.

Gels are stained with 5µg/ml ethidium bromide in 0.1M ammonium acetate, for 20-30 minutes and destained for at least 2h in 0.1M ammonium acetate. Destaining overnight results in slight diffusion of RNA bands, and does not significantly lower the background fluorescence. Gels are photographed under UV illumination as already described.

Laemmli polyacrylamide gels (Laemmli 1970)

A stock solution of 30% acrylamide, 0.2% N,N'-methylene-bis-acrylamide is kept at 4°C. A gel consists of a 5-20% (w/v) acrylamide separating gradient and a 5% (w/v) acrylamide stacking gel which extends a few cm below the comb. A gel is cast between glass plates as described on page 37. It is made as follows:

STOCK SOLUTIONS	5% ACRYLAMIDE	20% ACRYLAMIDE
Acrylamide/bis	4.2ml	16.8ml
1.875M tris-HCl pH 8.8	5.0ml	5.0ml
H <sub>2</sub> O	15.3ml	0.5ml
Glycerol	-	2.5ml
Degassed under vacuum		
10% SDS	0.25ml	0.25ml
TEMED	12.5 µl	6 µl
10% AMPS	125 µl	60 µl

The solutions are mixed, put one in each side of a gradient maker and pumped via two thin stainless steel tubes inserted between the glass plates. The tubes are slowly raised as the gel gradient is pumped in. The acrylamide is overlaid with TE-saturated butan-2-ol

and left to polymerise. Prior to pouring the stacking gel the butanol is washed away with distilled water and then 1 x running buffer (see below).

Stacking gel is made as follows:

23.3ml acrylamide stock solution

2.0ml 0.6M tris-HCl pH 6.8

14.4ml double distilled H<sub>2</sub>O

Degassed

0.2ml 10% SDS

10 µl TEMED

100 µl 10% AMPS

Sample buffer is

0.06M	tris-HCl pH 6.8
1%	2-mercaptoethanol
10%	glycerol
1%	SDS
	bromophenol blue

5 x running buffer is

144.9	g/l glycine
30	g/l tris base
0.5%	SDS

The gel is run at 8mA constant current for 15h and then fixed and stained as described on page 37.

### 2.3.b Precipitation of nucleic acids

RNA is ethanol precipitated by adjusting the solution to 0.3M sodium acetate (pH 5.5), adding two volumes of cold (-20°C)

ethanol, mixing and standing at  $-20^{\circ}\text{C}$  for at least two hours. RNA is recovered by centrifugation for 30 minutes at 10Krpm in a Sorvall HB4 swing out rotor, at  $4^{\circ}\text{C}$ .

DNA is isopropanol precipitated in a similar way by the addition of one volume of cold isopropanol. For rapid recovery of DNA the solution can be put in a methanol/dry ice bath for 10 minutes, or at  $-70^{\circ}\text{C}$  for 30 minutes. This is not advisable when very small quantities of DNA are involved.

### 2.3.c Phenol extraction

The following method of phenol extraction maximises recovery of nucleic acids, particularly when very small quantities (i.e. less than  $1\mu\text{g}$ ) are involved.

Phenol is redistilled, saturated with 1 x TE and stored in small aliquots at  $-20^{\circ}\text{C}$ . Prior to use phenol is warmed to  $37^{\circ}\text{C}$ . One tenth volume 4M NaCl and 0.4M EDTA and one volume of phenol are added to the solution to be extracted, which is also at  $37^{\circ}\text{C}$ , and vortexed. This mix is incubated at  $37^{\circ}\text{C}$  for 5 minutes, with vortexing intermittently to maintain the emulsion, spun briefly, (e.g. in the microfuge) to remove liquid from the walls of the tube, and chilled on ice. One volume of cold chloroform is added, vortexed and this is kept on ice for 5 minutes, again with occasional vortexing. Aqueous supernatant is separated from the phenol/chloroform by centrifugation and removed to a clean tube. The phenol/chloroform is back extracted with one volume NTE and the supernatant added to the original supernatant. Two volumes chloroform are added to the combined supernatants, vortexed and kept on ice for 5 minutes. The aqueous layer is recovered by centrifugation and ethanol or isopropanol precipitated.

### 2.3.d TCA and CTAB precipitation

Incorporation of radioactive label into proteins or nucleic acids is checked by precipitating a sample with either TCA or CTAB.

For TCA precipitation of DNA or RNA a sample is taken into 0.5ml H<sub>2</sub>O and 0.5ml carrier DNA (150µg/ml salmon sperm DNA) and 1ml 10% (w/v) TCA are added. The solution is vortexed and left on ice for 10 minutes.

TCA precipitation of proteins is essentially the same except that one tenth volume of 1M NaOH containing 10mg/ml of the same amino acid used for labelling is added prior to TCA addition in order to discharge tRNA.

For CTAB precipitation a sample is taken into 1ml H<sub>2</sub>O then 0.5ml yeast tRNA (2mg/ml in 1M sodium acetate pH 4.75) and 1ml 4% CTAB are added, at room temperature. The mixture is vortexed and processed immediately.

Precipitates are collected by vacuum filtration on Whatman GF/C glass fibre filters. Tubes and filters are washed: TCA precipitated samples with 5% TCA, then ethanol; CTAB precipitated samples with H<sub>2</sub>O. Filters are dried in a vacuum oven at about 80°C, put into glass scintillation vials, covered with scintillant (8g/l butyl-PBD in toluene) and counted.

### 2.3.e Sephadex G-75 column fractionation of incorporated and unincorporated nucleotides

After labelling DNA or RNA with a radioactive nucleotide (e.g. [ $\alpha$  <sup>32</sup>P]dCTP) it is necessary to separate unincorporated from incorporated label. Passing the reaction mix over a column of Sephadex G-75 gives rise to two peaks of radioactivity, the first

consists of larger molecules with incorporated label which are excluded from the Sephadex beads, the second contains the unincorporated nucleotides.

A siliconised glass column, about 18cm long and 8mm in diameter, is packed with Sephadex G-75 which has been equilibrated by washing several times with the required buffer. The buffer used depends on what the labelled sample is to be used for, details are given later in the relevant sections. The Sephadex is allowed to settle in the column under gravity at first and then buffer is run through to pack it down further.

Samples are applied, in about 100 $\mu$ l, to the surface of the G-75 and allowed to run in. Sample tubes are washed with about 100 $\mu$ l column buffer and this is also run into the column. Column buffer is then put gently on top of the G-75, taking care to minimise disturbance to the surface. Fractions are collected and aliquots from each CTAB precipitated to determine the position of fractions containing incorporated label. The column is then washed extensively with column buffer to remove all the free nucleotides, this can be conveniently monitored using a Geiger counter to detect the radioactive nucleotide. The column can be stored by adding a few drops of 2M sodium azide to the washing buffer, which prevents any bacterial or fungal growth. Before further use the sodium azide must be washed out of the column.

#### 2.3.f Deionising formamide

BioRad analytical grade mixed bed resin AG 501-X8(D), 20-50 mesh is used to deionise formamide. A small amount of resin is added to the solution and stirred at room temperature for 1h. If the indicator in the beads all turns from blue to brown more resin is

added until some beads remain blue. The solution is then filtered through Whatman filter paper and used immediately or stored in the refrigerator (for not more than two days) until required.

### 2.3.g Preparation of carrier calf thymus DNA

Calf thymus DNA is dissolved in TE and phenol extracted several times until the interface is clear. It is then ethanol precipitated, taken up in TE, sonicated and stored at  $-20^{\circ}\text{C}$  in small aliquots.

## 2.4

### Animals

Xenopus borealis tadpoles are raised from our own matings. Mature male and female frogs are injected in the evening with 150-300 units, depending on the animals' size, of human chorionic gonadotrophin (trade-name Pregnyl, from Organon). Two males and one female are kept in a small tank overnight. A perforated tray is positioned about 2cm from the bottom of the tank to prevent the animals damaging the eggs. Eggs are laid approximately 12h after hormone injection. The following morning adult animals are removed and embryos raised for about four days in the tank in which mating took place. After four days young tadpoles are swimming freely and are transferred to larger tanks by means of a round-ended pasteur pipette.

Tadpoles are fed on a mixture of nettle powder (250ml) and complan (100ml) which is homogenised in tap water and made up to about 1 litre. The food is stored at  $4^{\circ}\text{C}$ . The amount given to the animals is adjusted so that water in the tanks clears overnight, initially 5-10ml increasing to about 100ml per day just before metamorphosis. Over feeding gives rise to bacterial growth depriving tadpoles of oxygen which causes them to float on the surface and if

they are not immediately given clean water they may die. Water in the tanks is supplemented with about 1g ferrous ammonium sulphate and kept at approximately 23°C. Tanks are cleaned out only when necessary, i.e. when the tadpoles float, as they feed on algae which accumulate in the tanks. Small tadpoles are removed from tanks before water is pumped out, but larger ones are left in and water pumped out through a sieve.

## 2.5 Globin synthesis in isolated red blood cells

Blood cells are isolated from MS222 anaesthetised tadpoles and young frogs by heart puncture using siliconised boiling point capillaries which are drawn out to a point. The diameter of the drawn out capillary is kept as large as possible (about 0.5mm) in order to avoid lysing blood cells. Cells from each individual are kept separately in snap-cap tubes (Sarstedt). Tadpoles' developmental stages are determined by comparison with the tables of X. laevis development (Nieuwkoop and Faber, 1956). Young frogs are kept separately after they reach stage 66 and their age expressed in weeks after this time.

Blood cells are isolated and washed twice in a solution of Barth X and heparin containing 0.7µM NH<sub>4</sub>FeSO<sub>4</sub>. Heparin prevents blood from coagulating and iron maintains the cells' synthetic activity. Cells from each individual are split into two and resuspended in 300 µl of a medium optimised for globin synthesis, containing either [<sup>35</sup>S]-methionine (specific activity 7-10mCi/ml, > 800 Ci/mmol) or [<sup>3</sup>H]-leucine (specific activity 1mCi/ml, 61 Ci/mmol) : (200µl 10 x Eagles; 100µl 200mM Hapes pH 7.5; 200µl 5mM amino acids minus leucine or methionine; 200µl 40mM MgCl<sub>2</sub> in 200mM glucose; 200µl 2mM NH<sub>4</sub>FeSO<sub>4</sub>; 200 µl 20mM sodium citrate; 400µl



20mg/ml Sm plus penicillin; 2400  $\mu$ l foetal bovine serum; 600  $\mu$ Ci [<sup>3</sup>H]-leu or [<sup>35</sup>S]-met; H<sub>2</sub>O to 6000  $\mu$ l). The cells are incubated overnight at 25°C, on a rotating wheel in order to keep them suspended in the incubation medium. They must not be mixed too vigorously otherwise they lyse during the labelling period.

Globin is prepared by washing cells once in Barth X and once in lysis buffer followed by lysis in 100 $\mu$ l lysis buffer plus 1/40th volume 10% Nonidet NP40. Nuclei are spun down (2 minutes spin in microfuge) and the supernatant removed, one volume of glycerol added and the samples stored at -20°C.

Labelled globin samples are analysed on triton-acid-urea polyacrylamide gels by fluorography as described on page 36.

## 2.6 Preparation of *X. borealis* tadpole blood cell 9S mRNA

Stage 57-59 tadpoles (Nieuwkoop and Faber, 1956) are anaesthetized in MS222 and blood collected from the heart into a solution of Barth X and heparin on ice, as described in the previous section. Cells and solutions are kept at 4°C throughout the following procedure. Cells are washed twice in Barth X, twice in lysis buffer and resuspended in five cell volumes of lysis buffer plus 1/10th volume of a ribonucleoside-vanadyl complex to inhibit RNase activity (Berger and Birkenmeier, 1979). After lysis, by the addition of another five cell volumes lysis buffer containing 1/40th cell volume 10% Nonidet NP40, nuclei are removed by centrifugation (10 Krpm, 10 minutes, 4°C, Sorvall HB4 swing out rotor). The supernatant is adjusted to 20mM EDTA, 0.5% SDS and added to ten cell volumes NETS at room temperature. Total cytoplasmic RNA is prepared by two phenol-chloroform extractions (Penman, 1966) and removed from the aqueous phase by ethanol precipitation.

Approximately 100 OD260nm units of RNA are dissolved in 3ml NETS and loaded onto a 58ml gradient of 7-30% sucrose in NETS. Centrifugation is in a 3 x 65ml MSE swing out rotor for 22h at 23.5 Krpm at 25°C. The gradient is collected in 2.5ml fractions and peak poly A plus fractions, determined by <sup>3</sup>H-poly U hybridisation (Bishop et al, 1974) are pooled.

Poly A minus RNA is purified from the poly A plus peak by oligo dT column fractionation (Cabada et al, 1977). Both poly A plus and poly A minus RNA fractions are further purified on 4.2ml 7-30% sucrose in NETS gradients. (6 x 5ml Ti MSE swing out rotor, 4h, 50 Krpm, 25°C). Peak 9S fractions are pooled, as determined from the UV trace of the gradients.

RNA is stored in 0.3M sodium acetate pH 5.5 plus two volumes of ethanol, at -20°C, since it degrades rapidly if stored frozen in H<sub>2</sub>O. Prior to use RNA is washed in H<sub>2</sub>O followed by ethanol precipitation in 0.3M sodium acetate, three times in order to remove any residual SDS.

## 2.7 Translation of *X. borealis* tadpole 9S mRNA

RNA is translated in micrococcal nuclease treated rabbit reticulocyte lysate essentially as described by Pelham and Jackson (1976). One ml of lysate is treated on ice as follows: 25µl haemin stock (made by dissolving 0.62mg haemin in 50µl KOH, adding 0.9ml propylene glycol, 20µl 1M tris, 30µl double distilled water and adjusting to pH 7.2 with HCl) is added to the frozen lysate. After thawing 10µl of creatin kinase (5mg/ml in 50% glycerol) is added followed by a mix containing 50µl 0.2M creatin phosphate; 50µl 2M KCl and 10mM MgCl<sub>2</sub>; 50µl amino acid mix minus leu and minus his; 14µl 50mM ATP; 14µl 50mM GTP. At this stage a 25µl aliquot is

removed to check incorporation of labelled amino acids into endogenous RNA. 10  $\mu$ l 0.1M  $\text{CaCl}_2$  and 5  $\mu$ l 1mg/ml micrococcal nuclease is added to the remaining lysate which is incubated at 20°C for 15 minutes. The lysate is kept on ice while a 25  $\mu$ l aliquot is assayed to check that incorporation of labelled amino acids into endogenous RNA has been substantially reduced compared to the endogenous incorporation of untreated lysate. This assay is carried out for 40 minutes at 30°C using 2.5  $\mu$ Ci each of [ $^3\text{H}$ ]-leucine (61 Ci/mmol, 1m Ci/ml) and [ $^3\text{H}$ ]-histidine (58 ci/mmol, 1 mCi/ml). When endogenous incorporation has been reduced sufficiently 10  $\mu$ l 0.2M EGTA, neutralised with KOH, is added to the lysate to chelate  $\text{Mg}^{2+}$  and therefore prevent micrococcal nuclease from working. The lysate is then used immediately since storage at -70°C often severely reduces its activity.

1  $\mu$ g of X. borealis tadpole blood cell 9S (Xbt9S) poly A plus or poly A minus RNA is translated in 25  $\mu$ l treated lysate with 25  $\mu$ Ci each [ $^3\text{H}$ ]-leucine and [ $^3\text{H}$ ]-histidine. Labelled amino acids are dried down together and taken up in the required amount of lysate. RNA is added and the samples incubated at 30°C for up to 2h. 1  $\mu$ l samples are taken at intervals into 1ml  $\text{H}_2\text{O}$ , 180  $\mu$ l 10mg/ml each leucine and histidine in 1M NaOH and 20  $\mu$ l  $\text{H}_2\text{O}_2$  are added and the samples incubated for 5 minutes at 30°C to remove haemoglobin red which would interfere with scintillation counting. TCA precipitation and counting is as described on page 45. Remaining mix is stored at -20°C.

Translation products are analysed on triton-acid-urea gels by fluorography as described on page 36. The amount of sample loaded into each slot is adjusted so that the cpm in each track are the same.



## 2.8 Preparation of cDNA clones from X. borealis tadpole 9S poly

### A plus RNA

This is summarised in figure 15.

### 2.8.a cDNA synthesis

Initially cDNA is made from a small amount of RNA to check the reaction is working.  $1\mu\text{Ci}$  [ $^3\text{H}$ ]dCTP ( $1\text{mCi/ml}$ ,  $21\text{ Ci/mmol}$ ) is dried down and the following solutions added:  $20\mu\text{l}$  Mix-C (page 31),  $1\mu\text{l}$   $2\text{mM}$  dCTP,  $0.5\mu\text{g}$  RNA,  $2\mu\text{l}$  reverse transcriptase (i.e. 16 units, provided by Dr. J.W. Beard),  $1\mu\text{l}$   $\text{H}_2\text{O}$ . A control reaction mix is also made up in which RNA is replaced by  $\text{H}_2\text{O}$ . Mixes are incubated at  $37^\circ\text{C}$  for 45 minutes, then TCA precipitated and counted.

Larger quantities of cDNA are prepared as follows:  $10\mu\text{g}$  RNA is spun down ( $10\text{ Krpm}$ , 30 minutes,  $4^\circ\text{C}$ , Sorvall HB4 swing-out rotor) and the ethanol supernatant discarded.  $10\mu\text{Ci}$  [ $\alpha^{32}\text{P}$ ]dCTP is added to the RNA and dried down. RNA and label are dissolved in  $200\mu\text{l}$  Mix-C,  $20\mu\text{l}$   $20\text{mM}$  dCTP and  $10\mu\text{l}$   $\text{H}_2\text{O}$ , this is divided between nine tubes and  $2\mu\text{l}$  reverse transcriptase are added to each tube. Incubation is at  $37^\circ\text{C}$  for 30 minutes,  $46^\circ\text{C}$  for 30 minutes and  $51^\circ\text{C}$  for 30 minutes (modified from Zain et al, 1979).

cDNA/RNA is run on a Sephadex G-75 column in  $0.03\text{M}$  NaCl;  $5\text{mM}$  tris-HCl pH 7.0. Approximately  $0.5\text{ml}$  fractions are collected into siliconised scintillation vials and Cerenkov counted. Excluded peak fractions are pooled and ethanol precipitated.

### 2.8.b Tailing cDNA/RNA hybrids with dCTP

Under conditions described by Roychoudhury et al (1976) terminal deoxynucleotidyl transferase was used to add a homopolymer tract of dCTP to the 3' termini of cDNA/RNA hybrids.

cDNA/RNA is spun down, dried in a vacuum desiccator and

dissolved in 30  $\mu\text{l}$   $\text{H}_2\text{O}$ . To 10  $\mu\text{Ci}$  of dried [ $\alpha^{32}\text{P}$ ]dCTP is added 5  $\mu\text{l}$  cDNA/RNA (about 0.4  $\mu\text{g}$ ), 5  $\mu\text{l}$  10 x terminal transferase buffer (0.14M cacodylic acid, 0.3M tris-HCl pH 7.6, 1mM DTT), 5  $\mu\text{l}$  10mM  $\text{CoCl}_2$ , 2  $\mu\text{l}$  2mM dCTP, 33  $\mu\text{l}$   $\text{H}_2\text{O}$  and 2  $\mu\text{l}$  (i.e. 17 units of BRL enzyme) terminal transferase. The reaction is incubated at 37°C for 30 minutes. Samples are taken at 0 and 30 minutes and incorporation determined by TCA precipitation. Unincorporated nucleotides are removed on a Sephadex G-75 column, run in  $\text{H}_2\text{O}$ . Approximately 100  $\mu\text{l}$  fractions are collected and Cerenkov counts determined. The first peak fractions are pooled and ethanol precipitated.

#### 2.8.c 7-30% sucrose gradient of tailed cDNA/RNA

dC-tailed cDNA/RNA is spun down, dissolved in 100  $\mu\text{l}$  NTE and loaded onto a 2.2ml gradient of 7-30% sucrose in NTE. The tube is rinsed with 50  $\mu\text{l}$  NTE and this is added to the gradient. Centrifugation is in an MSE 3 x 3ml A1 swing out rotor at 38 Krpm for 10h at 4°C. Two drop fractions are collected from the gradient into sterilised, siliconised scintillation vials and Cerenkov counts determined. Fractions containing the largest molecules are pooled and ethanol precipitated with 3  $\mu\text{g}$  E.coli tRNA, those containing smaller molecules are stored frozen.

#### 2.8.d Preparation of pAT 153 for cloning

pAT 153 is cut with Pst I as follows: 100  $\mu\text{l}$  pAT 153 DNA (i.e. 100  $\mu\text{g}$ ), 100  $\mu\text{l}$  Pst I (i.e. 100 units), 25  $\mu\text{l}$   $\text{H}_2\text{O}$  and 25  $\mu\text{l}$  10 x Pst I buffer are incubated at 37°C for 2h. 300  $\mu\text{l}$  TE is added and the solution is phenol extracted. DNA is precipitated from the supernatant with ethanol (two volumes) at -70°C for 3h, collected

by centrifugation and passed over a Sephadex G-25 column in H<sub>2</sub>O. 350  $\mu$ l fractions are collected and peak fractions, determined from OD260nm, pooled and ethanol precipitated.

Pst I cut pAT 153 is checked in a transformation assay (see later) to ensure that the number of transformants per  $\mu$ g is reduced compared with uncut pAT 153.

Pst I cut pAT 153 is tailed with dGTP (Roychoudhury *et al*, 1976) in the following reaction mix: 10 $\mu$ Ci [<sup>3</sup>H]dGTP dried down, 20  $\mu$ l cut pAT 153 (i.e. 20 $\mu$ g), 10 $\mu$ l 10 x terminal transferase buffer, 10  $\mu$ l 10mM CoCl<sub>2</sub>, 4  $\mu$ l 2mM dGTP, 55  $\mu$ l H<sub>2</sub>O and 0.5  $\mu$ l terminal transferase. This is incubated at 37°C for 6h and 1 $\mu$ l samples TCA precipitated at intervals. 1 $\mu$ l of terminal transferase is added after 1h 15 minutes and again after 2h 15 minutes.

Tailed pAT 153 is run on a Sephadex G-75 column in H<sub>2</sub>O. 200  $\mu$ l fractions are collected and 2 $\mu$ l from each counted. Fractions in the first peak are pooled and OD260nm determined to estimate percentage recovery. The DNA was dried down and taken up in H<sub>2</sub>O at a concentration of about 0.1 $\mu$ g/ $\mu$ l.

#### 2.8.e Annealing reactions

cDNA/RNA fractions are spun down, dried in a vacuum desiccator and taken up in H<sub>2</sub>O.

Annealing mixes are as follows:

	[dC]cDNA/RNA	Sterile			
		pAT 153	4M NaCl	10 x TE	H <sub>2</sub> O
1.	10μl i.e. 0.025μg	2.5 μl 0.25μg dG-pAT	25μl	50μl	412.5μl
2.	-	0.5 μl 0.05μg dG-pAT	5μl	10μl	84.5μl
3.	-	0.01 μg or 0.001μg uncut pAT	5μl	10μl	84.5μl

1. dC tailed-cDNA/RNA hybrid is annealed to dG-tailed, Pst I cut pAT 153.

2. Control to show background level of transformation with dG-tailed Pst I cut pAT 153 alone.

3. Controls to check that transformation is working efficiently with uncut pAT 153.

The mixes are incubated at 60°C for 2h then left to cool overnight at room temperature.

#### 2.8.f Transformation of E. coli HB101 with annealed pAT 153-cDNA

A single HB101 (Boyer and Roulland-Dussoix, 1969) colony is picked into 10ml L-broth and grown up at 37°C overnight with shaking. 0.5ml of overnight culture is diluted into 25ml L-broth and grown to a cell density of 40 measured in a Klett-Summerson photometer. Cells are chilled for 10 minutes on ice, spun down (Sorvall HB4 rotor 5 Krpm, 4°C, 5 minutes), washed in 10ml 0.1M MgCl<sub>2</sub>, resuspended gently in 10ml cold 50mM CaCl<sub>2</sub> and left on ice for 15 minutes. They are collected by centrifugation, resuspended in 3ml cold 50mM CaCl<sub>2</sub> and left on ice for at least 30 minutes prior to use.

200µl of competent cells are added to samples of DNA from each annealing reaction and left on ice for 30 minutes. They are heat shocked at 42-45°C for 2 minutes and put at 37°C for 2 minutes. One ml of L-broth, prewarmed to 37°C, is added to each mix and incubated for 2h at 37°C. DNA mixes are added to 2.5ml aliquots of BBL top agar plus 15 µg/ml Tc, poured onto 15 µg/ml Tc plates and incubated at 37°C overnight.

#### 2.8.g Screening HB101 Transformants for ampicillin sensitivity

Transformants are picked onto Tc plates. Approximately 80 are arranged in a grid on one plate. Colonies are grown overnight at 37°C, replica plated onto Tc/Ap and Tc plates and grown overnight. Those colonies which appear to be Tc<sup>R</sup>Ap<sup>S</sup> are picked onto fresh Tc plates and rescreened for ampicillin sensitivity.

2.9.

#### Colony Hybridisation

Tc<sup>R</sup>Ap<sup>S</sup> transformants are screened by colony hybridisation



(modified from Hanahan and Meselson, 1980) to identify those containing DNA complementary to Xbt9S and Xba9S RNA.

Nitrocellulose filters (Schleicher and Schuell) are cut to fit petri dishes, washed three times by boiling in H<sub>2</sub>O and autoclaved in a sealed package between layers of damp blotting paper. Sterile filters are placed on Tc/Sm plates and dried at 37°C overnight. Colonies to be screened are transferred from master grids (approximately 60 colonies per plate) to nitrocellulose by replica plating. Two filters and one master plate are made from each replica and grown overnight at 37°C.

Filters are placed for 3 minutes each on two successive dishes of blotting paper soaked in 0.5M NaOH to denature the colonies. They are washed in the following solutions: 1M tris-HCl pH 7.4 for 3 minutes; 1.5M NaCl, 0.5M tris-HCl pH 7.4 for 3 minutes; 0.3M NaCl for 1 to 2 minutes, air dried and baked between sheets of blotting paper in a vacuum oven at 80°C for 2h. After baking some filters stick to the blotting paper, but are easily removed by wetting the paper with 0.3M NaCl. Filters are stored between layers of blotting paper at room temperature.

Colonies are screened with kinase labelled RNA (Donis-Keller et al, 1977). Xbt9S poly A plus RNA, Xbt9S poly A minus RNA, Xba9S poly A plus RNA and X. tropicalis blood cell rRNA were used. RNA is first treated with sodium hydroxide to increase the number of 5' ends available to accept a labelled phosphate: 2.5µg RNA is taken up in 50 µl H<sub>2</sub>O plus 5µl 1M NaOH and left on ice for 1h. The solution is neutralised with 5µl 1M tris-HCl pH 8.0, 5µl 1M HCl and ethanol precipitated at -20°C overnight by addition of 1µl 4M NaCl and 125µl ethanol.

RNA is spun down (10 minutes in microfuge), dried in a vacuum desiccator and the 5' labelling procedure of Donis-Keller et al (1977) followed: RNA is dissolved in 70µl 10mM tris HCl pH 7.4, 1mM spermidine, 0.1mM EDTA, heated to 50°C for 3 minutes and chilled in ice-water. 10µl 10 x kinase buffer (0.5M tris-HCl pH 9.0, 0.1M MgCl<sub>2</sub>, 0.05M DTT), 10µl [ $\gamma$ <sup>32</sup>P] rATP (i.e. 100 µCi 10mCi/ml, > 5000 Ci/mmol), 5 µl 0.02mM ATP, 4 units T4 polynucleotide kinase and H<sub>2</sub>O to make 100µl final volume are added and incubation is at 37°C for 1h. A 1 µl sample is TCA precipitated to check incorporation and to the rest is added 100µl 3M sodium acetate pH 5, 4 µl 5mg/ml E. coli tRNA and 600µl ethanol. This is stood at -20°C for at least 30 minutes.

Labelled RNA is spun down, dried, taken up in 5 x SET and passed over a column of Sephadex G-75 in 5 x SET. Approximately 200µl fractions are collected, 1 µl from each TCA precipitated and peak fractions pooled and used as a hybridisation probe.

Filters are prehybridised at 37°C in about 200ml of 5 x SET plus 50% dimethylformamide for at least 30 minutes. They are hybridised in a polythene bag containing 15ml of 5 x SET, 50% dimethylformamide plus the probe. Up to 12 filters are put together in one bag. Hybridisation is at 37°C overnight with the bag clamped vertically between two perspex sheets so that air bubbles are forced above the filters.

Filters are washed at 37°C in the following solutions for 30 minutes each: 200ml 5 x SET; 200ml 2 x SET; 200ml 2 x SET plus 20 µg/ml RNase A (heat treated at 85°C, 15 minutes to remove any DNase activity); 500ml 2 x SET; then 500ml 2 x SET overnight. They are dried, covered in Saran Wrap and preflashed Kodak X-omat XS-1 X-ray film is exposed to the filters at -70°C, with an intensifying

screen.

## 2.10 Determination of Insert Sizes in *X. borealis* tadpole globin

### cDNA clones

Small samples of plasmid DNA are prepared from randomly selected clones essentially as described by Birnboim and Doly (1979).

10ml L-broth plus 25µg/ml Tc and 200µg/ml Sm is inoculated with a single colony and grown with shaking at 37°C overnight. Cells are harvested by centrifugation at 5 Krpm for 5 minutes at 4°C (Sorvall HB4 swing out rotor), resuspended by vortexing in 300µl lysis solution (2mg/ml lysozyme, 25mM tris-HCl pH 8.0, 10mM EDTA pH 8.0, 10% glucose) and left on ice for 30 minutes. 600µl alkaline SDS (0.2M NaOH, 1% SDS) is then added and the clear, viscous suspension left on ice for 5 minutes. 450µl high salt solution (3M sodium acetate pH 5) is added, the suspension mixed gently and left on ice for 1h, with occasional mixing. The coarse white precipitate is removed by centrifugation at 10 Krpm for 20 minutes. Approximately 1200µl of supernatant is taken into a clean tube, care being taken to avoid any floating precipitate. The supernatant is ethanol precipitated at -20°C for 30 minutes, DNA collected by centrifugation, taken up in 100µl 0.3M sodium acetate pH 5.0 and reprecipitated. Plasmid DNA is spun down (10 minutes, microfuge), dried and taken up in 130µl TE, giving an approximate concentration of 1µg plasmid DNA per 10µl.

Plasmid DNAs are digested with Pst I and EcoRI separately to obtain an estimate of the cDNA insert size. RNase A (10µg/ml) is added to digests to remove RNA which runs with small DNA fragments on agarose gels. Approximately 20 units of restriction enzyme is necessary to digest 1µg of plasmid DNA prepared in this way.

Digestion products are separated on 1.4% agarose gels as described on page 38.

## 2.11 Preparation of Plasmid DNA for detailed Analysis

Larger quantities of DNA are prepared from a scaled up version of the Birnboim preparation. A 10ml overnight culture is grown up under selection and diluted the following day into 500ml L-broth plus 25 µg/ml Tc and 200 µg/ml Sm. This is shaken at 37°C to an approximate OD<sub>550nm</sub> of 0.8 to 1.0 and then 50µg/ml chloramphenicol added and the cells left shaking overnight at 37°C. Cells are chilled on ice, harvested by centrifugation (Sorvall HS4 swing-out rotor, 5 Krpm, 5 minutes, 4°C) and resuspended in 10ml lysis solution. 20ml alkaline SDS and 15ml high salt solution are used as described for mini-preps on page 59.

After isopropanol precipitation DNA is banded in caesium chloride (Radloff et al, 1967). It is dissolved in 5ml TE and 0.5ml 20mg/ml ethidium bromide is added. Approximately 6g CsCl<sub>2</sub> is dissolved in each DNA solution and the density adjusted, by the addition of more CsCl<sub>2</sub> or TE, so that 5ml weighs between 7.6g and 8.0g. DNA solutions are put into centrifuge tubes with stainless steel caps and the tubes filled with parafin oil. Centrifugation is for 48h at 20°C at 38 Krpm in an MSE 8 x 40 Ti or 10 x 10 Ti fixed angle rotor. Plasmid bands are collected under UV illumination by inserting a needle through the tube just below the bands and collecting the DNA into plastic syringes. Ethidium bromide is removed with TE-saturated butan-1-ol by shaking the DNA solution with ten successive volumes of butanol, discarding the top layer each time. DNA solutions are dialysed against six 1l changes of TE to remove butanol and CsCl<sub>2</sub>, isopropanol precipitated and taken up at

about 1mg/ml in TE. DNA concentration is determined from optical density and a sample run on an agarose gel. Samples of DNA to be used frequently are stored at 4°C, the main stock is kept frozen at -20°C.

### 2.12 Restriction analysis of clones with inserts greater than 600 nucleotides

Those clones with inserts of 600 nucleotides or more are selected for further study and sorted into classes based on restriction sites they contain. Single digests with BamHI, EcoRI, Hind III, Pst I and Pvu II provide this data. Sites within each clone are mapped with respect to each other by double digests of DNA from a member of each class. Buffers used for double digests are: BamHI/Pst I, Hind III/BamHI, BamHI buffer; Hind III/Pst I, Hind III buffer; EcoRI/Pst I, EcoRI/Hind III, EcoRI buffer; Pvu II/Taq I, Pvu II/Bam HI Taq I/Pst I, universal buffer. All reactions are at 37°C except when Taq I is included in a double digest: incubation is then at 37°C with the first enzyme, then Taq I is added and the temperature raised to 65°C. Digest products are analysed on 1.4% agarose or 8% acrylamide gels depending on expected fragment sizes.

### 2.13 Hybridisation - selection of RNAs homologous to cloned sequences

DNA is partially depurinated (Kindle and Firtel, 1978) to prevent reannealing and increase the hybridisation efficiency. 10µl DNA (i.e. 10 µg), 4 µl 1M sodium acetate pH4.2 and 6 µl H<sub>2</sub>O are incubated for 15 min at 55°C. The DNA is then denatured in 0.2M NaOH by the addition of 10 µl H<sub>2</sub>O, 50 µl 20 x SET and 20 µl 1M NaOH. This is left for 5 minutes at room temperature, chilled on ice and

neutralised with 100 $\mu$ l NB (1M HCl (1): 1M tris-HCl pH 7.4 (10): 20 x SET (10)). Samples are then frozen in dry ice/methanol.

DNA is applied under slight vacuum to small areas (about 3-4mm square) of nitrocellulose (Schleicher and Schuell 0.45 $\mu$ m pore diameter) which has been extensively washed with H<sub>2</sub>O and 10 x SET. The DNA is kept on ice and loaded from a capillary as it thaws, thus reducing the risk of it reannealing. The filter is washed two or three times with 10 x SET followed by 3 x SET. Areas containing DNA are cut out, placed in sterilised Sarstedt tubes and baked for 3-4h at 80 $^{\circ}$ C in a vacuum oven.

Hybridisations are modified from the method described by Ricciardi et al, 1978. Filters are prehybridised individually for 1h at 50 $^{\circ}$ C, in Sarstedt tubes, in 70% formamide (BDH, specially purified for biochemistry, deionised just before use), 5 $\mu$ g/ml poly A 0.4M NaCl, 0.01M Pipes pH 7.0. They are then transferred to one tube and hybridised at 50 $^{\circ}$ C overnight in 100 $\mu$ l of the same buffer plus 5 $\mu$ g Xbt9S poly A plus mRNA. All filters must be covered with buffer.

The buffer is removed to a clean tube and kept as the "unbound" fraction. Filters are washed in the following solutions by vortexing and discarding the buffer: 10 x 1ml 1 x SET, 0.5% SDS at 65 $^{\circ}$ C; 2 x 1ml 10mM tris pH 7.8, 2mM EDTA at 50 $^{\circ}$ C. They are finally washed in 10mM tris pH 7.8, 2mM EDTA for 3 minutes at 50 $^{\circ}$ C and put into individual tubes. Bound RNA is eluted with 300 $\mu$ l double distilled water at 100 $^{\circ}$ C for 1 minute, frozen in dry ice/methanol and ethanol precipitated with 2 $\mu$ g Xenopus tRNA and three volumes ethanol. RNA pellets are washed twice with 70% ethanol, dried and taken up in 4 $\mu$ l double distilled water (unbound RNA is taken up in 10 $\mu$ l). 1 $\mu$ l of each sample is translated in rabbit reticulocyte lysate as described on page 50.

Filters are dried and stored in sterile Sarstedt tubes as they can be used several times.

## 2.14 Hybridisation of RNA bound to nitrocellulose

### 2.14.a Binding RNA to filters

#### Dot Blots

RNA is spotted directly onto nitrocellulose as described by Thomas (1980). Nitrocellulose is wet with H<sub>2</sub>O equilibrated by soaking for 1h, in 20 x SET and dried under a lamp. Varying amounts of RNA in 1 µl of H<sub>2</sub>O are spotted, through a grid, onto the nitrocellulose. Outer edges of the grid are marked in pencil on the nitrocellulose for later reference. A range from 10 to 200ng poly A plus and 100 to 500ng poly A minus Xbt9S RNA is used. Equivalent amounts of rRNA (from X. borealis tadpole blood cells) are used as controls for non-specific background hybridisation. After loading blots are air dried and baked for 1½h in a vacuum oven.

#### Northern transfers

RNA is run on 1.5% agarose formaldehyde gels as described on page 41. The centre four wells are loaded with 1µg rRNA, 1µg poly A plus, 10µg poly A minus Xbt9S RNA and 1µg rRNA. These tracks are cut out and transferred to nitrocellulose after electrophoresis. On each side of the centre wells 5µg rRNA and 5µg poly A plus Xbt9S RNA are run as visual marker controls. After electrophoresis the outer part of the gel is stained, destained and photographed as described on page 42.

RNA is transferred to nitrocellulose using a Bio Rad "Trans Blot Transfer Apparatus", and model 160/1.6 power supply. The gel holder is placed in a tray of transfer buffer (25mM sodium phosphate

pH 5.8) and the gel placed on top of one Scotch Brite pad which is covered with Whatman 3MM paper. A piece of nitrocellulose, the same size as the gel, is wet with transfer buffer then carefully placed on the gel ensuring there are no air bubbles between gel and filter. Well positions are marked on the nitrocellulose by cutting out one corner level with the wells. The nitrocellulose is covered with Whatman 3MM paper and another Scotch Brite pad and the holder closed and put into the transfer chamber, which contains about 3 litres transfer buffer to completely cover the gel. RNA is transferred by electrophoresis at 27V, about 0.5A for 4h. The filter is then air dried and baked in a vacuum oven at 80°C for 1½h.

#### 2.14.b Preparation of probes for hybridisations

##### Nick-translation of cloned cDNAs

(modified from Rigby et al 1977)

A major  $\alpha$ - and  $\beta$ - tadpole globin cDNA clone and a major adult globin clone have been used as probes. 10 $\mu$ g plasmid DNA (i.e. about 10 $\mu$ l) plus 50 $\mu$ l DNase I mix are incubated at 37°C for 20 minutes, then phenol extracted as described on page 44. The supernatant is precipitated with 150 $\mu$ l isopropanol at -20°C for at least 1h. DNA is recovered by centrifugation, (a small pellet is visible), dried under vacuum, dissolved in 90 $\mu$ l H<sub>2</sub>O and stored frozen. Assuming 90% recovery this gives a DNA concentration of 0.1 $\mu$ g/ $\mu$ l. 2 $\mu$ l of this DNA is run on a 1% agarose mini-gel to check that it is nicked.

0.2 $\mu$ g nicked DNA is labelled as follows: 30 $\mu$ Ci [ $\alpha$  <sup>32</sup>P]dCTP (1mCi/ml, > 400 Ci/mmol) is dried down, dissolved in H<sub>2</sub>O, dried again, taken up in 50 $\mu$ l Pol I mix-C (page 34) and transferred to a Sarstedt tube. 0.5 $\mu$ l is removed to a Whatman GF/C filter to give total cpm in the reaction. 2 units of DNA polymerase I and 2 $\mu$ l



nicked DNA are added to the remaining mix, vortexed and incubated at 37°C for about 40 minutes, until the reaction stops. 0.5µl samples are removed at intervals and CTAB precipitated. Incorporated label is separated from unincorporated on a Sephadex G-75 column run in H<sub>2</sub>O. Approximately 400µl fractions are collected, 1µl from each CTAB precipitated, peak fractions pooled and stored frozen until required.

#### Preparation of cDNA probe

20 µCi [ $\alpha$  <sup>32</sup>P]dCTP (1mCi/ml, > 400 Ci/mmol) is dried down, washed twice with H<sub>2</sub>O and taken up in 20 µl mix-C, 2 µl (i.e. 1µg) Xbt9S poly A plus RNA, 1 µl H<sub>2</sub>O and 1 µl reverse transcriptase. Incubation is at 37°C for 20 minutes to allow priming, then 42°C for 1h, measured from when the water temperature reaches 42°C (Zain et al, 1979). A 1 µl sample is TCA precipitated to check incorporation. RNA is removed from cDNA by adding NaOH to 0.3M and incubating at 37°C for 2h. The solution is neutralised by addition of HCl and one quarter volume 20 x SET. Unincorporated label is removed on a Sephadex G-75 column run in H<sub>2</sub>O.

#### 2.14.c Hybridisations (Thomas, 1980)

Filters are prehybridised for about 20-24h at 42°C in polythene bags in a shaking water bath in the following buffer: 50% formamide, 5 x SET, 50mM sodium phosphate pH 6.5, 0.02% each BSA, Ficoll, polyvinylpyrrolidone, 5 µg/ml poly A and 250 µg/ml sonicated, denatured calf thymus DNA. Formamide is AnalaR grade and is deionised as already described. Calf thymus DNA is heated to 100°C for 10 minutes immediately before being used. About 30ml buffer is used for four filters. Hybridisations are carried out in the same buffer plus probe at 42°C for 24h. Calf thymus DNA and probe are denatured together and added to about 15-20ml buffer. Probes are

kept at 4°C after use and reused for up to two weeks. They are denatured by boiling for 10 minutes prior to each use. Filters are washed four times with 500ml 2 x SET, 0.1% SDS at room temperature for 15 minutes each, then twice with 500ml 0.1 x SET, 0.1% SDS at 50°C for 30 minutes to 1h each. They are blotted dry, covered in Saran wrap and autoradiographed using preflashed Kodak X-omat XS-1 film and an intensifying screen at -70°C. Some exposures are obtained at room temperature without an intensifying screen.

2.15

#### DNA sequencing

DNA was sequenced using the chain-termination method of Sanger et al (1977).

#### 2.15.a Stock solutions

BC1G            5-bromo-4-chloro-3-indoyl-beta-galactoside.  
20mg/ml in dimethylformamide, stored at -20°C.

1PTG            Isopropyl-beta-D-thio-galactopyranoside  
20mg/ml in H<sub>2</sub>O, stored at -20°C.

TE<sup>0.1mM</sup>        10mM tris-HCl pH 7.5, 0.1mM EDTA.

#### T4 DNA ligase dilution buffer

0.05 M KCl

0.01 M tris-HCl pH 7.4

1 mM DTT

0.1 mM EDTA

200µg/ml BSA

50% glycerol

10 x T4 DNA ligase Mix

600 $\mu$ l 1M tris-HCl pH 7.2  
100 $\mu$ l 0.1M EDTA  
100 $\mu$ l 1M MgCl<sub>2</sub>  
100 $\mu$ l 1M DTT  
40 $\mu$ l H<sub>2</sub>O

M13 primer

A single-stranded synthetic DNA 15 bases in length, was first described by Messing et al, 1981. The primer used was obtained from New England Biolabs and has the following sequence:

5' T C C C A G T C A C G A C G T 3'

2.5 $\mu$ g, received dried down in a snap-cap tube is reconstituted by the addition of 1ml of H<sub>2</sub>O and stored at -20°C in small aliquots.

10 x Hin buffer

66 mM tris-HCl pH 7.4  
66 mM MgCl<sub>2</sub>  
0.5 M NaCl  
Stored at -20°C

A solution of 0.1M DTT is stored at -20°C. 10 $\mu$ l 0.1M DTT is added to 90 $\mu$ l 10 x Hin just before the buffer is required.

N<sup>o</sup> mixes

Made fresh every week and stored at -20°C.

	T <sup>o</sup> (μl)	G <sup>o</sup> (μl)	A <sup>o</sup> (μl)	C <sup>o</sup> (μl)
0.5mM TTP	1	20	20	-
0.5mM dGTP	20	1	20	20
0.5mM dATP	20	20	1	20
0.5mM dCTP	-	-	-	1
50mM tris-HCl	5	5	5	5
pH 8; 1mM EDTA				

#### Dideoxynucleotide stocks

0.5mM ddTTP  
0.5mM ddGTP  
0.5mM ddCTP  
1.0mM ddATP

Concentrations of these stock solutions have to be optimised by trial sequencing reactions with each new batch of dideoxynucleotides.

#### DNA polymerase, Klenow fragment

Klenow is obtained from Boehringer at a concentration of 125 units/25μl and stored at -20°C. 1μl enzyme is diluted with 3.7μl dilution buffer (50mM KPO<sub>4</sub> pH 7.0; 0.25mM DTT; 50% glycerol) as required. The enzyme does not keep once it has been diluted.

#### Sanger dyes

0.3% xylene cyanol FF  
0.3% bromophenol blue  
10mM EDTA  
made up in deionised formamide.

10 x TBE buffer

108g/l tris base

9.3g/l EDTA (disodium salt)

55g/l boric acid

Sequencing gels (Sanger and Coulson, 1978)

38% acrylamide, 2% N,N'-methylene bis-acrylamide stock solution kept at 4°C.

Gel contains            21 g urea  
                             10 ml acrylamide stock  
                             5 ml 10 x TBE buffer  
                             1.6ml 1.6% AMPS (80mg in 5ml water)  
Made to 50ml with H<sub>2</sub>O.

Polymerisation is commenced with 50µl TEMED. Gels are poured between glass plates, 20cm x 40cm, separated by "plastikard" spacers 0.35mm thick. Combs are put in position after gels have been poured. Polymerisation starts after about 10 minutes and gels are left for 1h before use.

NB. Plastikard must be extensively washed in detergent before it is first used otherwise it inhibits polymerisation of acrylamide.

M13 host - E.coli K12 JM101

lac, pro, sup E, thi

F' tra D36, pro AB lac I<sup>q</sup>, Z delta M15.

JM101 is maintained on minimal plus glucose plates (see page 35) to prevent loss of the plasmid.

### 2.15.b Subcloning into M13mp8

M13mp8 replicative form (RF), prepared as described by Sanger (1980) was a gift from Dr. D. Finnegan. 2 $\mu$ g RF is digested with Pst I for 2h at 37 $^{\circ}$ C (10 $\mu$ l RF, 2 $\mu$ l 10 x Pst I buffer, 0.2 $\mu$ l i.e. 2 units Pst I, 7.8 $\mu$ l H<sub>2</sub>O). 0.2 $\mu$ g digested RF is run on a 1% agarose mini gel to check that digestion is complete and the remainder heated to 70 $^{\circ}$ C for 10 minutes and diluted to 10ng/ $\mu$ l with TE<sup>0.1mM</sup>.

Cut RF is checked by transformation to ensure that it gives only a few plaques (about ten) per ng DNA. Competent cells are made as follows: 100ml L-broth is inoculated with a single JM101 colony. Cells are grown to OD<sub>650nm</sub> = 0.2, cooled, harvested (Sorvall HS4 rotor, 5 Krpm, 5 minutes, 4 $^{\circ}$ C), resuspended in 50ml cold 0.1M CaCl<sub>2</sub> and left on ice for about 30 minutes. They are then centrifuged, resuspended in 10ml cold 50mM CaCl<sub>2</sub> and either used immediately or left on ice until required (they can be kept for up to 24h on ice, Dagert and Ehrlich, 1979).

1ng and 10ng cut RF and 1ng uncut RF are used to transform 0.2ml competent JM101. DNA is added to the cells and kept on ice for about 40 minutes. Cells are heat shocked at 42 $^{\circ}$ C for 2 minutes, added to 3ml molten BBL top-agar containing 30 $\mu$ l each BC1G and IPTG and 0.2ml exponential JM101 (OD<sub>650nm</sub> = 0.3). They are plated on BBL plates and incubated at 37 $^{\circ}$ C overnight.

Cut RF is religated to check that transformation efficiency is restored to approximately that of uncut RF and that the majority of plaques produced on indicator plates are blue. Ligations are carried out in washed, sterilised Sarstedt tubes at 10 $^{\circ}$ C for about 24h in the following reaction mix: 1 $\mu$ l Pst I cut RF (i.e. 10ng), 1 $\mu$ l T4 ligase mix, 1 $\mu$ l 10mM ATP, 1 $\mu$ l ligase (2, 0.2 or 0.02 units BRL T4 DNA ligase), 6 $\mu$ l sterile H<sub>2</sub>O. 1ng and 5ng DNA from each ligation

are used to transform competent JM101, as described above.

DNA to be cloned into M13 is cut with Pst I and the products analysed on 1.4% agarose mini-gels. Since cDNA was cloned into the Pst I site of pAT 153 by G-C homopolymer tailing cleavage of recombinant plasmids with Pst I gives two fragments, one is linear pAT 153 (3657 nts), the other inserted globin cDNA sequence (about 600 nts). Smaller fragments are preferentially selected in ligation reactions, therefore there is no need to separate the two fragments to subclone the globin sequences into M13.

Pst I cut DNA is ligated to mp8 RF for 22h at 10°C in the following mix: 2µl vector (i.e. 20ng Pst I cut mp8 RF), 1µl Pst I cut DNA (i.e. 40ng), 1µl 10mM ATP, 1µl 10 x T4 ligase mix, 4µl H<sub>2</sub>O and 0.2 units T4 DNA ligase. Initially 2µl of each ligation mix is used to transform JM101 as previously described. The remainder is left at 10°C until results of the first transformations are known. Amounts of DNA then used in further transformations are adjusted so there are about 1-200 plaques per plate.

White plaques are arranged in grids on BBL plates plus 3ml BBL top agar containing 0.2ml exponential JM101 and grown overnight at 37°C. They are screened for globin sequences by a modification of the plaque hybridisation method described by Benton and Davis (1977). Nitrocellulose filters are placed on plates containing the plaques for 1 minute. Plaques are denatured by placing filters, plaque side up, on blotting paper soaked in 0.5M NaOH, 3M NaCl for 2 minutes. Filters are rinsed in a solution containing 0.5M tris-HCl pH 7.4, 3M NaCl, transferred to a beaker of 2 x SET, blotted dry and baked in a vacuum oven at 80°C for 1 h.

Since M13 is a single strand phage and the insert may be in either orientation it is essential to identify which strand a phage

contains. Duplicate filters are made from each plate and hybridised to either an RNA or a cDNA probe made as described on pages 57 and 65 respectively. Phages containing RNA-like inserts hybridise cDNA but not RNA probes, while phages containing cDNA-like inserts hybridise RNA and not cDNA probes.

Hybridisations are as described for colony hybridisation except that 50% formamide is used in place of dimethylformamide and 250 g/ml carrier calf thymus DNA (see page 47) is added to the cDNA hybridisation. Filters are agitated on a rotary shaker overnight at 37°C, rather than kept upright as described earlier.

#### 2.15.c Preparation of single stranded template DNA

Preparation of sequencing templates from M13 mp8 recombinants is adapted from Sanger (1980). An overnight culture from a single JM101 colony is grown up in L-broth. For 20 templates 1ml of overnight culture is diluted into 20ml L-broth and grown at 37°C, with shaking, to OD<sub>650</sub> = 0.3, 1ml aliquots are put into Bijoux tubes, a white plaque toothpicked into each one and shaken at 37°C for 4½h exactly. Single white plaques are first purified by picking into 50µl M13 phage buffer and plating 100µl of a 10<sup>-6</sup> dilution onto BBL indicator plates. A single white plaque is then used for template preparation. Cells are spun down for 5 minutes in a microfuge, and supernatants transferred to clean Sarstedt tubes. 200 µl of 2.5M NaCl, 20% PEG 6000 is added to each supernatant, mixed by inverting and left for thirty minutes. Phage are spun down (5 minutes, microfuge) and PEG supernatants removed using a pasteur pipette pulled to a fine tip. Final traces of supernatant are removed by recentrifuging and using a drawn out capillary attached to a syringe. Pellets are taken up in 100µl TE<sup>0.1mM</sup> and 50µl phenol, vortexed for



10 seconds, left for 10 minutes, vortexed again and centrifuged for 1 minute. The aqueous layers are taken off, avoiding all traces of phenol, and extracted with 100 $\mu$ l chloroform. Supernatants are then ethanol precipitated with 10 $\mu$ l 3M sodium acetate pH 5.5 and 250 $\mu$ l ethanol. Precipitates are collected by centrifugation, washed with 1ml cold ethanol, dried in a vacuum desiccator and taken up in 50 $\mu$ l TE<sup>0.1mM</sup>. Templates are stored at -20°C.

#### 2.15.d One track screening of templates

1  $\mu$ l of a solution containing 4 $\mu$ l M13 primer, 4 $\mu$ l 10 x Hin buffer and 12 $\mu$ l H<sub>2</sub>O is added to each of 16 pulled capillaries. 1 $\mu$ l of template DNA is added to each and the solutions moved to the wide part of the capillaries which are then sealed at both ends in a flame. Primer and template are annealed by placing the capillaries in test tubes of water in a beaker of boiling water for 3 minutes and then allowing them to cool slowly in the tubes of water to room temperature.

Meanwhile 10  $\mu$ Ci [ $\alpha$  <sup>32</sup>P] dCTP (1mCi/ml, > 400 Ci/mmol) is dried down in a short siliconised glass test tube, taken up in 20 $\mu$ l T<sup>o</sup> plus 20  $\mu$ l 0.5mM ddTTP and 2  $\mu$ l aliquots transferred to 16 microcap capillaries in short siliconised glass test tubes.

When template/primer mixes are cool the 2 $\mu$ l aliquots of label are added to each one and mixed in the glass tubes. 0.3 $\mu$ l diluted Klenow is added to each capillary and mixed by taking the solution up into the capillary and expelling it into the tube ten times, avoiding bubbling as this destroys the enzyme. Capillaries are left at room temperature for 15-30 minutes. 1  $\mu$ l 0.5mM dCTP is added to each capillary, mixed and again left for 15-30 minutes. This prevents non-specific termination due to insufficient dCTP in the reaction.

Solutions are expelled into the tubes and kept at 4°C until gels are ready.

Immediately prior to loading gels 4µl of Sanger dyes are added to samples and they are boiled in open tubes for 1 minute. About half each sample is loaded into the slots of an 8% acrylamide gel. Wells are rinsed with 1 x TBE buffer just before loading samples to remove urea which has diffused out of the gel and which would prevent samples from layering under the running buffer. Electrophoresis is at 1.2 - 1.4 KV (so gels feel warm to touch), for approximately 4-6h, until the bromophenol blue dye reaches the bottom of the gel. Gels are fixed for 10 minutes in 10% acetic acid, rinsed in distilled H<sub>2</sub>O (5-10 minutes) and dried onto Whatman 3MM paper. Dupont Cronex 4 film is exposed to gels at room temperature for about 20h.

#### 2.15.e Sequencing

Four templates can be conveniently sequenced at once. 5µCi [ $\alpha$  <sup>32</sup>P] dCTP (1mCi/ml, > 400 Ci/mmol) is dried down into each of three short siliconised glass tubes, 5µCi [ $\alpha$  <sup>32</sup>P] TTP (1mCi/ml, > 400 Ci/mmol) in another and dissolved in 5µl N<sup>0</sup> plus 5µl ddN (see page 68). [ $\alpha$  <sup>32</sup>P] TTP is used for the C reaction. Four 2µl aliquots of each mix are measured into microcap capillaries which are stood in short siliconised glass tubes, labelled with template and T, C, G or A.

8µl template, 1µl 10 x Hin buffer and 1µl M13 primer are taken up in a capillary, mixed on Nescofilm and the capillary sealed. This is repeated for each template. Primer and template are annealed by boiling as described on page 73 and 2µl aliquots are measured into drawn out capillaries.

Label aliquots are added to template/primer capillaries, so that for each template there is a T, C, G and A reaction. 0.3 $\mu$ l diluted Klenow is added to each, mixed and incubated as described in the last section. 0.5mM TTP is used as cold chase in the C reaction.

Samples are treated as described on page 74 and loaded onto two gels, half on each gel. One gel is run for about 6-8h, until the xylene cyanol is almost at the bottom. More dye is then loaded onto the "long" gel and electrophoresis continues until the second xylene cyanol band reaches the bottom (about 10-12h total). Gels are treated and autoradiographed as on page 74.

CHAPTER THREE

ANALYSIS OF GLOBINS AND 9S mRNA ISOLATED FROM  
XENOPUS BOREALIS TADPOLE RBCs

### 3.1 INTRODUCTION

Haemoglobin synthesis in all vertebrates investigated so far (reviewed by Maclean and Jurd, 1972) undergoes a series of sequential changes during development. In Xenopus the switch between tadpole and adult haemoglobin synthesis has been shown to occur during metamorphosis (Just et al, 1977; Hentschel et al, 1979). The haemoglobins of tadpoles and adults have been shown to be separable immunologically (Just et al, 1980) and the globin subunits are thought to be completely distinct sets, although there may be some overlap between minor components (Hentschel et al, 1979). This system is therefore particularly suitable for examining regulation of a family of genes during development and a considerable amount is now known about adult X. laevis globin genes (see Chapter 1 for references).

Mature Xenopus erythrocytes synthesise very little, if any, RNA (Maclean et al, 1973). Therefore in order to study globin synthesis in adult Xenopus it is necessary to inject animals with phenylhydrazine (Thomas and Maclean, 1975) which kills circulating RBCs and results in a wave of erythropoiesis. Globin mRNA can then be isolated from the immature erythrocytes which are released into the blood. It was assumed by analogy that in order to isolate globin mRNA from Xenopus tadpoles they would also have to be made anaemic by injection of phenylhydrazine (Perlman et al, 1977). However, as will be described in this chapter, it has proved unnecessary to treat tadpoles with phenylhydrazine, since they are naturally anaemic and their blood contains many immature cells active in RNA and protein synthesis. RNA may be isolated from animals raised in entirely

normal physiological conditions eliminating any question of whether globin species seen are those normally present or are produced as a response to phenylhydrazine, as is known to be the case for chicken and sheep (Salser et al, 1979; Richards and Wells, 1980; Wood et al, 1979) and also possibly for adult X. laevis (Maclean and Jurd, 1972; Hentschel et al, 1979).

Widmer et al (1981) have recently prepared cDNA clones for two  $\alpha$ - and two  $\beta$ -globins from X. laevis tadpole globin mRNA. However, they raised animals in 0.4% thiourea to prevent onset of morphological metamorphosis (Just et al, 1977) and then induced anaemia by transferring the tadpoles to 3mg/l phenylhydrazine for 30 minutes. This work is therefore open to criticism on the grounds that the physiology of the tadpoles used has been severely upset and the sequences isolated may not normally be expressed at this stage of development.

Most previous work on Xenopus globins has used the species X. laevis but, as shown in this chapter, the same transition in haemoglobin synthesis also occurs during metamorphosis of X. borealis. X. borealis tadpoles are at least twice the size of X. laevis and must contain more blood and hence more globin mRNA. Use of X. borealis to study tadpole globin mRNA has reduced the problem of obtaining sufficient RNA.

Since X. borealis has not been used before in studies of globin gene expression it was necessary to obtain information about the proteins before attempting to isolate globin mRNA and make cDNA clones. Globins of X. borealis are shown in this chapter to be

separable on denaturing polyacrylamide gels, as is the case for X. laevis globins (Hentschel et al, 1979; Hosbach et al, 1982).

To isolate cDNA clones coding for tadpole globin sequences mRNA which does not contain a high proportion of adult globin-coding sequences was used. Transition from tadpole to adult globin synthesis in X. borealis occurs during metamorphosis in a similar manner to that described for X. laevis (Just et al, 1977; Hentschel et al, 1979). Labelling isolated RBCs in culture with radioactive amino acids and analysing the products on polyacrylamide gels by fluorography provides a more sensitive assay for onset of adult globin synthesis and, as shown in this chapter, adult globin synthesis can be detected by stage 57 (Nieuwkoop and Faber, 1956). Amounts of adult globin mRNA in blood of animals after this stage of development can be expected to increase as the transition to entirely adult globin production proceeds and RNA was isolated therefore at an early stage (stages 57-59) in the transition.

## RESULTS

### 3.2 Tadpole blood contains many immature RBCs

Three points demonstrating that tadpoles are naturally anaemic are outlined below.

(1) Examination of blood smears by light microscopy shows some tadpole RBCs to be similar morphologically to immature RBCs of anaemic adults (see Thomas and Maclean, 1975, for photographs of

anaemic adult RBCs). X. borealis tadpole blood contains many immature erythrocytes and dividing cells, as has also been shown for Rana blood (Benbassat, 1970).

(2) Thomas and Maclean (1975) showed that differentiating erythroid cells in anaemic adult X. laevis are active in protein synthesis.

X. borealis tadpole RBCs also synthesis protein and [<sup>3</sup>H]-leucine labelled RBC products, separated on polyacrylamide gels and fluorographed, consist of a few major proteins which are assumed to be globins (see page 81).

(3) Tadpole RBCs contain RNA while non-anaemic adult erythrocytes do not. Approximately 0.4µg poly A plus 9S mRNA can be extracted from one stage 57 X. borealis tadpole. Details of RNA preparation from tadpole RBCs are given in section 3.5.

### 3.3 Tadpole globins

X. borealis tadpole and adult globins can be completely separated on triton-acid-urea polyacrylamide gels (figure 1). Tadpole globin polypeptides, defined as the major polypeptides extracted from tadpole RBCs, consist of three major and five minor species (figure 1). In subsequent experiments both major and minor species are shown to hybridise with sequenced globin cDNA sequences (Chapter 5) thus proving their authenticity as globins.

Each track in figure 2 shows globins isolated from one animal but unlike adults (figure 3) no globin polymorphism has been detected. Figure 3 shows X. borealis adult globin polymorphism,



Figure 1. X. borealis tadpole globins separated on a triton-acid-urea polyacrylamide gel

Details of sample and gel preparation are given in materials and methods. Each track shows globins isolated from RBCs of an individual animal. RBCs were isolated from MS222 anaesthetised adult frogs by heart puncture and globins isolated as from tadpole RBCs. Approximately 200 $\mu$ l samples were obtained from each tadpole and about 4ml from each adult. Samples (as indicated below) were mixed with 10  $\mu$ l triton-acid-urea sample buffer, heated at 100 $^{\circ}$ C for 3 minutes, to denature haemoglobin into its constituent globin subunits, and loaded into slots of a 10% triton-acid-urea polyacrylamide gel. Electrophoresis was at 7mA constant current for 16h at 20 $^{\circ}$ C, then the gel was fixed, stained with Coomassie blue R250, destained and photographed.

Tadpole globins consist of three major proteins, labelled 1, 2, 3 track 3 and five minor proteins, 1a, 1b, 3a, 3b, 3c track 3. Adult globins were run as markers to demonstrate the complete separation between tadpole and adult RBC proteins on these gels.

Track 1 0.5 $\mu$ l adult X. laevis globins

2 0.5 $\mu$ l adult X. borealis globins

3 0.5 $\mu$ l stage 57 tadpole X. borealis globins

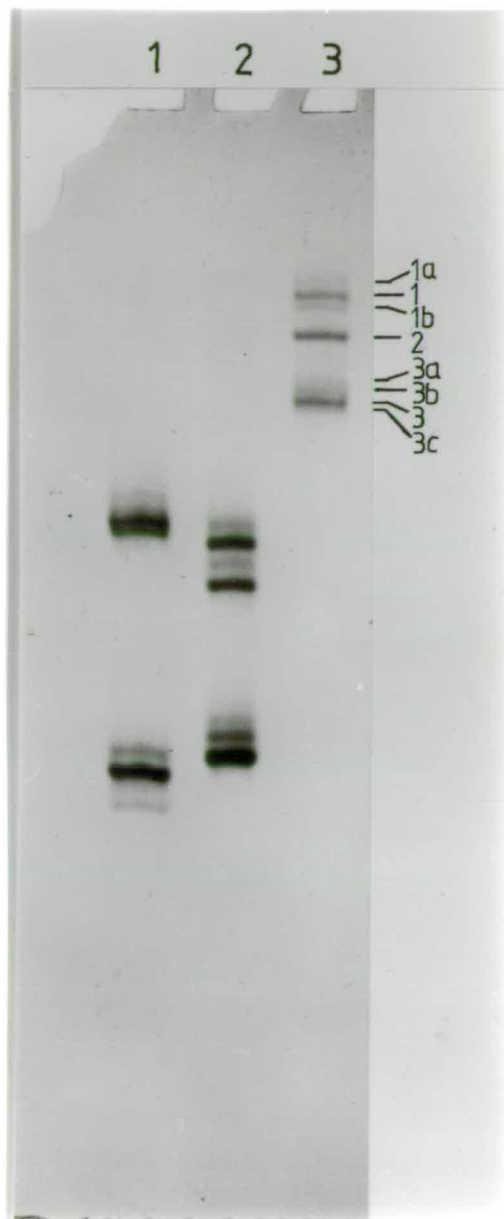


Figure 2 Triton-acid-urea polyacrylamide gel of globins from  
X. borealis tadpole RBCs

For experimental details see legend to figure 1 and materials and methods. Approximately 1/100th (i.e. 2 $\mu$ l) of the total globin isolated from an individual tadpole was loaded into each slot. The poorer resolution of globin components compared to figure 1 is due to inferior quality of acrylamide used in this gel.

Tracks 1-4 stage 57 tadpole globins

5-8 stage 58 tadpole globins

9 0.5 $\mu$ l adult X. borealis globins

10 0.5 $\mu$ l adult hybrid borealis/laevis globins

11 0.5 $\mu$ l adult X. laevis globins.

Inset 1a, 1, 1b, 2, 3 indicate indicate individual globins which can be detected by staining as shown in figure 1.

1 2 3 4 5 6 7 8 9 10 11

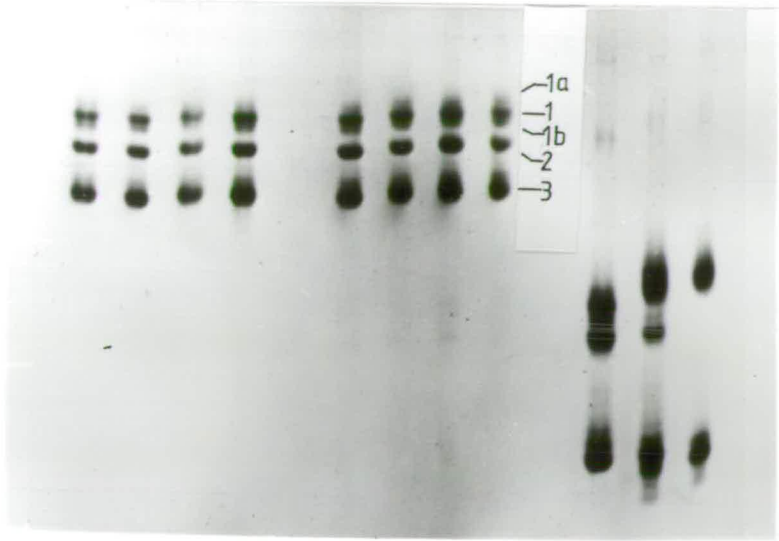


Figure 3. Triton-acid-urea polyacrylamide gel of globins  
from young X. borealis frogs

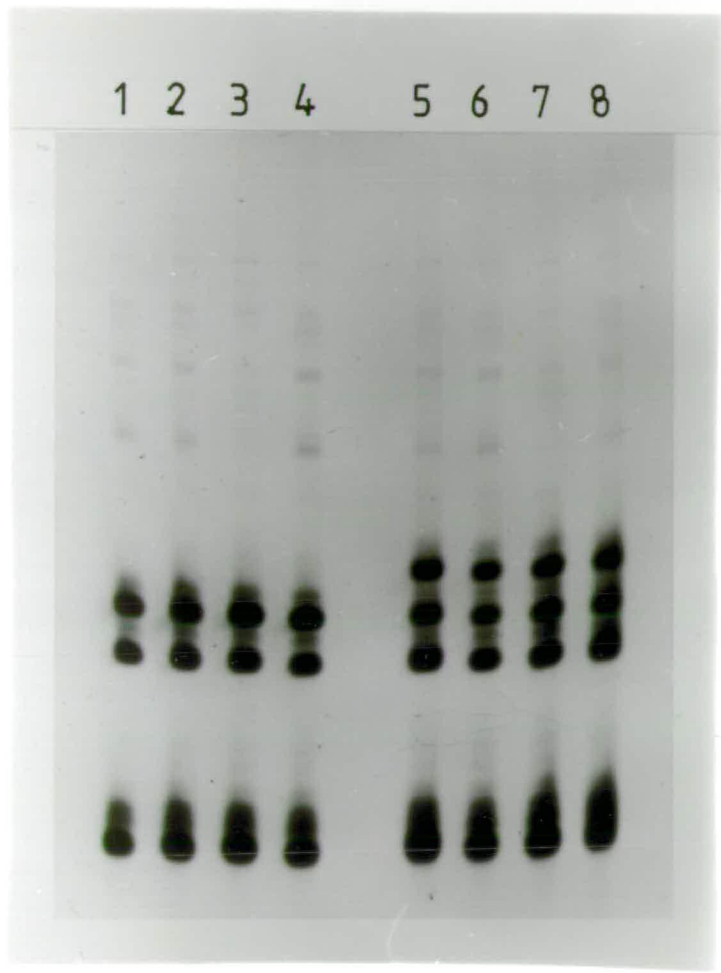
Globins were isolated from X. borealis frogs 10 to 12 weeks after metamorphosis (determined from the time they reached stage 66, Nieuwkoop and Faber, 1956), as described in the legend to figure 1. Each track shows approximately 1/1000th (i.e. 1  $\mu$ l) of the total globin sample prepared from an individual. Electrophoresis was at 7mA constant current at 20°C for 16h and the gel was fixed, stained with Coomassie blue R250 and photographed as described in materials and methods. The upper half of the gel shows traces of tadpole globins still detectable by staining 12 weeks after metamorphosis.

Tracks 1-4 globins from frogs 10 weeks after metamorphosis.

Tracks 5-8 globins from frogs 12 weeks after metamorphosis.

The slowest migrating band, tracks 5-8, is polymorphic. Its presence or absence is not dependent on the frogs' age and it is purely fortuitous that the four animals in each group show the same pattern of globin bands. The second slowest migrating band has also been observed to be polymorphic, and is not present in some animals.

1 2 3 4 5 6 7 8



i.e. the absence (tracks 1-4) or presence (tracks 5-8) of the slowest migrating major component. It has been reported that genetic polymorphism of adult X. laevis is infrequent (Hentschel et al, 1979). The second slowest running major component (figure 3) is also polymorphic. It is purely fortuitous that animals chosen for this gel show only two of the three possible combinations.

Figure 4 shows tadpole and adult globins run on an SDS polyacrylamide gel (Laemmli, 1970). Tadpole globins (tracks 4 and 6) consist of two components at least one of which is smaller than adult globin (tracks 5 and 7). Molecular weight markers are shown alongside the globin tracks in figure 4 but it is impossible to estimate molecular weights of the globin species due to anomalous migration of the markers.

#### 3.4 Globin synthesis in isolated RBCs

Blood cells from tadpoles at various stages of development were incubated with either [<sup>3</sup>H]-leucine or [<sup>35</sup>S]-methionine and labelled proteins analysed on triton-acid-urea polyacrylamide gels by fluorography. [<sup>35</sup>S]-methionine was used as it allows a more sensitive and therefore earlier detection of adult globin synthesis. However there is evidence which suggests tadpole globins may not contain methionine (figure 5), so cells were also labelled with [<sup>3</sup>H]-leucine to permit a direct comparison of relative amounts of adult and tadpole globins being synthesised, assuming equal leucine contents.

Figure 4. 5-20% polyacrylamide gradient (Laemmli) gel of  
X. borealis tadpole and adult globins

Tadpole and adult globins were isolated from individual animals, as described in materials and methods, from RBCs obtained by heart puncture. A 5-20% polyacrylamide gradient gel containing the discontinuous buffer system described by Laemmli (1970) was loaded with approximately 1/10th the amount of globin used for the triton-acid-urea gels in figure 2. The gel was run at 8mA constant current for 15h at 20°C then fixed and stained with Coomassie blue R250 as described in materials and methods.

- Tracks 1 and 2 Molecular weight markers cytochrome C (12000) and  $\beta$ -lactoglobulin (18000).
- Track 3 Molecular weight markers bovine serum albumin (64000), penicillinase (28000) and lysosyme (13900).
- Track 4 0.2 $\mu$ l stage 57 X. borealis tadpole globins.
- Track 5 0.05 $\mu$ l X. borealis adult globins.
- Track 6 0.4 $\mu$ l stage 57 X. borealis tadpole globins.
- Track 7 0.05 $\mu$ l X. borealis adult globins.
- Tracks 8 and 9 Molecular weight marker myoglobin (17000).



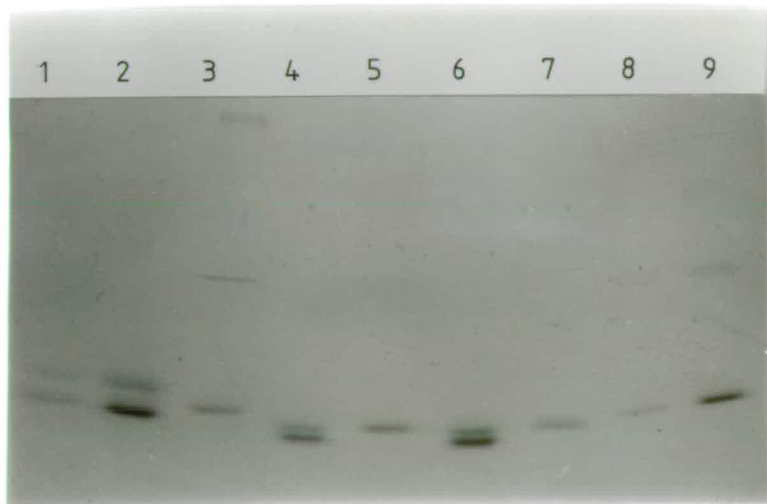


Figure 5. Fluorograph of [<sup>3</sup>H]-leucine and [<sup>35</sup>S]-methionine  
labelled proteins from X. borealis tadpole RBCs

RBCs were isolated from individual tadpoles at various stages of development as indicated below. The RBCs were extracted into a solution of Barth X and heparin containing 0.7 $\mu$ M NH<sub>4</sub>FeSO<sub>4</sub> on ice in order to maintain their synthetic activity. Cells from each animal were divided into two and labelled overnight half by culturing in a medium containing [<sup>35</sup>S]-methionine and half in a medium containing [<sup>3</sup>H]-leucine. Products were recovered by lysing the cells in the presence of Mg<sup>2+</sup> and spinning out nuclei. [<sup>35</sup>S]-methionine and [<sup>3</sup>H]-leucine labelled products were separated on the same 10% triton-acid-urea polyacrylamide gel (7mA constant current, 16h at 20<sup>o</sup>C) and analysed by fluorography (Laskey and Mills, 1975). Details of experimental procedures are given in materials and methods. Since [<sup>35</sup>S]-labelled tracks expose more quickly than [<sup>3</sup>H]-labelled tracks the gel was exposed for 1 day (right hand tracks 1-6) and 10 days (left hand tracks 1-6) and the fluorographs cut and realigned for the photograph.

RBCs were isolated from animals at the following stages of development:

Track 1, 57/58; 2, 57; 3, 57; 4, 53/54; 5, 55; 6, 54.

Data for this figure were obtained by P.J. Ford.

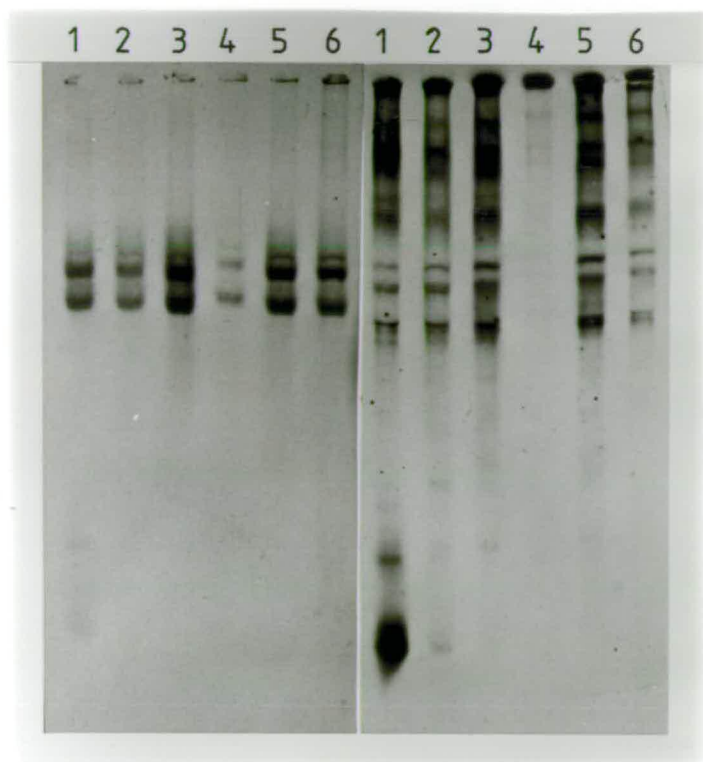


Figure 6 Fluorographs of  $^3\text{H}$ -leucine and  $^{35}\text{S}$ -methionine  
labelled proteins from *X. borealis* tadpole RBCs

RBCs were isolated from stage 57 and 58 *X. borealis* tadpoles and labelled in culture with [ $^{35}\text{S}$ ]-methionine or [ $^3\text{H}$ ]-leucine as described in the legend to figure 5 and materials and methods. [ $^3\text{H}$ ]-leucine and [ $^{35}\text{S}$ ]-methionine labelled products were run on separate triton-acid-urea polyacrylamide gels, (7mA constant current, 16h at 20°C), since their fluorographs require different exposure times. Each track in figure 6a shows products from cells of the same animal as the corresponding track in figure 6b. The fluorographs were overexposed (figure 6a 16 days, figure 6b 3 days) to allow detection of adult globin synthesis.

Figure 6a [ $^3\text{H}$ ]-leucine labelled proteins. 2 $\mu\text{l}$  of labelled sample, 1.5  $\times$  10<sup>6</sup> cells was loaded per track and the fluorograph was exposed for 16 days.

Figure 6b [ $^{35}\text{S}$ ]-methionine labelled proteins. 3 $\mu\text{l}$  of labelled sample, 1.5  $\times$  10<sup>6</sup> cells was loaded per track, fluorograph exposed for 3 days.

Figures 6a and 6b

Tracks 1-4 globins synthesised in RBCs isolated from stage 57 tadpoles.

Tracks 5-8 globins synthesised in RBCs isolated from stage 58 tadpoles.

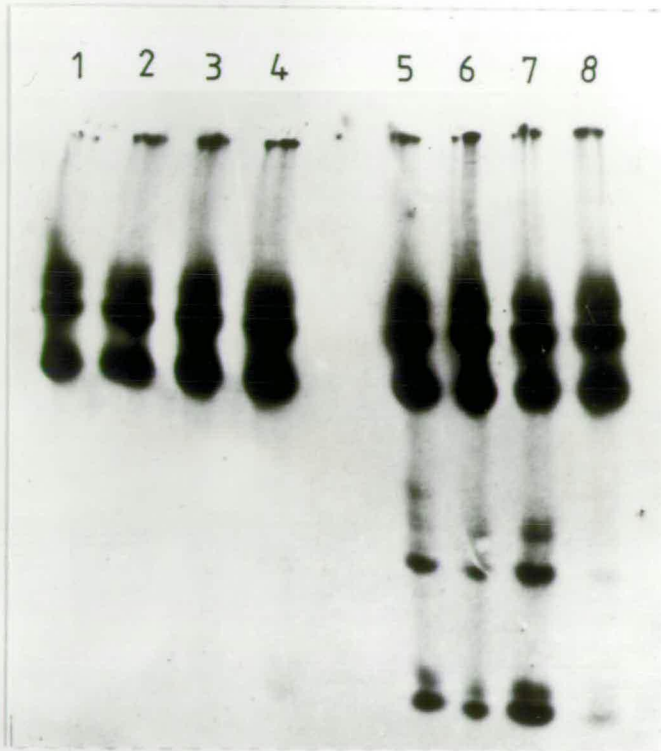
Figure 6 shows fluorographs of proteins from stage 57 and 58 animals. When labelled at stage 57 with [ $^3\text{H}$ ]-leucine (figure 6a) no adult globin synthesis is apparent, however labelling with [ $^{35}\text{S}$ ]-methionine clearly shows that it is taking place. Cells from stage 55 and 56 tadpoles incorporated very little label, presumably because of the small number of cells cultured and, in the case of [ $^{35}\text{S}$ ]-methionine, the low level of adult globin synthesis. None of the products were analysed on gels.

These results have been confirmed in a separate experiment: figure 5 tracks 5 and 6 shows labelled proteins from stages 55 and 54 respectively, no adult globins can be seen. They are present, however, in proteins from stage 57/58 (track 1) and traces can be seen in track 3 (stage 57).

Figure 6b shows that the relative amounts of adult globin synthesised by individual stage 57 tadpoles varies greatly. This suggests that there is a period, presumably beginning just before stage 57, during which adult globin synthesis is switched on. The precise point at which individual animals start to synthesise adult globin apparently varies within this period.

Both tadpole and adult globins are synthesised for some time after stage 57. Hentschel et al (1979) reported the presence of both tadpole and adult globins in blood of young X. laevis frogs three to four weeks after metamorphosis. Cells from X. borealis animals up to stage 61 have been incubated with [ $^3\text{H}$ ]-leucine and both tadpole and adult globins are labelled (figure 7). Figure 7a suggests that synthesis of the slowest migrating tadpole globin species

a.



b.

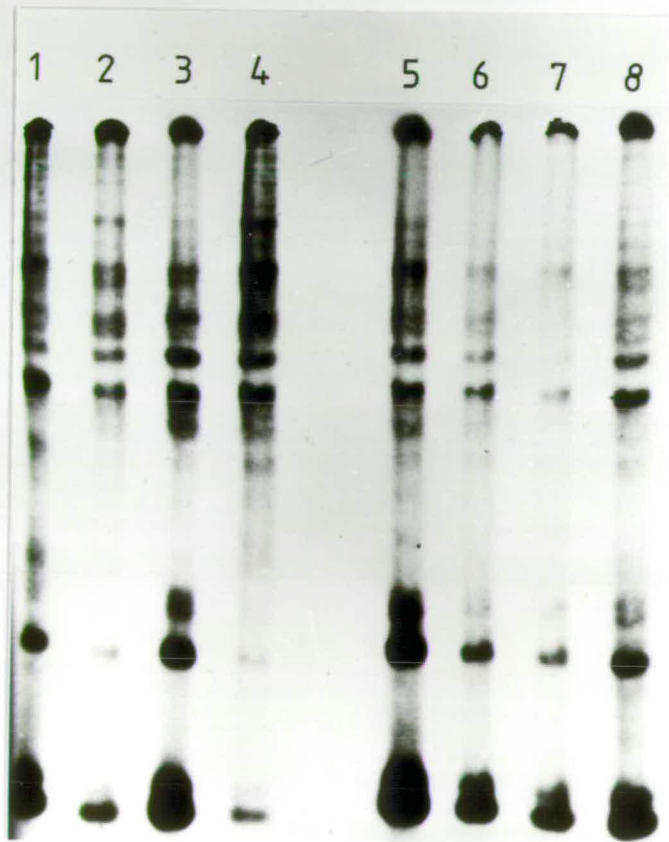


Figure 7 Fluorographs and stained gels of globins isolated from stage 59 to 61 X. borealis RBCs cultured with [<sup>3</sup>H]-leucine

RBCs were isolated from individual stage 59 to 61 tadpoles and labelled in culture with [<sup>3</sup>H]-leucine. Globin samples were run on triton-acid-urea 10% polyacrylamide gels (7mA constant current, 16h at 20°C) and the gels stained with Coomassie blue R250, destained and photographed prior to being prepared for fluorography. Experimental details are given in materials and methods.

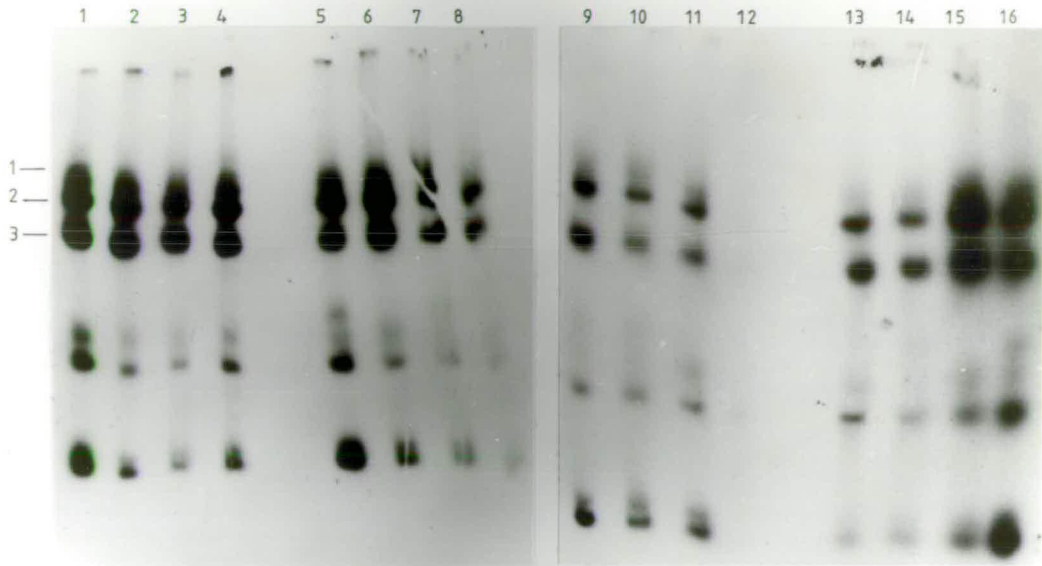
RBCs were isolated from animals at the following stages of development:

Tracks 1-4	stage 59
Tracks 5-12	stage 60
Tracks 13-16	stage 61

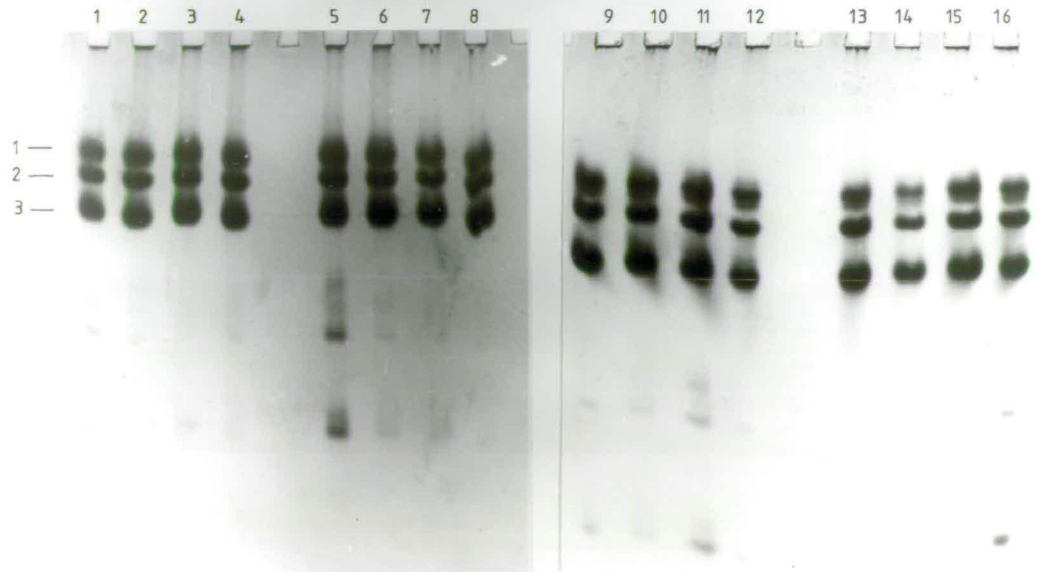
Figure 7a - fluorographs

Figure 7b - stained gels

a.



b.





(labelled 1, figure 7b) is the first to be switched off during the transition from tadpole to adult globin synthesis. Whereas the protein gels in figure 7b clearly demonstrate the presence of component 1 in stage 61 tadpole RBCs, figure 7a shows that it is apparently no longer being synthesised. In contrast to this figures 6 and 7 show that the different adult globin species are apparently co-ordinately regulated and synthesis of them all begins simultaneously.

Figure 3 shows that X. borealis tadpole globins are still present in RBCs of animals twelve weeks after stage 66. However, it is not possible to detect synthesis of either tadpole or adult globins by labelling cells of young frogs. Incorporation of [<sup>3</sup>H]-leucine or [<sup>35</sup>S]-methionine into cells from animals one week after stage 66 is too low for analysis by fluorography. The young frogs' RBCs now consist mainly of mature erythrocytes which synthesise little, if any, globin. Tadpole globin synthesis in the blood has ceased by this stage and Perlman et al (1977) could detect no tadpole globin mRNA sequences in RNA prepared from anaemic adult blood cells. Complete disappearance of tadpole globins from the blood will depend on the life span of mature erythrocytes, reported to be about 100 days by comparison with other amphibia (Maclean et al, 1973; Cline and Waldman, 1962; Forman and Just, 1976). The possibility that RBCs containing tadpole globins are recruited to the blood having matured elsewhere at this stage has not been investigated.

### 3.5 RNA preparation

Perlman et al (1977) showed that X. laevis adult and tadpole globin mRNAs do not cross hybridise and that tadpole blood cells contain a low level of adult globin message. The stage from which to isolate tadpole globin mRNA is a compromise between using larger animals in order to maximise the amount of blood extracted and avoiding stages at which substantial amounts of adult mRNA are produced. The previous section showed that tadpole globin synthesis still predominates at stages 57-59 and these stages were used for total RNA preparation.

Tadpole globin mRNA was initially purified on a sucrose gradient, details of which are shown in figure 8. Globin 9S mRNA sediments as a small peak between 18S and 4+5S RNA. [<sup>3</sup>H]-poly U hybridisation (dotted line, figure 8) shows most poly A plus RNA from blood cells sediments in this region. Inclusion of a ribonucleoside-vanadyl complex RNase inhibitor in the cell lysis mix improves RNA quality, 9S poly A plus RNA being of a much more uniform size. The 9S fraction is further purified by passage over an oligo dT cellulose column, followed by sedimentation on a small sucrose gradient (figures 9 and 10).

Table 2 gives details of the quantities of RNA isolated from the blood of stage 57-59 tadpoles. The values shown are minimum estimates because RNA losses during extraction have not been accounted for. Actual values in different experiments vary greatly, due partly to differences in extraction efficiency and partly to variation between batches of animals. The total amount of blood RNA

Figure 8 Initial purification of tadpole 9S RNA on a 7-30%  
sucrose in NETS gradient

Approximately 100 OD260nm units of total RNA isolated from stage 57-59 X. borealis tadpole RBCs by phenol extraction was fractionated on a 58ml 7-30% sucrose in NETS gradient (see materials and methods for details). Centrifugation was in an MSE SW65 rotor at 23.5 Krpm for 22h at 25°C. A thin stainless steel tube was then inserted into the bottom of the centrifuge tube, the gradient was pumped through a UV scanner and 2.5ml fractions were collected. Peak poly A containing fractions, as determined by [<sup>3</sup>H]-poly U hybridisation (Bishop et al, 1974, see legend to Table 2) were pooled and RNA recovered by ethanol precipitation prior to fractionation on oligo dT cellulose. ( — ) OD260nm trace of gradient. 28S, 18S, 9S and 4+5S peaks can be seen from left to right respectively. The 9S peak is shown on a scale expanded 10 times compared to the other peaks. (-o---o-) cpm [<sup>3</sup>H]-poly U hybridising to each fraction. Peak poly U hybridisation coincides with the 9S peak as determined by optical density.

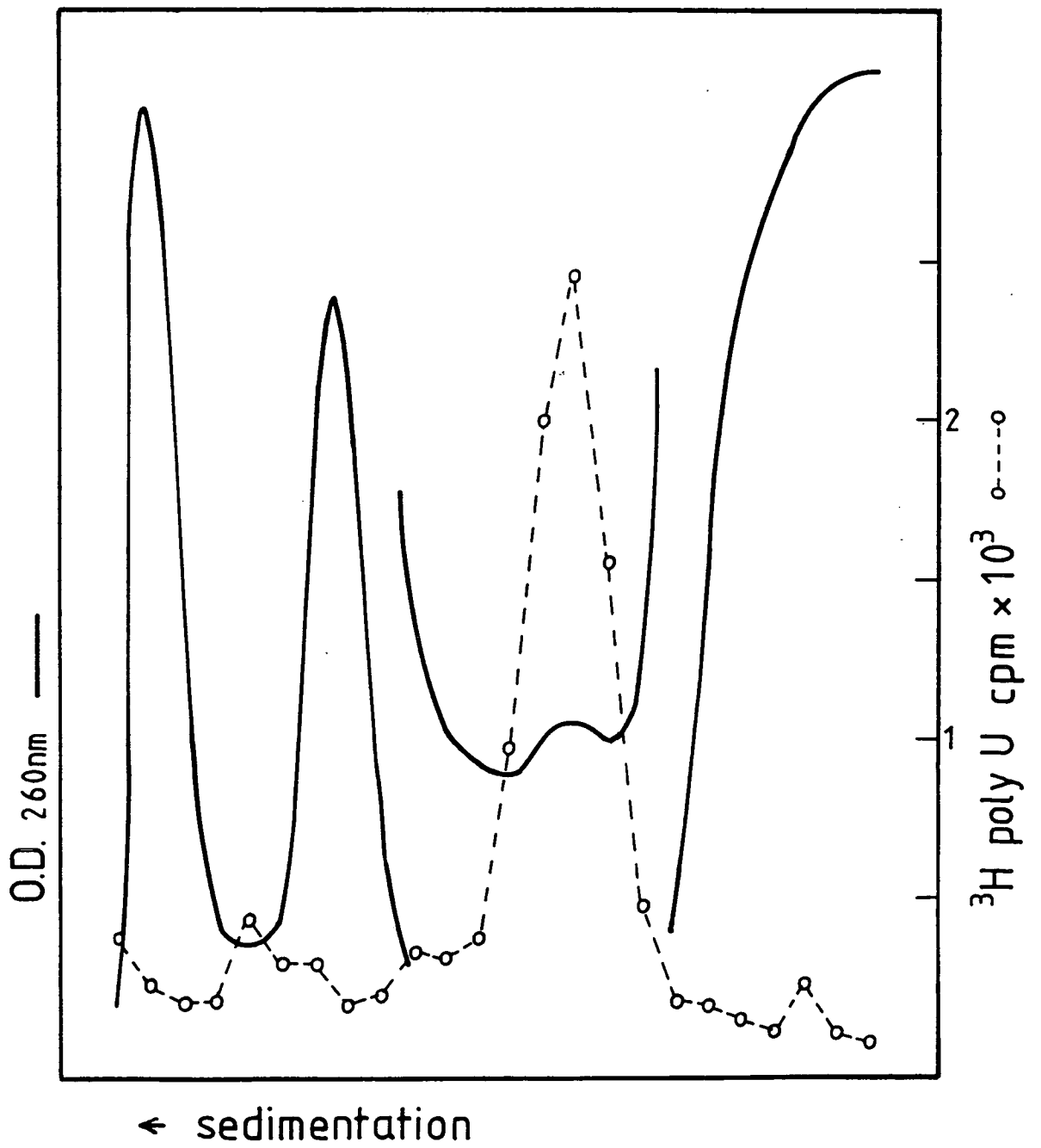


Figure 9. Final purification of tadpole 9S poly A plus RNA on a  
7-30% sucrose in NETS gradient

Poly A plus RNA from oligo dT column fractionation was further purified on a 4.2ml 7-30% sucrose in NETS gradient. Approximately 50 to 100 $\mu$ g RNA was loaded on to one gradient, calculated from its OD<sub>260nm</sub>. About 50  $\mu$ g each of 28S, 18S and 4+5S RNA were run in parallel on a separate gradient as markers. Sedimentation is from right to left as shown by the arrows. 9S RNA sediments between the 18S and 4+5S peaks. Centrifugation was in an MSE 6 x 5ml Titanium swinging bucket rotor at 50Krpm for 4h at 25<sup>o</sup>C.

(a) 28S, 18S and 4+5S RNA markers.

(b) Poly A plus RNA from oligo dT column fractionation. The small peak and shoulder sedimenting below the main 9S peak are contaminating 28+18S RNA. A small amount of 4+5S RNA, sedimenting more slowly than the 9S RNA, is also present.

O.D. 260 nm

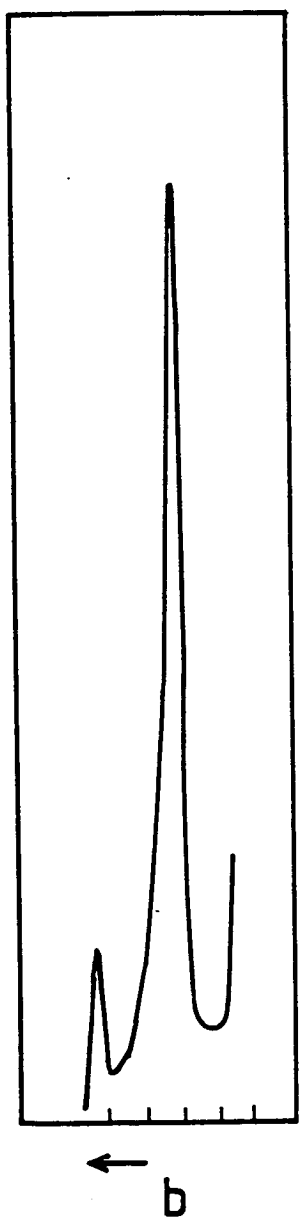
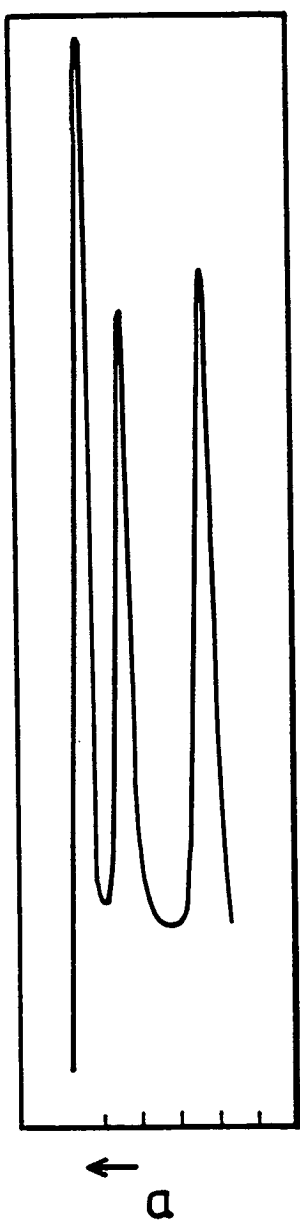


Figure 10. Final purification of tadpole 9S poly A minus RNA on a 7-30% sucrose in NETS gradient

9S poly A minus RNA from oligo dT column fractionation was further purified on a 4.2ml 7-30% sucrose in NETS gradient, as described for poly A plus RNA in the legend to figure 9.

- (a) 4+5S marker RNA.
- (b) 9S poly A minus RNA. ↓ indicates the 9S peak. There is considerable contaminating 28S, 18S and 4+5S RNA and also a shoulder of unknown composition sedimenting slightly slower than the 9S RNA.
- (c) 28S+18S marker RNA.

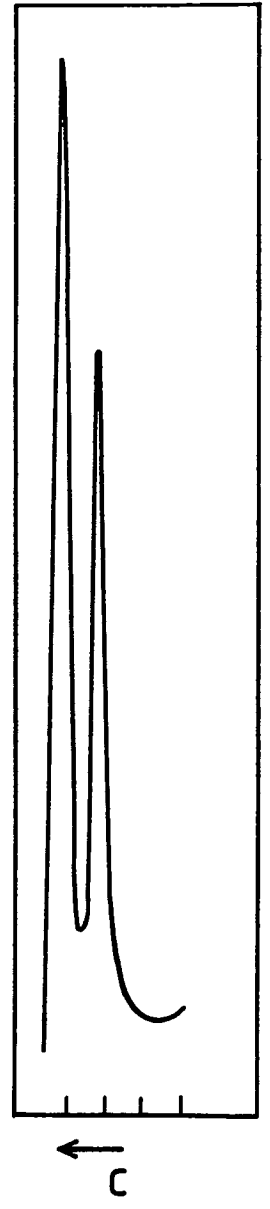
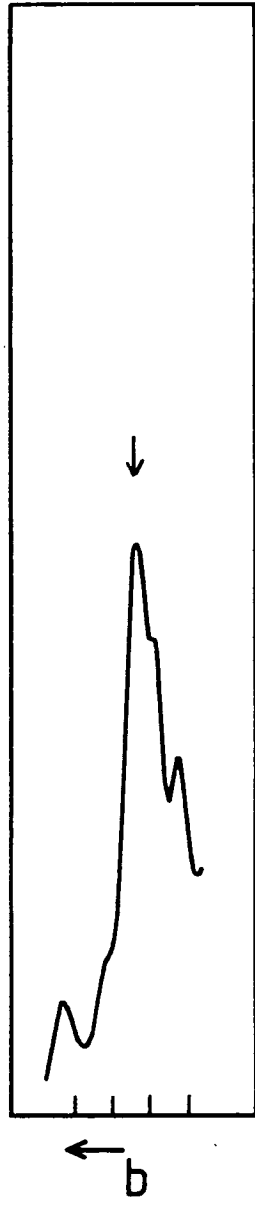
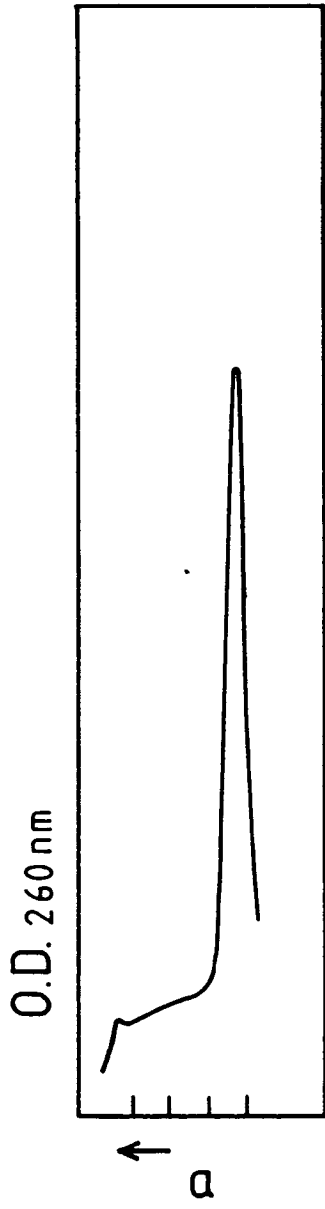




TABLE 2. Maximum amounts of RNA extracted from stage 57-59

X. borealis tadpole blood

Purification step	<sup>1</sup> μg total RNA/tadpole	<sup>2</sup> μg total A+ RNA/tadpole	<sup>3</sup> μg 9S"A+" RNA/tadpole	<sup>5</sup> μg 9S"A-" RNA/tadpole
60 ml SUCROSE GRADIENT	80	3.7	0.64	-
OLIGO dT COLUMN	-	-	0.4	1.7
4.2 ml SUCROSE GRADIENT	-	-	0.4	-

Cytoplasmic RNA was extracted from pooled blood cells of stage 57-59 X. borealis tadpoles by phenol extraction (Penman, 1966). (1) The total amount of RNA extracted was determined, after the sample had been washed three times to remove phenol and chloroform, by spectrophotometry, at 260nm. The OD 260:280nm ratio was measured to check that it was about 2 and that most of the proteins had been removed. (2) Amounts of poly A plus RNA recovered from 60ml sucrose gradients were calculated by [<sup>3</sup>H]poly U hybridisation (Bishop et al, 1974) in the following way:

$$\mu\text{g poly A plus RNA} = \frac{\text{cpm in fraction}}{\text{specific activity poly U}} \times 10,$$

assuming the poly A tail represents 10% of the total globin mRNA (Laird and McCarthy, 1968). (3) Peak fractions from 60ml gradients which hybridised to poly U were pooled, ethanol precipitated, and fractionated on oligo dT cellulose to remove any contaminating poly A minus RNA. The amount of poly A plus RNA was again determined

by [<sup>3</sup>H]poly U hybridisation. (4) 9S poly A plus RNA from oligo dT column fractionation was further purified on a small sucrose in NETS gradient and the peak fractions pooled according to a trace from a UV scanner (figure 9). The amount of 9S RNA was calculated from the area under the peak (figure 9) and the OD<sub>260nm</sub> of the sample: A small proportion of contaminating RNA is still present as shown by the fact that the peak does not start or finish at the base line (figure 9). The amount of this RNA was calculated, by determining the area between the base line and the start of the peak, and subtracted from the total OD<sub>260nm</sub>. (5) The amount of 9S poly A minus RNA from oligo dT column fractionation was calculated from its OD<sub>260nm</sub>. A large proportion of this RNA is contaminating ribosomal and transfer RNA (figure 10) and the recovery from the small sucrose gradient was not quantified.

Values presented here are the maximum amounts of RNA which were extracted from stage 57-59 tadpoles. Amounts extracted in different experiments varied greatly as discussed in the text, and these figures merely provide an idea of the approximate amounts of RNA which it is possible to isolate from these animals. Amounts of globin mRNA in 9S RNA preparations were not determined.

extracted per tadpole is approximately 80 $\mu$ g, of this 3.7 $\mu$ g is poly A plus RNA determined by poly U hybridisation to total RNA. 9S poly A plus RNA can be further purified from this fraction yielding approximately 0.4 $\mu$ g per tadpole, determined by optical density.

Figure 11 shows X. borealis tadpole blood cell 9S (Xbt9S) poly A plus RNA and X. borealis adult anaemic blood cell 9S poly A plus RNA run on a polyacrylamide gel. Xbt9S poly A plus RNA is heterogeneous, one component being approximately the same size as adult X. laevis globin RNA, one larger and one smaller.

Figure 12a shows Xbt9S poly A plus and poly A minus RNA run on a formaldehyde-agarose gel. Again the poly A plus RNA is seen to be heterogeneous. The same amounts of poly A plus and poly A minus RNA (5 $\mu$ g) were loaded into each well and it appears that the poly A plus RNA fraction contains a greater proportion of the two slowest migrating components, whereas the fastest running RNA is present in equal amounts in both samples. The smallest RNA component apparently does not prime cDNA synthesis: figure 12b shows an RNA gel blot of total Xbt9S cDNA hybridised to Xbt9S poly A plus and poly A minus RNA. Only the two larger species hybridise, there is no detectable hybridisation in the region of the small RNA.

Sizes of RNA components in each band were estimated from the photograph shown in figure 12a and two other formaldehyde agarose gels (figure 25b and not shown) to be  $848 \pm 66$ ,  $680 \pm 56$  and  $443 \pm 21$  nucleotides.

Figure 11. Polyacrylamide gel of *X. laevis* adult and *X. borealis*  
tadpole 9S RNA preparations

*X. borealis* tadpole 9S poly A plus RNA was prepared from stage 57-59 tadpoles as described in materials and methods and figures 8 to 10. *X. laevis* adult 9S poly A plus RNA was isolated from RBCs of animals two weeks after injection of phenylhydrazine, and purified in the same way as tadpole 9S RNA.

2  $\mu$ g of each RNA preparation were separated on a 4% polyacrylamide gel containing 8M urea, run in tris borate buffer (108 g/l tris; 55 g/l boric acid; 9.5 g/l Na<sub>2</sub>EDTA), overnight at 10mA constant current at 25°C. The gel was stained with 10 $\mu$ g/ml ethidium bromide and photographed.

Track 1 *X. laevis* adult 9S poly A plus RNA

Track 2 *X. borealis* tadpole 9S poly A plus RNA

↓ indicates direction of migration.

1 2

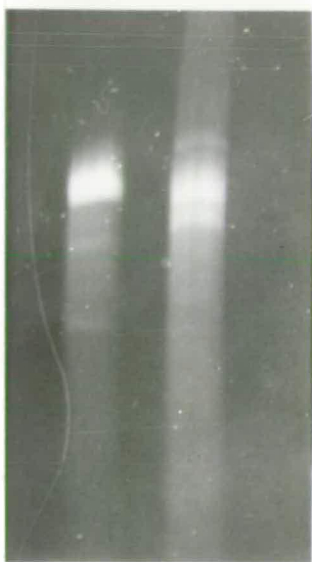


Figure 12. 1.5% formaldehyde-agarose gel of X. borealis tadpole  
9S poly A plus and poly A minus RNA

(12a) RNA samples (as indicated below) were dried down, taken up in 18 $\mu$ l sample buffer and heated to 55 $^{\circ}$ C for 15 minutes. 2 $\mu$ l 10 x dyes were added and samples layered under gel buffer into the wells of a 1.5% agarose gel containing 2.2M formaldehyde. Details of gel and sample preparation are given in materials and methods. Electrophoresis was at 70V constant voltage for about 14h, until the dye had travelled two thirds the length of the gel. The gel was stained with 5 $\mu$ g/ml ethidium bromide in 0.1M ammonium acetate for 30 minutes, destained for 2h in 0.1M ammonium acetate and photographed.

Track 1      10 $\mu$ g 4+5S RNA  
          2      5 $\mu$ g 18+28S RNA  
          3      5 $\mu$ g X. borealis tadpole 9S poly A plus RNA  
          4      5 $\mu$ g X. borealis tadpole 9S poly A minus RNA

(12b) Xbt9S poly A plus and poly A minus RNA separated on a formaldehyde-agarose gel as shown in figure 12a was transferred to nitrocellulose and hybridised to cDNA prepared from Xbt9S poly A plus RNA. Experimental details are given in materials and methods.

Track 1      1 $\mu$ g poly A plus RNA  
          2      10 $\mu$ g poly A minus RNA

Arrows indicate positions of the two major bands (848  $\pm$  66 and 680  $\pm$  56 nucleotides) seen by ethidium bromide staining of gels.

a.



b.

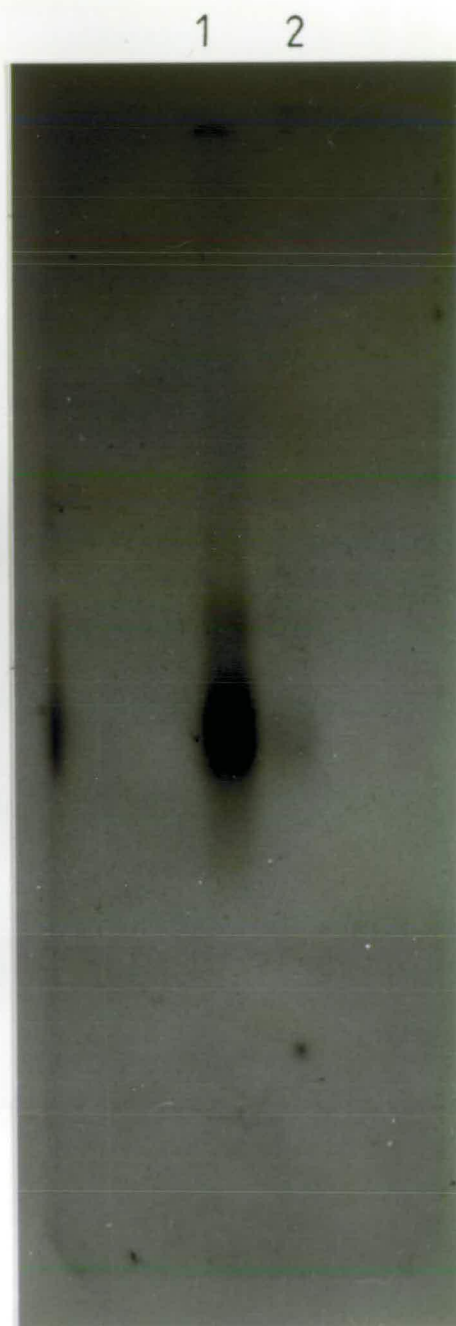


Figure 13. Fluorograph of translation products from *X. borealis* tadpole blood cell 9S poly A<sup>+</sup> and poly A<sup>-</sup> RNA preparations

1 µg of 9S poly A<sup>+</sup> and poly A<sup>-</sup> RNA from two separate preparations and an additional 9S poly A<sup>-</sup> RNA sample (all isolated as described in materials and methods and figures 8-10) were each translated in 25 µl micrococcal nuclease treated rabbit reticulocyte lysate containing 25 µCi [<sup>3</sup>H]-leucine and 25 µCi [<sup>3</sup>H]-histidine. Incorporation data is given in Table 3. Products were analysed by fluorography on 10% triton-acid-urea polyacrylamide gels. Electrophoresis was at 7mA constant current for 16h at 22°C. Gels were fixed, stained with Coomassie blue R250 and photographed prior to being treated for fluorography as described by Laskey and Mills (1975). The fluorograph was exposed for 3 days at -70°C. Further experimental details are given in materials and methods.

Track	Sample	cpm/µl lysate	µl loaded onto gel
1	<u><i>X. borealis</i></u> tadpole 9S poly A <sup>+</sup> RNA (1)	51 694	1
2	<u><i>X. borealis</i></u> tadpole 9S poly A <sup>+</sup> RNA (2)	97 540	0.5
3	<u><i>X. borealis</i></u> tadpole 9S poly A <sup>-</sup> RNA (1)	8 739	6
4	<u><i>X. borealis</i></u> tadpole 9S poly A <sup>-</sup> RNA (2)	4 156	12
5	<u><i>X. borealis</i></u> tadpole 9S poly A <sup>-</sup> RNA (3)	9 000	6



Approximately 50 000 cpm were loaded into each slot of the gel to ensure that all tracks exposed in the same time. This resulted in a large amount of rabbit globin being present in poly A<sup>-</sup> tracks which has affected migration of samples within the gel. Bands in the A<sup>-</sup> tracks do not appear to comigrate exactly with those in A<sup>+</sup> tracks, but the overall pattern of proteins is characteristic of globins isolated from X. borealis tadpole RBCs.

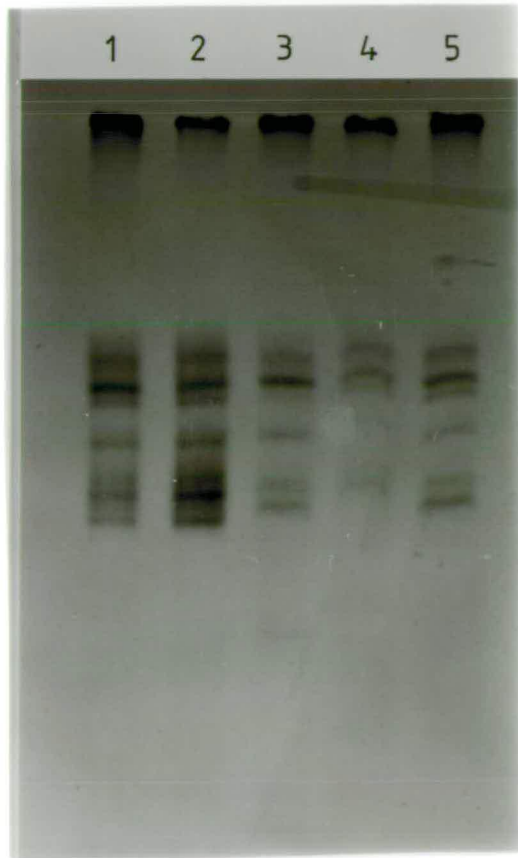


TABLE 3. Incorporation of [<sup>3</sup>H]-leucine and [<sup>3</sup>H]-histidine stimulated by Xbt9S poly A plus and poly A minus RNA in rabbit reticulocyte lysate.

Time point	cpm per $\mu$ l lysate x 10 <sup>3</sup>					
	-mRNA <sup>1</sup>	A <sub>1</sub> <sup>+</sup>	A <sub>2</sub> <sup>+</sup>	A <sub>1</sub> <sup>-</sup>	A <sub>2</sub> <sup>-</sup>	A <sub>3</sub> <sup>-</sup>
10	0.4	24	14	0.6	1.9	3.0
20	0.7	47	28	1.2	2.6	4.4
40	1.3	64	37	3.0	4.8	8.0
65		84	51	3.0	6.5	8.6
90		97	50	3.0	7.2	8.8
initial rate of incorporation cpm/min	30	2350	1400	70	140	150
relative rates of incorporation	1	78	47	2	5	5

1 $\mu$ g of each of three X. borealis tadpole blood cell 9S RNA preparations (labelled with subscripts 1, 2, 3 above) was translated in 25 $\mu$ l micrococcal-nuclease treated rabbit reticulocyte lysate containing 25 $\mu$ Ci each [<sup>3</sup>H]-histidine and [<sup>3</sup>H]-leucine. Incubation was at 30<sup>o</sup>C for 90 minutes and samples were then stored at -20<sup>o</sup>C. 1 $\mu$ l aliquots were TCA precipitated at intervals during incubation and incorporation of label determined. This data is presented graphically in figure 14, from which initial rates of incorporation were determined.

(1) Incorporation due to translation of endogenous rabbit reticulocyte mRNA.

Figure 14. Incorporation of [<sup>3</sup>H]-leucine and [<sup>3</sup>H]-histidine  
stimulated by Xbt9S poly A plus and poly A minus RNA  
in rabbit reticulocyte lysate

See legend to Table 3 and materials and methods for experimental details.

Key to symbols: incorporation due to translation of

△—△ endogenous mRNA;

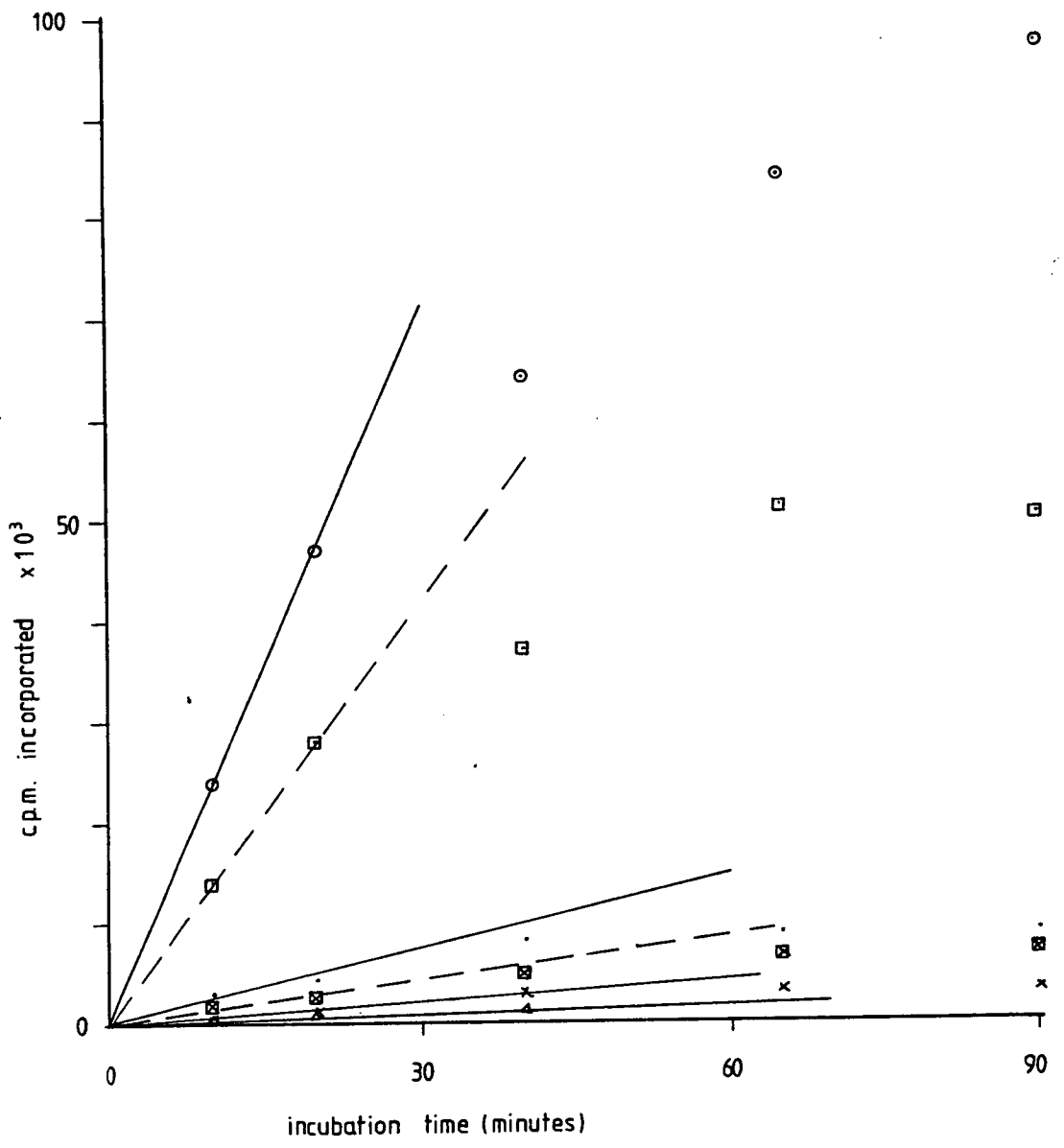
⊖—⊖ preparation 1, Xbt9S poly A plus RNA;

⊖—⊖ preparation 2, Xbt9S poly A plus RNA;

×—× preparation 1, Xbt9S poly A minus RNA;

⊖—⊖ preparation 2, Xbt9S poly A minus RNA;

●—● preparation 3, Xbt9S poly A minus RNA.



### 3.6 In vitro translation of X. borealis tadpole blood cell 9S mRNA

Purified stage 57-59 Xbt9S poly A plus and poly A minus RNA was translated in micrococcal nuclease treated rabbit reticulocyte lysate. Products from two independent mRNA preparations were shown to comigrate with tadpole globins on triton-acid-urea polyacrylamide gels (figure 13). No detectable adult globin synthesis was stimulated by these mRNAs (figure 13). Poly A minus RNA from oligo dT column fractionation also stimulates tadpole globin synthesis (figure 13) but to a much lesser extent. Details of incorporation are given in Table 3 and figure 14 from which it can be seen that initial rates of incorporation stimulated per  $\mu\text{g}$  poly A plus RNA are 10 to 40 times greater than those of the corresponding poly A minus RNA preparations. However, figure 13 shows that all globin mRNA species appear to be present in poly A minus RNA, although at a reduced level compared to poly A plus RNA.

### 3.7 DISCUSSION

Xenopus borealis tadpoles are naturally anaemic and it is therefore not necessary to treat them with phenylhydrazine to enable globin message preparation. Other developing systems have also been shown to be naturally anaemic e.g. foetal mice (Cole and Tarbutt, 1973) and Rana tadpoles (Benbassat, 1970).

Since there is no need to induce anaemia in tadpoles, the globin mRNAs studied are the products of an entirely normal physiological system. RNA and globin species which can be extracted from tadpole RBCs are therefore those which occur in normal

development. In adult animals, where anaemia has to be induced in order to isolate RNA, it is always uncertain whether all RNA and globin species detected in RBCs during recovery from anaemia are normally present in developing adult erythrocytes. The appearance of a novel globin species in immature erythrocytes of adult X. laevis recovering from anaemia has been reported (Maclean and Jurd, 1972; Hentschel et al, 1979). However no such novel globin has been detected in RBCs from anaemic adult X. borealis.

It has been reported that some of the minor tadpole globins in X. laevis may be common to both tadpole and adult RBCs (Hentschel et al, 1979). However, inclusion of triton-X 100 in gels improves resolution and figures 1,2 and 3 show that there is a clear and complete separation between X. borealis tadpole and adult globins using this system. Hosbach et al (1982) have recently also shown this for X. laevis globins using a similar gel system.

Figure 4 suggests that tadpole globins are heterogeneous in size, at least one component being smaller than the adult globin band. Data from the RNA gel shown in figure 11 is consistent with the idea of a heterogeneous population of globins. However, there is a prominent band migrating more slowly than adult mRNA and no tadpole globins larger than the adult proteins have been detected. This could most easily be explained by variations in the lengths of 3' and 5' untranslated regions of the mRNAs. Evidence that this band contains globin-coding RNA is presented in Chapter 5.

The labelling experiments with isolated RBCs suggest that X. borealis tadpole globins may not contain methionine. The proteins

have not been sequenced so the possibility that they contain only one or two methionine residues cannot be definitely excluded. However, incubating RBCs with [<sup>35</sup>S]-methionine clearly does not label the tadpole globins satisfactorily and it is therefore necessary to use other labelled amino acids when comparing relative amounts of adult and tadpole globins synthesised. Lack of methionine in globins of other anuran larvae has also been observed. Watt and Riggs (1975) determined the amino acid composition of the  $\alpha$  and  $\beta$  chains from four major haemoglobin components of Rana catesbeiana tadpoles and found that they contain no methionine (nor cysteine). Since Rana tadpole globins did not contain cysteine no attempt was made to label X. borealis tadpole RBCs with [<sup>35</sup>S]-cysteine.

The work of Hentschel et al (1979) suggested that the transition from tadpole to adult globin synthesis occurs during metamorphosis, adult globin species being first detectable in RBCs of young X. laevis frogs three to four weeks after metamorphosis. However, this work did not involve labelling the globins. By labelling tadpole blood cells in culture with [<sup>35</sup>S]-methionine it is possible to detect synthesis of adult globin species in RBCs from stage 57 X. borealis tadpoles. Labelling with [<sup>3</sup>H]-leucine shows that tadpole globin synthesis still predominates at this stage.

These results suggest that the switch from tadpole to adult globin synthesis starts sometime just before stage 57 and adult globin synthesis then gradually supercedes tadpole globin synthesis. The precise point at which adult globin synthesis begins varies between individuals. Some tadpole globin species can still be detected, by staining gels, twelve weeks after metamorphosis, but



it is no longer possible to label either tadpole or adult globins at this stage by incubation of RBCs. Thus cessation of tadpole globin synthesis within the circulation system occurs between stage 66 and twelve weeks thereafter. The possibility that cells containing tadpole globins are still being released into the circulation, having matured elsewhere, has not been investigated nor has any attempt been made to reactivate tadpole globin synthesis during this period by, for instance, inducing anaemia.

RNA preparation from blood cells of stage 57-59 tadpoles yields about 0.4  $\mu$ g 9S poly A plus RNA per animal, approximately four times as much as can be obtained from X. laevis tadpoles (Widmer et al, 1981). Use of X. borealis therefore significantly reduces the problem of obtaining reasonable quantities of 9S RNA.

The proportion of 9S RNA which codes for globins has not been estimated, but experiments involving RNA gel blot hybridisations (see Chapter 5) show that the two major 9S poly A plus RNA bands seen on formaldehyde-agarose gels (figure 12a) hybridise to  $\alpha$ - or  $\beta$ -globin cDNA probes. Poly A minus RNA gives fainter and more diffuse bands on formaldehyde-agarose gels (figure 12a) and appears to contain less of the two major globin-coding bands.

In vitro translation by rabbit reticulocyte lysate shows that both poly A plus and poly A minus 9S RNA fractions contain sequences coding for all globins normally isolated from tadpole RBCs. Poly A minus RNA is 10 to 40 times less efficient at incorporating labelled amino acids into proteins in the reticulocyte lysate (Table 3, figure 14), suggesting it contains 10 to 40 times less coding RNA, than

poly A plus RNA. This is in agreement with data from dot blot and RNA gel blot hybridisations (chapter 5).

Since poly A minus RNA is translated, (albeit at a low level), to give normal globin products it is possible that a small proportion of globin message may contain very short, or non-existent poly A regions. The limit of resolution of oligo dT column fractionation under the conditions used is about 10 to 20 A residues (Cabada et al, 1977) and poly A minus RNA was passed over the column until no more RNA bound, to ensure that as much poly A plus RNA as possible had been removed. The actual number of A residues on the "poly A minus" RNA was not determined and so it is not known if any molecules completely lack a poly A segment. The possibility that poly A is removed during RNA extraction was not investigated.

CHAPTER FOUR

PREPARATION OF XENOPUS BOREALIS TADPOLE GLOBIN cDNA CLONES

#### 4.1 INTRODUCTION

In order to investigate the regulation of globin gene expression during metamorphosis of Xenopus it is necessary to construct specific probes which can be used to identify individual globin genes. Adult globin cDNA clones have been prepared from X. laevis blood cell 9S RNA (Hentschel et al, 1979; Kay et al, 1980; Widmer et al, 1981) and more recently cDNA clones coding for two  $\alpha$ - and two  $\beta$ -globins from X. laevis tadpoles have also been isolated (Widmer et al, 1981).

It was thought that X. borealis would be a useful organism in which to study globin gene expression since, as mentioned in the previous chapter, the tadpoles are large and it is therefore relatively easy to isolate globin mRNA from them. cDNA clones coding for adult X. borealis globins have been isolated and partially characterised (P.J. Ford unpublished). Isolation of 9S RNA from X. borealis tadpole RBCs was described in the previous chapter and it was shown to contain sequences coding for all globins normally present in tadpole RBCs, as judged by electrophoresis on triton-acid-urea polyacrylamide gels. The preparation of cDNA clones from this RNA is described below.

X. laevis adult globin mRNA has been shown to be about  $670 \pm 50$  nucleotides long (Kay et al, 1980). X. borealis tadpole 9S poly A plus RNA is heterogeneous (see previous section), one component being larger and one approximately the same size as adult X. laevis globin mRNA. Selection of cDNA clones with inserts over 600 nucleotides long should therefore identify those containing almost full length

copies of globin mRNA.

Partial restriction analysis of clones containing large inserts subdivides them into groups each probably containing the same cDNA sequence (Widmer et al, 1981). The largest clones from each group can then be selected for further analysis.

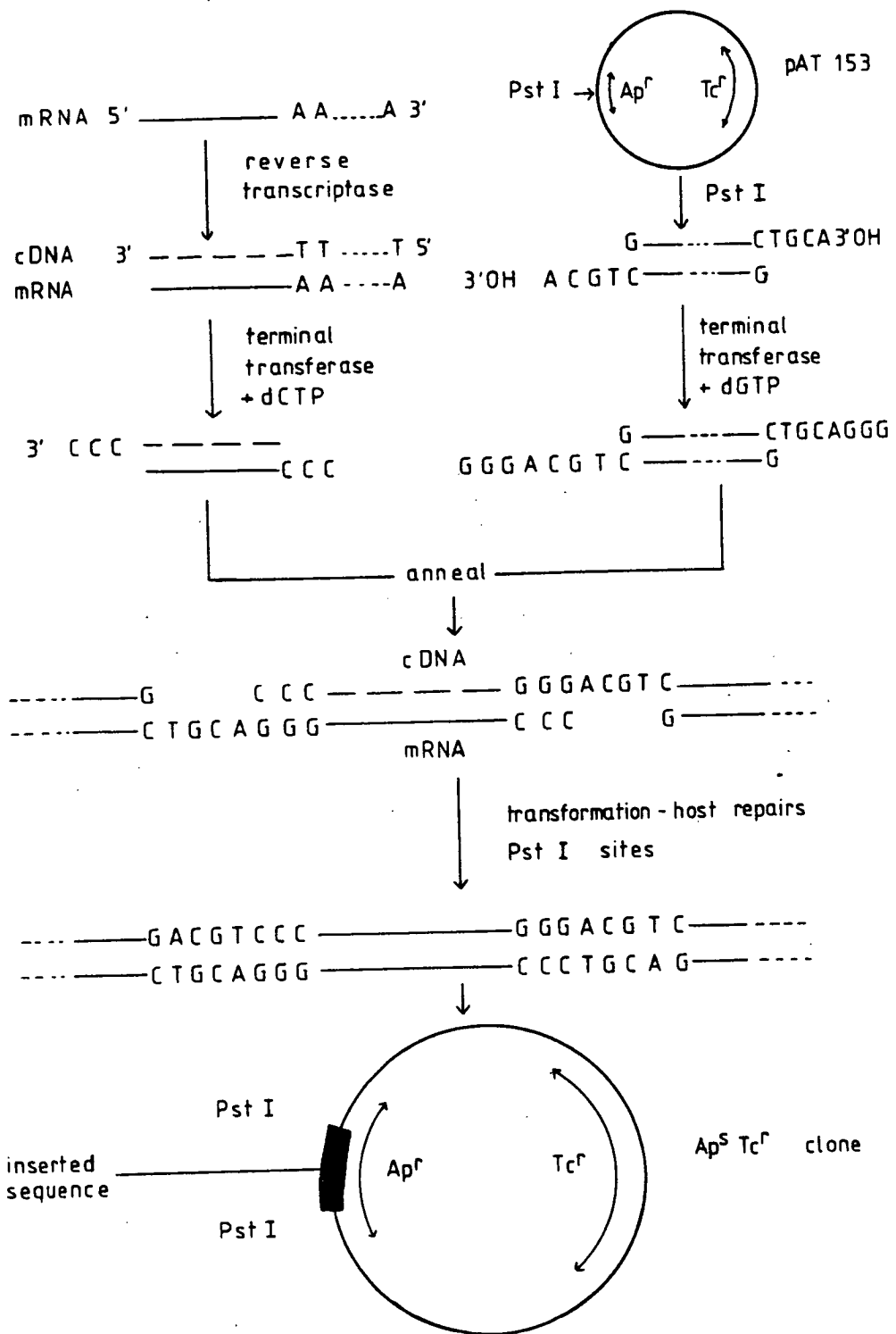
## RESULTS

### 4.2 Preparation and identification of clones containing sequences complementary to X. borealis tadpole blood cell 9S poly A plus RNA

Figure 15 outlines the strategy used to clone X. borealis tadpole globin cDNA sequences. 3.1 $\mu$ g cDNA was made from 10 $\mu$ g Xbt9S poly A plus RNA, (see Appendix I for calculation), taken up in 30 $\mu$ l H<sub>2</sub>O and a dC tail added to 5 $\mu$ l as described by Roychoudhury et al (1976). On average 56 dC residues were added to each 3' terminus, calculated from the radioactivity incorporated and assuming all cDNA molecules were full length copies of globin mRNA (see Appendix II for calculation). This estimate is likely to be too high since many cDNA molecules will be incomplete copies and small molecules contribute a disproportionate number of 3' termini to the tailing reaction. Tailed cDNA/RNA molecules were size-fractionated on a 7-30% sucrose in NTE gradient, small fractions collected and the amount of nucleic acid in each calculated (Appendix III). Fractions were pooled so each sample contained sufficient cDNA for one annealing reaction (approximately 0.025  $\mu$ g in the two samples containing the largest molecules, and 0.05 $\mu$ g in the next two) and ethanol precipitated with

Figure 15. Strategy for cloning cDNA/RNA hybrids

X. borealis tadpole 9S poly A plus RNA from RBCs is the template for cDNA synthesis. cDNA/RNA hybrids are tailed with dCTP by terminal transferase (Roychoudhury et al, 1976), size fractionated on a sucrose gradient and annealed to Pst I cut, T-tailed pAT 153. After transformation into E. coli HB101, host repair mechanisms regenerate a Pst I recognition sequence on either side of the insert (↓). Insertion at the Pst I site of pAT 153 inactivates the ampicillin resistance gene and transformants are Ap<sup>S</sup>Tc<sup>R</sup>.



3  $\mu\text{g}$  E. coli tRNA plus two volumes ethanol. Fractions containing smaller molecules were pooled in approximately 0.05  $\mu\text{g}$  aliquots and stored frozen at  $-20^{\circ}\text{C}$ , in case they were later required.

pAT 153 was cut with Pst I to give a reduction from  $10^6$  transformants per  $\mu\text{g}$  uncut pAT 153 to  $10^4$  per  $\mu\text{g}$  cut pAT 153. Approximately 27 dG residues were added to each 3' terminus (see Appendix II) of the Pst I cut pAT 153 in an analogous tailing reaction to that described for cDNA/RNA molecules.

E. coli K12 strain HB101 was transformed with dG tailed pAT 153 annealed to dC tailed cDNA from the four sucrose gradient fractions containing the largest molecules.  $1.8 \times 10^4$  and  $2.9 \times 10^4$  transformants per  $\mu\text{g}$  cDNA were obtained from the first and second largest fractions respectively. The two smaller fractions gave  $4.1 \times 10^3$  and  $1.1 \times 10^4$  transformants per  $\mu\text{g}$  cDNA. About 1000 transformants from the two fractions containing the largest cDNA molecules were kept for further study. Approximately 75% of these were found to be ampicillin sensitive when screened on 50  $\mu\text{g}/\text{ml}$  Ap plates.

337  $\text{Ap}^{\text{S}}\text{Tc}^{\text{R}}$  clones were screened by colony hybridisation with kinase labelled Xbt9S poly A plus RNA (specific activity approximately  $3.0 \times 10^6$  cpm/ $\mu\text{g}$ ) and X. borealis anaemic adult blood cell 9S poly A plus RNA (specific activity approximately  $1.3 \times 10^7$  cpm/ $\mu\text{g}$ ) separately to identify those containing tadpole globin cDNA sequences. Table 4 and figure 16a,c,e,f show the results of these experiments. Of the 337 clones screened about 60% (i.e. 203) showed some degree of hybridisation to Xbt9S poly A plus RNA (figure 16a,e



TABLE 4. Colony Hybridisation Data showing number of clones hybridising to Xbt9S and anaemic adult X. borealis blood cell 9S poly A plus RNA

Hybridisation signal	No. of clones	% clones
<sup>1</sup>		
Tadpole		
xxx	71	21
xx	67	20
x	65	19
0	134	40
<sup>2</sup>		
Adult + Tadpole	6	1.8
Adult only	30	8.9

Experimental details are given in materials and methods.

Colonies were grown and their DNA fixed on nitrocellulose filters. Hybridisation in 50% dimethylformamide, 5 x SET plus kinase labelled RNA was overnight at 37°C, then filters were washed and autoradiographed.

1. Filters were hybridised with 2.5µg kinase labelled Xbt9S poly A plus RNA (specific activity  $3.0 \times 10^6$  cpm/µg). Hybridisation signals were scored on a zero to three scale (figure 16a): 0 = no detectable signal above background; xxx = maximum signal detected. Scoring was very subjective and therefore does not give an absolute estimate of the amount of hybridisation.

2. The same clones, on new filters, were hybridised to 2.5µg kinase labelled X. borealis adult 9S mRNA (specific activity  $1.3 \times 10^7$  cpm/µg) as described above (figure 16c). Those clones which hybridised were then sorted into classes according to whether or not they also hybridised to tadpole mRNA (figure 16e).

Figure 16. Screening Xbt9S clones by colony hybridisation

Clones containing tadpole blood cell 9S sequences were selected from  $\text{Ap}^{\text{S}}\text{Tc}^{\text{r}}$  transformants by colony hybridisation with kinase labelled Xbt9S and Xba9S poly A plus RNA (specific activities  $3.0 \times 10^6$  and  $1.3 \times 10^7$  cpm/ $\mu\text{g}$  respectively). The same clones were also screened with Xbt9S poly A minus RNA (specific activity  $7.2 \times 10^6$  cpm/ $\mu\text{g}$ ) and as a negative control with X. tropicalis blood cell rRNA (specific activity  $2.6 \times 10^6$  cpm/ $\mu\text{g}$ ). Details of experimental procedures are given in materials and methods. Autoradiographs were exposed at  $-70^{\circ}\text{C}$  using Kodak intensifying screens for the following times:

- (a) Xbt9S poly A plus RNA hybridisation, 20h exposure.
- (b) Xbt9S poly A minus RNA hybridisation, 24h exposure.
- (c) Xba9S poly A plus RNA hybridisation, 2 day exposure.
- (d) rRNA hybridisation, 20h exposure.

(a)-(d) Duplicate filters are arranged in pairs from left to right and numbered 1 to 6.

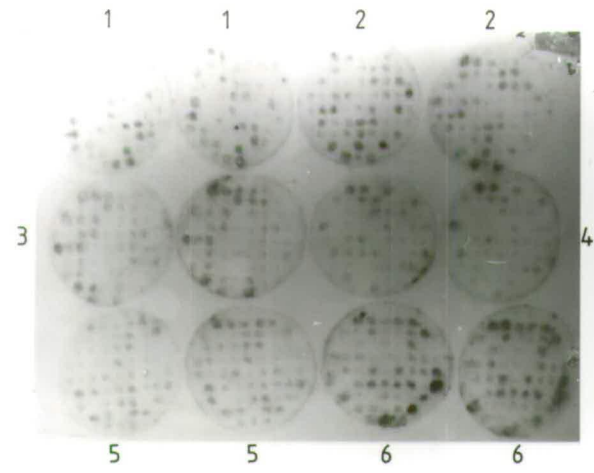
(e)and(f) Data from the autoradiographs shown in 16a and 16c is summarised diagrammatically. Filters are shown once only and numbered as in 16a-d.

(e) Clones hybridising to Xbt9S poly A plus RNA are denoted by open circles (○, compare 16a), those hybridising to Xba9S poly A plus RNA by closed circles (●, compare figure 16c) and those hybridising to both tadpole and adult mRNA by half filled circles (◐). + indicates

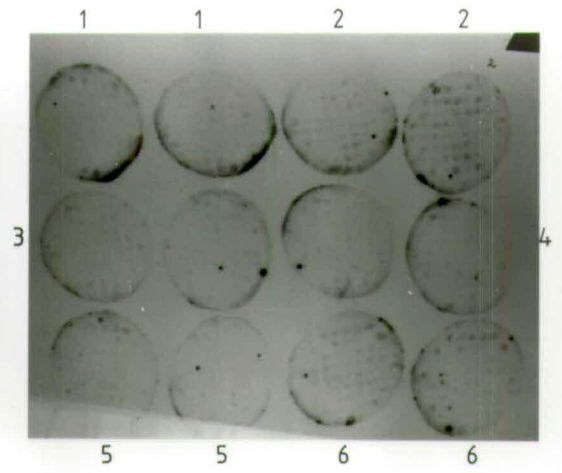
positions of clones which were known to contain inserted sequences, determined by Pst I digestion, and which were used as positive controls on each filter. - indicates positions of negative controls (pAT 153) on each filter.

(f) Positions of clones subsequently analysed are shown: a, 6.14; b, 1.34; c, 1.43; d, 1.48; e, 2.22; f, 2.42; g, 2.43; h, 2.50; i, 2.82; j, 4.2; k, 4.10; l, 4.11; m, 4.17; n, 4.19; o, 4.74; p, 4.84; q, 5.35; r, 5.47.

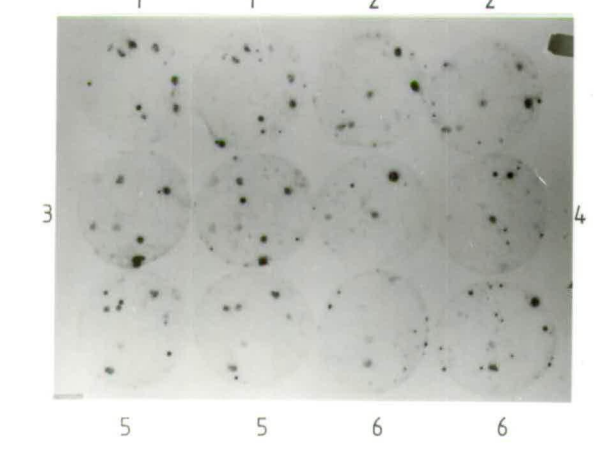
a.



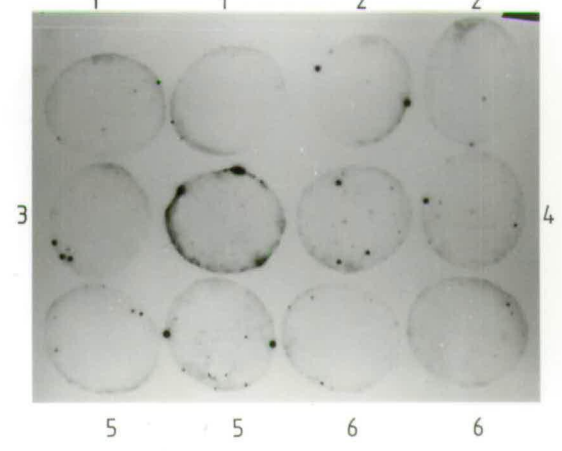
b.



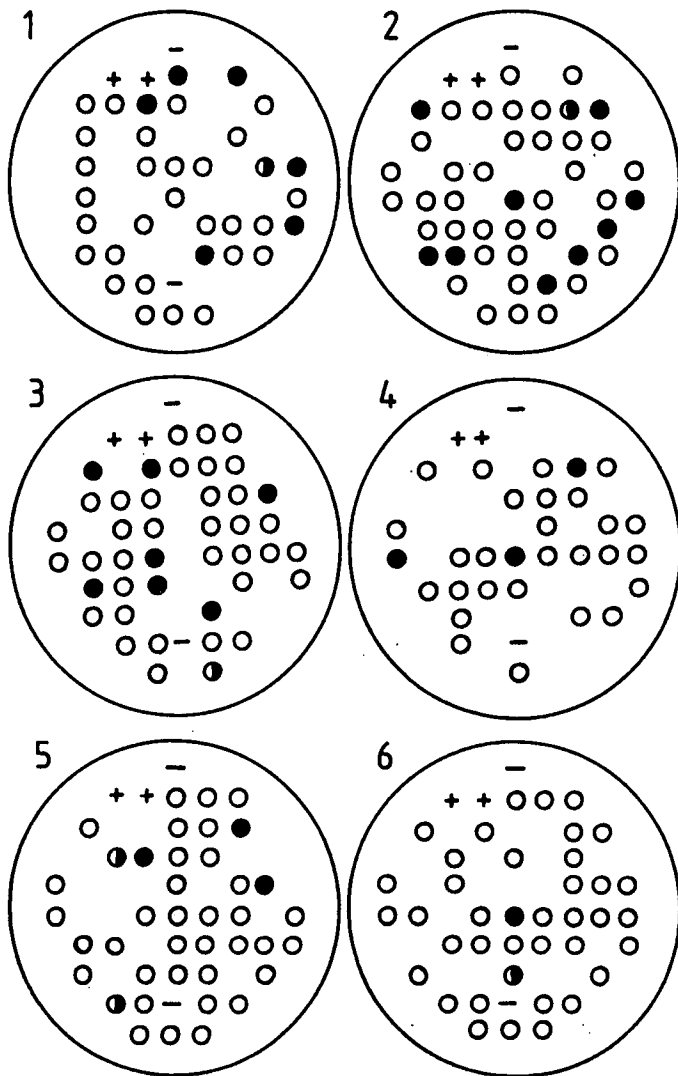
c.



d.



e.



f.

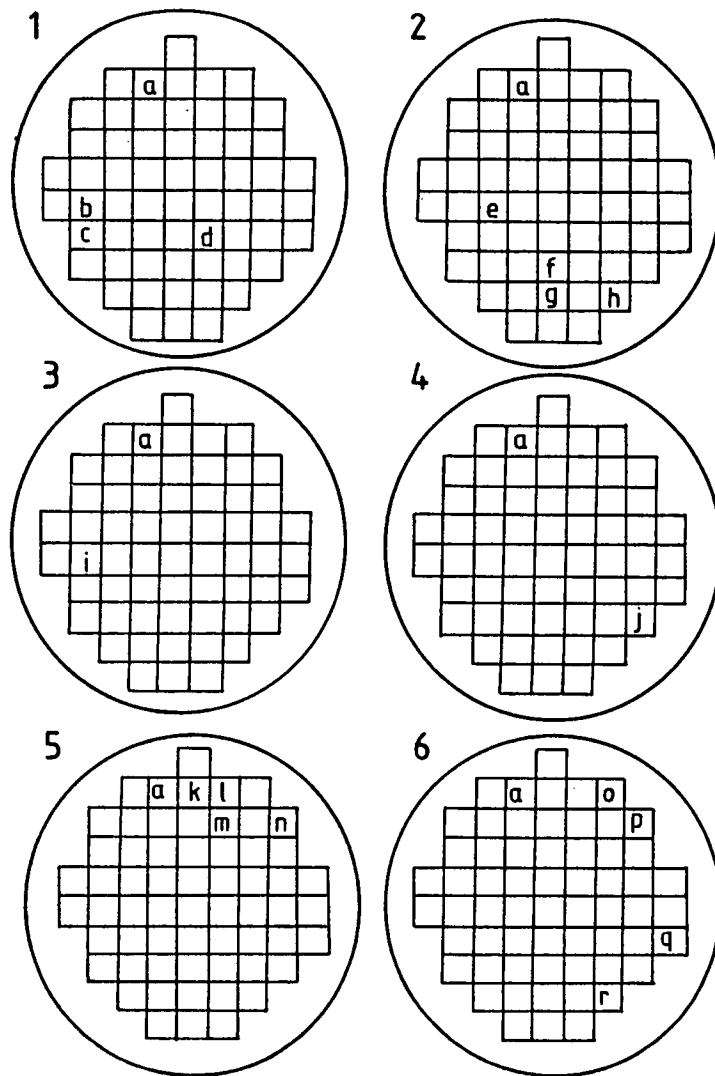


Table 4). Only 3% (i.e. 6) of the clones which hybridised to tadpole message also hybridised to adult globin mRNA (figure 16e, Table 4). About 9% (i.e. 30) of the clones hybridised only to the adult probe (figure 16c,e, Table 4).

Kinase labelled Xbt9S poly A minus RNA (specific activity approximately  $7.2 \times 10^6$  cpm/ $\mu$ g) was also hybridised to the same 337 clones (figure 16b). The amount of hybridisation was overall much lower than, but the general pattern appeared to be similar to, that obtained with Xbt9S poly A plus RNA (compare figure 16a and b).

Clones were screened with kinase labelled X. tropicalis blood cell rRNA (specific activity about  $2.6 \times 10^6$  cpm/ $\mu$ g) to check that hybridisation with 9S mRNA was not due to contaminating rRNA in the 9S RNA preparations. None of the clones showed any increase above the background hybridisation to control pAT 153 (figure 16d).

#### 4.3 Determination of insert sizes in X. borealis tadpole globin cDNA clones

DNA was prepared from randomly selected clones (Birnboim and Doly, 1979) and digested with EcoRI and Pst I separately. EcoRI cuts once within pAT 153 and therefore generates a linear molecule, unless the cDNA insert contains an EcoRI site in which case two fragments are produced. Pst I should cut out the cDNA sequence precisely, since cloning into the single Pst I site of pAT 153 by dG-dC tailing regenerates a Pst I recognition sequence at either end of the insert (figure 15). Figure 17 shows DNA from such digests separated on a 1.4% agarose gel. One of the clones (track 14, figure 17) contains

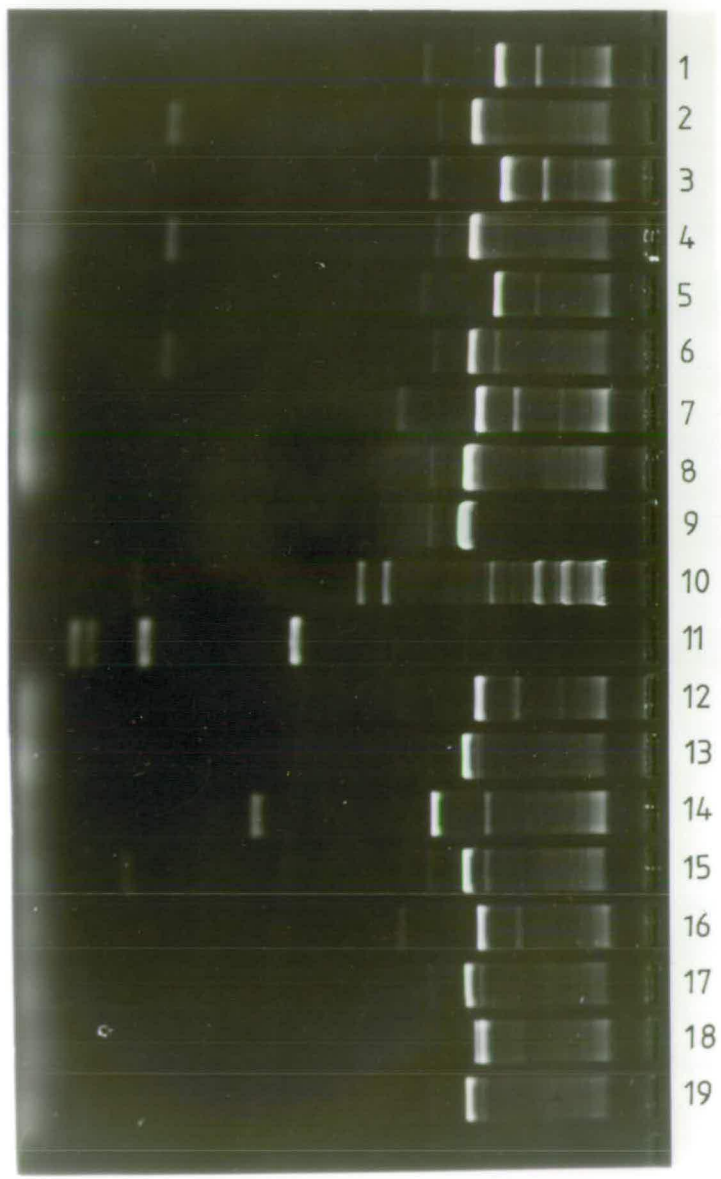
Figure 17. Agarose gel electrophoresis of Pst I and EcoRI restriction digest products from randomly selected clones

1µg of plasmid DNA, prepared by a modification of the method of Birnboim and Doly (1979), was digested with approximately 20 units of Pst I or EcoRI under conditions given in materials and methods. Products were electrophoresed through a 1.4% agarose gel at 70V constant voltage for 7h, then the gel was stained with 1µg/ml ethidium bromide for 30 minutes, destained in water for 30 minutes and photographed under UV illumination.

Tracks 1, 3, 5, 7, 12, 14, 16, 18      EcoRI restriction fragments  
Tracks 2, 4, 6, 8, 13, 15, 17, 19      Pst I restriction fragments

Size markers were run in tracks 9 (EcoRI cut pAt 153, 3657 nucleotides), 10 (Hind III cut  $\lambda^+$ , 23000, 9400, 6600, 4400, 2300, 2000, 560 nucleotides) and 11 (Taq I cut pAT 153, 1444, 609, 368, 314 nucleotides).

The fastest running bands in tracks 2, 4, 6 and 15 are cDNA sequences excised by Pst I. Track 14 shows that this clone has an EcoRI site within its cDNA sequence.





an EcoRI site within the inserted sequence. In some Pst I tracks smaller bands, corresponding to the inserts, cannot be seen e.g. tracks 8, 13, 17 (figure 17). Comparison of the EcoRI fragments with the large Pst I fragments (which are linear pAT 153) gives an indication of insert sizes. In all cases where no insert can be seen on Pst I digestion (figure 17) the sizes of these fragments are very similar, indicating that the inserts are small and have probably run off the bottom of the gel. Fragment sizes were determined as described in Appendix IV (modified from Southern 1979).

In order to determine insert lengths more accurately clones which were thought to have inserts greater than 500 nucleotides were restricted with Pst I and products analysed on longer 1.4% agarose gels (page 39). Table 6 gives insert lengths of clones containing cDNA sequences of at least 600 nucleotides.

The association between colony hybridisation signal with kinase labelled Xbt9S poly A plus RNA and insert size is shown in Table 5 and figure 18. In general clones containing larger inserts give more intense signals. However, some clones containing large inserts (i.e. over 600 nucleotides) give weak hybridisation signals, (e.g. three were classified as x, Table 5), and these presumably represent rare sequences in the total RNA population used for hybridisation.

#### 4.4 Restriction analysis of clones with inserts greater than 600 nucleotides

DNA was prepared from 500ml cultures of individual clones by a modification of the method of Birnboim and Doly (1979) and further

TABLE 5. Association between colony hybridisation signal and insert size.

Hybridisation class	no. clones	<sup>1</sup> i >600		600>i >500		500>i >400		i <400	
		no.	%	no.	%	no.	%	no.	%
xxx	37	11	30	12	32	10	27	4	11
xx	25	4	16	5	20	6	24	10	40
x	28	3	11	3	11	1	3	21	75
0	39	1	3	2	5	3	8	33	84
Total	129	19	15	22	17	20	15.5	68	52.5

129 clones were randomly selected and insert sizes determined by Pst I digestion and agarose gel electrophoresis. The same clones were also scored for signal intensity (see figure 16a) when hybridised to Xbt9S poly A plus RNA (see Table 4 and figure 16a for details). They were then sorted into classes according to insert size and hybridisation signal. The numbers and percentages of clones in each class are given.

1. i = insert size in nucleotides.

## TABLE 6

Restriction digests were as described in materials and methods. Clones were sorted into classes based on hybridisation signal in a colony hybridisation experiment with Xbt9S poly A plus RNA (figure 16a) and common restriction enzyme recognition sites. Values in brackets after clone numbers are insert sizes in nucleotides determined by Pst I digestion.

1. By comparison with restriction maps of X. laevis tadpole cDNA clones (Widmer et al, 1981) the clones shown here have been tentatively identified as  $\alpha$ - or  $\beta$ -globin coding sequences. Sequencing studies (Chapter 6) confirm these assignments for the two major groups.

2. These clones were not checked for Taq I sites.

3. pAT 153 contains no Pvu II site. When Pvu II cuts more than once in the insert it generates a large band, which is smaller than linear clone DNA, plus a number of small bands depending on how many Pvu II sites there are. 5.47 cut with Pvu II gave a large band which was smaller than EcoRI cut 5.47 indicating at least 2 Pvu II recognition sites. No smaller bands could be detected, therefore the exact number of sites could not be determined.

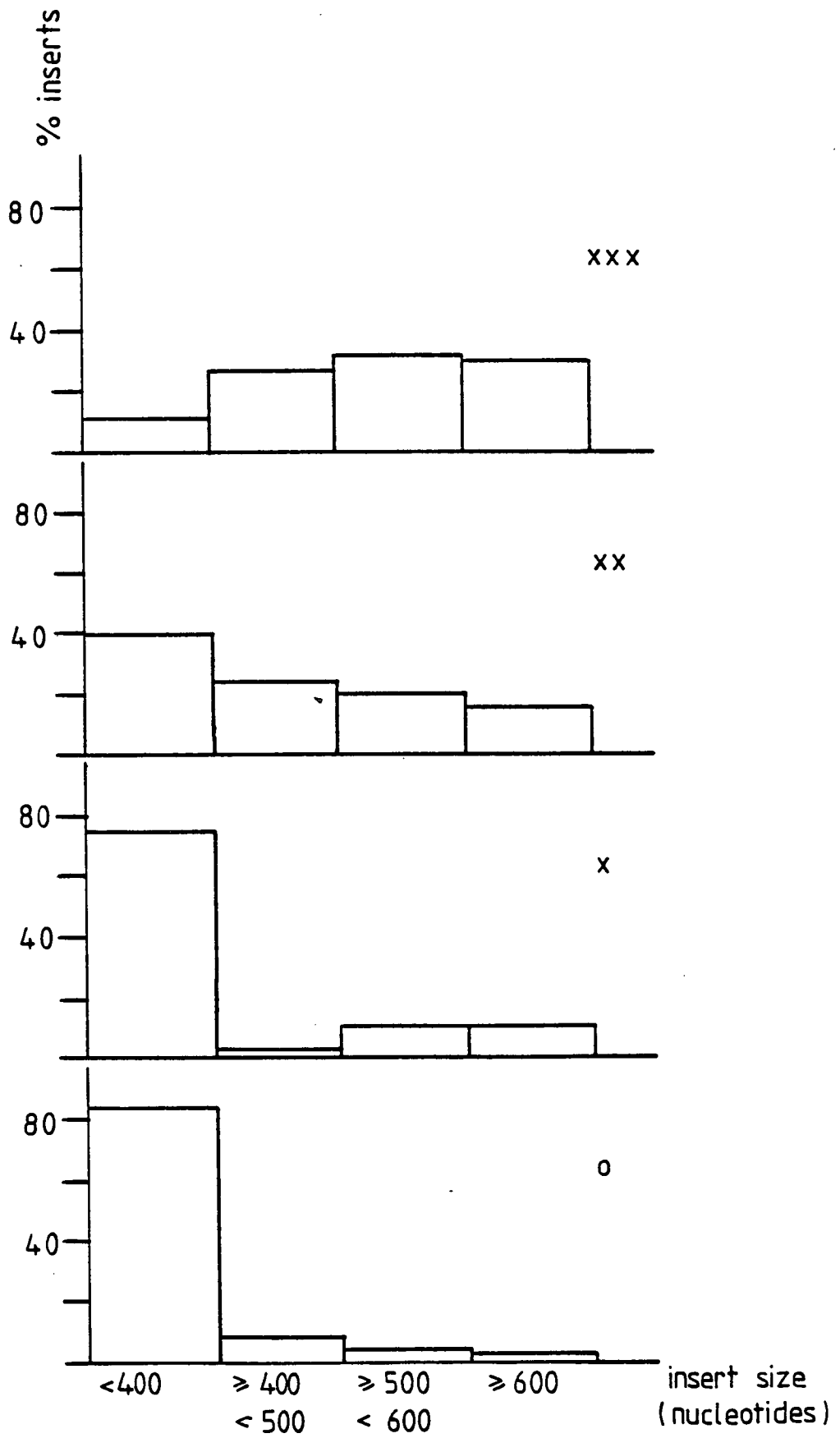
Positions of restriction sites relative to one another are shown for one member of each class in figure 19.

TABLE 6. Restriction analysis of clones with inserts  
greater than 600 nucleotides

Type of globin	Restriction sites	Hybridisation signal			
		xxx	xx	x	0
α	Bam HI, 1 Pvu II	1.34(642)		2.82(639)	
		1.43(720)			
	No EcoRI, Pst I,	4.84(703)			
	Hind III	1.48(614)			
α	EcoRI, 2 Hind III	2.22(634)	4.2(615)		
	No Pst I, Bam HI, Pvu II				
β	Pvu II, Taq I	2.42(619)	4.11(604)		
		2.48(613)	4.10(610)		
	No EcoRI, Pst I,	2.50(640)			
	Hind III, Bam HI	6.14(692)			
		2 (	5.35(618)		
	? Taq I (	4.19(610)			
	(	4.74(602)			
β	3 Pvu II		3 4.17(610)		
	2 Pvu II	5.47(623)			
	No EcoRI, Pst I, Bam HI, Hind III				
	Pst I, Pvu II				2.13(730)
	No EcoRI, Bam HI Hind III				

Figure 18. Association between colony hybridisation signal  
and insert size

See legend to Table 5 for details. Data presented in Table 5. are drawn as a histogram. The percentage of clones having inserts within the ranges shown on the x axis are given for each class of hybridisation signal.



purified on caesium chloride gradients. DNAs were digested with various restriction enzymes and clones sorted into classes based on common restriction sites. Table 6 summarises these results and shows that clones fall into two major and three minor groups. Clone 2.13 was not investigated further as it falls in a class of its own and shows no detectable hybridisation to Xbt9S poly A plus RNA.

Positions of restriction sites relative to one another were determined by double digests. This data is shown in Table 7 and figure 19.

#### 4.5 DISCUSSION

Insertion of a cDNA/RNA hybrid molecule into an E. coli plasmid (ColE1) was first described by Wood and Lee (1976). Their method was later modified by Zain et al (1979): dA-tailed adenovirus 2 fibre mRNA/cDNA was annealed to T-tailed Pst I cut pBR 322 and used to transform E. coli, since terminal transferase will catalyse the addition of homopolymer tracts to hybrid RNA/cDNA molecules using the same conditions described by Roychoudhury et al (1976) for tailing of duplex DNA fragments. The construction of recombinant clones from cDNA/RNA molecules has several advantages over the more conventional cloning of double stranded cDNA.

(1) Synthesis of a second cDNA strand is primed from a hairpin loop at the 3' terminus of the first strand, which is cut with S1 nuclease prior to tailing and inserting the cDNA into a plasmid (Efstratiadis et al, 1976). The hairpin loop can form without correct base-pairing at all positions and any mismatch is corrected after transformation

## TABLE 7

DNA was restricted with various enzymes, as indicated in the table, under conditions given in materials and methods. Products were separated on 1.4% agarose or 8% polyacrylamide gels and fragment sizes determined by a modification of the method described by Southern (1979), using a Texas Instruments SR56 programmable calculator (see Appendix IV for programme). In some cases in order to obtain reasonable size estimates of large and small fragments samples had to be run on both agarose and polyacrylamide gels.

### Notes:

1. The size (in nucleotides) of the insert indicated is that calculated from migration of the small fragment on Pst I digestion.
2. Where necessary positions of sites were determined by difference between fragments expected from pAT 153 and fragments obtained from clone DNA. The final map for each clone is derived from comparison of all digest data. Sizes are very approximate, especially when deduced by difference.
3. Addition of restriction fragment sizes within each insert predicts the total insert length. This is a less accurate estimate than obtained by measuring the small Pst I fragment (see note 1).
4. DNA fragments separated on 1.4% agarose gels.
5. DNA fragments separated on 8% polyacrylamide gels.

### Abbreviations:

Pv = Pvu II; B = Bam HI; H = Hind III; R = EcoRI; T = Taq I;  
d = doublet; p = partial digestion product; i = fragment derived from insert (only shown for Taq I digestion).



TABLE 7. Restriction digest data used to construct restriction maps shown in Figure 19

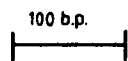
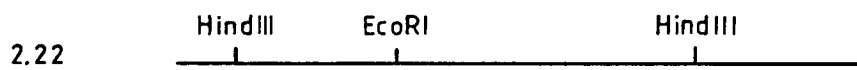
Clone (size) <sup>1</sup>	Restriction digest	Fragment sizes	Position of sites within insert <sup>2</sup>	deduced insert size <sup>3</sup>
1.43 (720)	Bam HI <sup>4</sup>	2826 1429	<u>296 , 302</u>	598
	Bam HI/Pst I <sup>4</sup>	2081 1448 372 <sup>d</sup>	<u>372 , 372</u>	744
	Pvu II/Pst I <sup>5</sup>	pAT 572 98	<u>98 , 572</u>	670
	Pvu II/Bam HI	2906 <sup>4P</sup> 2620 <sup>4</sup> 1503 <sup>4</sup> 283 <sup>5</sup>	<u>92 , 283 , 374</u> Pv B	749
2.22 (634)	Hind III <sup>4</sup>	4332 <sup>P</sup> 3379 <sup>P</sup> 2950 1334 <sup>P</sup> 918	<u>76 , 429 , 135</u>	640
	EcoRI <sup>4</sup>	3200 1250	<u>500 , 295</u>	795
	Hind III/Pst I <sup>5</sup>	2230 888 409 150 61	<u>61 , 409 , 150</u>	620
	EcoRI/Pst I <sup>5</sup>	2230 812 404 215	<u>404 , 215</u>	619
	EcoRI/Hind III <sup>5</sup>	2270 984 281 150	<u>61 , 154 , 269 , 135</u> H R H	619

TABLE 7 (continued)

Clone (size)	Restriction digest	Fragment sizes	Position of sites within insert	Deduced insert size
2.50 (640)	Taq I <sup>4</sup>	1332 <sup>i</sup>	<u>298, 363</u>	661
		773 <sup>i</sup>		
		609		
		459		
		368		
		314		
	Taq I/Pst I <sup>5</sup>	1077	<u>300, 372</u>	672
		562		
		422		
		403		
		372 <sup>i</sup>		
		315		
		300 <sup>i</sup>		
		141		
	Pvu II/Pst I <sup>5</sup>	pAT	<u>532, 106</u>	638
		532		
		106		
	Pvu II/Taq I <sup>5</sup>	1444 <sup>P</sup>	<u>124, 212, 274</u> Pv T	610
		1029 <sup>i</sup>		
		684 <sup>i</sup>		
		609		
		475		
		368		
		315		
		312		
		212 <sup>i</sup>		
		141		
	Pvu II/Taq I <sup>4</sup>	1345 <sup>P</sup>		
		1158		

Figure 19. Restriction sites found in *Xenopus borealis* tadpole  
9S cDNA sequences

Restriction sites were mapped with respect to one another using the data given in Table 7. The position of sites in 2.22 and the Bam HI site in 1.43 are very approximate, those in 2.50 and the Pvu II site in 1.43 were determined from the DNA sequence (Chapter 6). Horizontal lines represent cloned cDNA sequences excised from recombinant plasmids by Pst I digestion. Table 6 gives details of those restriction enzymes lacking recognition sites within the cDNA sequences.



by host repair mechanisms, which may result in introduction of mutations in cloned sequences (Weaver et al, 1981).

(2) The S1 nuclease treatment mentioned above results in loss of several nucleotides from the 5' end of the mRNA sequence, thus preventing a complete copy of the mRNA from being made. For this reason the 5' non-coding region of X. laevis adult  $\beta$ -globin cDNA is incomplete (Williams et al, 1980).

(3) Synthesis of a full length cDNA copy of mRNA is not 100 per cent efficient, many copies terminate part way through the message. Second strand synthesis also suffers from the same problem and therefore further reduces the probability of obtaining full length copies.

(4) Inserting cDNA/RNA hybrid molecules directly into plasmids eliminates several steps in the cloning procedure (i.e. alkaline hydrolysis of the RNA strand, second strand synthesis, S1 nuclease treatment). The cDNA/RNA therefore does not have to be recovered by, for instance, ethanol precipitation so many times and so less material is lost. This is particularly important when only small amounts of mRNA are available.

Size selection of tailed cDNA/RNA molecules on sucrose gradients did not eliminate all small molecules. There are several possible explanations for this. Since the cDNA remains linked to RNA a molecule containing a short cDNA strand could be "carried" by the longer RNA strand and sediment with larger molecules. On annealing with tailed plasmid the 5' end of the RNA remains single stranded

(figure 20a) and will be removed by host repair mechanisms after transformation.

Alternatively nicks in either the RNA or cDNA strand could act as sites for incorporation of nucleotides during the tailing reaction (Jackson and Berg, personal communication reported in Lobban and Kaiser, 1973). Annealing to plasmid followed by transformation results in molecules containing short stretches of DNA from randomly selected regions of the message (figure 20b). This possibility could be easily tested by analysing recombinants with small inserts.

Clones containing sequences complementary to tadpole globin mRNA were tentatively identified by colony hybridisation. RNA used in the preparation of these clones was isolated from stage 57-59 X. borealis tadpoles. In the previous chapter it was shown that RBCs of animals at these stages of development synthesise both adult and tadpole globins, although tadpole globin synthesis accounts for most protein synthesis. cDNA clones made from this RNA will therefore contain a low level of adult globin-coding sequences. Anaemic X. laevis adult globin mRNA does not contain any tadpole globin sequences and tadpole and adult globin sequences do not cross hybridise (Perlman et al, 1977). The same was assumed to be true for X. borealis. Thus hybridisation to adult 9S RNA will only identify those clones coding for adult globin message whereas hybridisation to tadpole 9S mRNA will identify clones containing either tadpole or adult sequences. Comparison of the two sets of data allows clones coding for tadpole sequences to be selected. At this stage it is not known whether all clones so chosen code for globins since any other polyadenylated 9S message will have been cloned too. However, it is

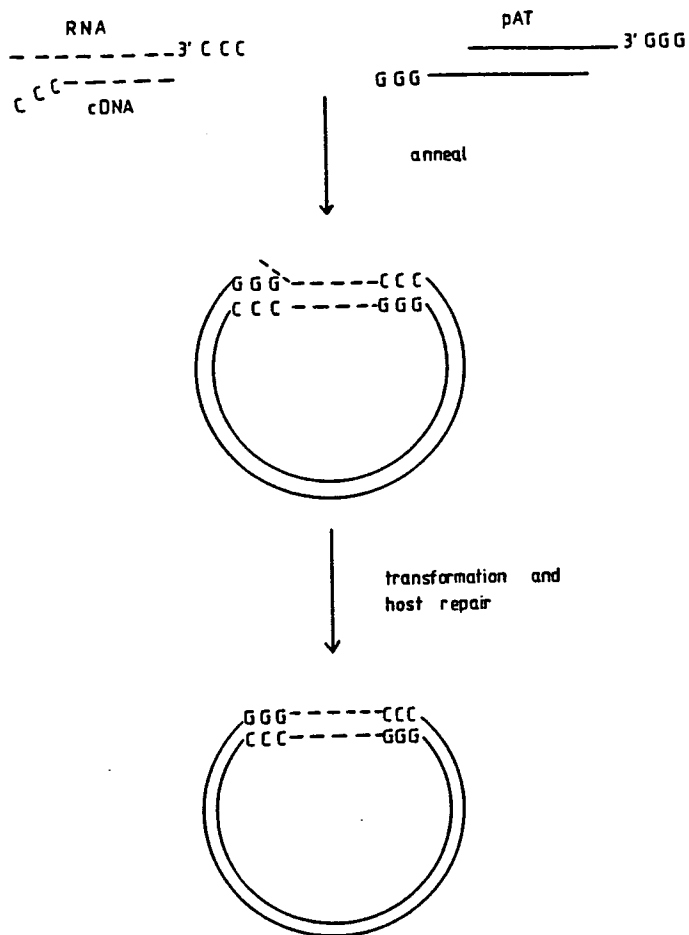
Figure 20. Hypothetical ways of producing clones with small inserts

A considerable proportion of clones contained small cDNA inserts (see figure 18 and text) despite the fact that dC-tailed cDNA/RNA hybrids were size fractionated on sucrose gradients prior to annealing with dG-tailed pAT 153.

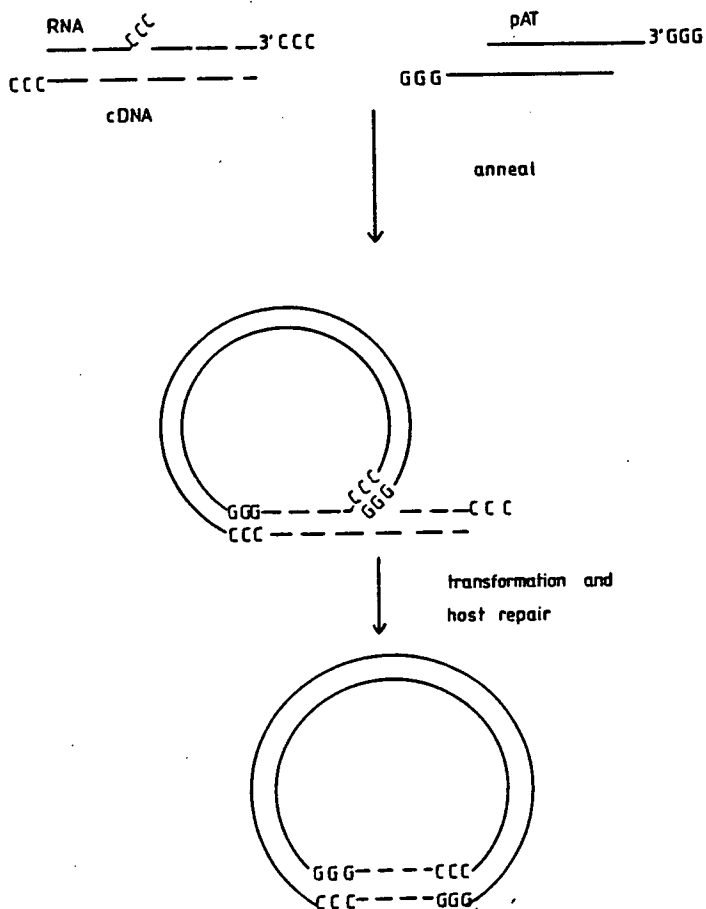
(20a) cDNA synthesis terminates prematurely and only a partial copy of the RNA is made. Tailing with dCTP results in the 3' ends of the RNA and cDNA possessing a poly C tract. The 5' end of the mRNA cannot base pair with the cDNA, and on annealing to dG-tailed pAT 153 a single stranded RNA tail is produced. After transformation this is removed by host repair mechanisms, resulting in a molecule containing the 3' part of the message, but lacking any 5' regions.

(20b) A nick is introduced into the RNA (or cDNA) strand. Terminal transferase catalyses the addition of dC residues to the 3' side of the nick as well as to the 3' ends of the RNA/cDNA hybrid. Annealing of dG-tailed pAT 153 to the poly dC tract extending from the nick, transformation and subsequent repair by the host results in a molecule containing the 5' part of the message. Clones produced by this mechanism will contain copies of various parts of the message, since nicks are introduced randomly in the RNA or cDNA.

a.



b.





assumed that the majority of clones do code for globins, since globin mRNA is the predominant mRNA in RBCs.

Colony hybridisation with Xbt9S poly A minus RNA gave the results expected if, as was suggested in the last chapter, the poly A minus RNA fraction contains a small proportion of all globin-coding sequences. Figure 16 shows that the overall level of hybridisation was lower, suggesting the poly A minus RNA contains fewer complementary sequences, but the general pattern was similar to that obtained with poly A plus RNA, indicating that a similar set of sequences are present.

It was hoped that colony hybridisation with Xbt9S poly A plus RNA might provide a simple way of selecting clones containing large inserts. The principle involved is as follows: the more DNA complementary to hybridising RNA a clone contains the more RNA it is able to hybridise. This means that clones with large inserts should give more intense signals on autoradiography than those with small inserts. The data shown in Table 5 and figure 18 show that, in general, this is the case. However an obvious disadvantage of selecting clones by this method alone is that those coding for sequences which are only rarely represented in the hybridising RNA population may be missed because they will only give weak signals. When trying to isolate a mixed population of sequences, some of which are likely to be less represented than others, it is therefore advisable to analyse randomly selected clones.

Since X. laevis adult globin mRNA is  $670 \pm 50$  nucleotides long (Kay et al, 1980) and X. borealis tadpole 9S mRNA has been shown to

contain components of approximately the same size and slightly larger (Chapter 3, figure 11) clones containing inserts over 600 nucleotides in length were chosen for further study. X. borealis 9S RNA also contains a component which is smaller than X. laevis adult mRNA (figure 11) and clones complementary to this RNA, if they existed, would not have been selected. However, it was shown that this RNA does not prime cDNA synthesis (figure 12b), and neither  $\alpha$  - nor  $\beta$  - tadpole globin cDNA sequences hybridise to it (Chapter 5).

Based on common restriction sites 19 clones were sorted into five classes. One class contained only one member and did not detectably hybridise to Xbt9S poly A plus RNA or to X. borealis adult globin mRNA. This clone (2.13, Table 6) was not investigated further, but it is probable that it contains a non-globin sequence which is present at a low level in Xbt9S poly A plus RNA. Hybrid-selection of RNA, from total Xbt9S poly A plus RNA, homologous to 2.13 cDNA would show whether or not it coded for globin and analysis of the translation products (assuming they were not globins) on SDS-polyacrylamide gels by fluorography would provide an estimate of the protein size. Sequencing the cDNA would show directly whether it coded for globin. The other four classes could be tentatively classified as consisting of  $\alpha$  - or  $\beta$ -globin coding sequences by comparison with the restriction data from X. laevis tadpole globin cDNA clones (Widmer et al, 1980). This was subsequently confirmed for the two major classes by DNA sequencing (Chapter 6).

CHAPTER FIVE

CORRELATION OF XENOPUS BOREALIS TADPOLE  $\alpha$ - AND  $\beta$ -GLOBIN  
cDNA SEQUENCES WITH PROTEIN AND RNA BANDS ON GELS

## 5.1 INTRODUCTION

Globins of X. borealis tadpole RBCs have been shown to consist of three major and five minor species (Chapter 3, figure 1). cDNA clones prepared from Xbt9S poly A plus RNA and described in Chapter 4 might therefore be expected to contain sequences coding for each one of these proteins. Hentschel et al (1979) showed by hybrid-arrested translation (Paterson et al, 1977) that X. laevis adult globin cDNA clones coded for separate proteins on denaturing polyacrylamide gels, indicating that there are at least three genes coding for adult X. laevis globins. In this chapter Xbt9S cDNA clones are shown, by positive selection of homologous RNA and cell-free translation (Ricciardi et al, 1979), to code for different protein bands seen on triton-acid-urea polyacrylamide gels.

X. borealis tadpole RBC 9S RNA is heterogenous (see Chapter 3, figure 12) and it was thought likely that the different bands code for different globin species. In order to establish the coding capacity of individual RNA bands nick-translated X. borealis tadpole  $\alpha$ - and  $\beta$ -globin cDNA clones were used in RNA gel blot hybridisations (Thomas, 1980).

## RESULTS

### 5.2 Hybridisation-selection of homologous RNAs

10 $\mu$ g of DNA from individual clones was partially depurinated, denatured, applied to small areas of nitrocellulose and hybridised to

5 µg Xbt9S poly A plus RNA for 16h in 70% formamide, 0.4M NaCl at 50°C (method modified from Ricciardi et al, 1979). The RNA Cot of the reaction was 3.84 and the calculations given in Appendix VI show that 99.9% of any RNA sequence present at a level of 1% or more in the total hybridising RNA population should have hybridised in 2.8h. Bound RNA was eluted, translated in rabbit reticulocyte lysate and products analysed on triton-acid-urea polyacrylamide gels by fluorography. Incorporation of label into products was at a variable low level (see Table 8) but no attempt was made to load the same amount of labelled product, 10µl of each sample were loaded into gel slots, since the rate of migration of individual components is sensitive to total protein loaded in this system.

Figure 21 shows autoradiographs from two such experiments. As negative controls filters containing plasmid DNA from a X. borealis adult globin cDNA clone and pAT 153 DNA were also included in the hybridisation. No proteins could be detected when samples eluted from these filters were translated (figure 21b tracks 1 and 6) indicating that RNA had not detectably hybridised to these DNAs.

Initially poly A was omitted from hybridisation solutions and clone 6.14 DNA cross hybridised with RNA coding for all translation products of Xbt9S poly A plus RNA (figure 21a track 3). This is presumably because clone 6.14 possesses a particularly long poly A-T tract. Poly A was added to subsequent hybridisations to compete with RNA for poly T rich regions in the DNA and, as expected, under these conditions clone 6.14 hybridised to RNA coding for one specific set of bands (figure 21c track 7).

TABLE 8. In vitro translation data for hybridisation-  
selected RNAs

RNA source <sup>1</sup>	cpm/ $\mu$ l lysate x 10 after incubation <sup>4</sup> for time (mins)							Fluorograph figure
	10	20	30	40	60	90	120	
-mRNA <sup>3</sup>	0.5	0.8		1.1				
2.50	1.1	1.2		2.9			5.5	
6.14	0.6	1.4		4.0			6.1	19a
2.82	1.6	1.3		5.1			12.2	
1.43	0.5	0.8		1.9			2.7	
adult	0.3	0.4		2.0			2.0	
2.22	1.7	2.5		4.1			2.7	
4.17	0.3	0.8		0.6			1.8	19b
pAT <sup>2</sup>	0.5	0.5		0.5			1.6	
unbound <sup>2</sup>	5.6	7.5		13.5			28	
unbound <sup>2</sup>			0.9		1.6	2.0	2.6	
4.17			1.2		1.2	2.1	2.5	
pAT			1.0		1.2	2.3	2.5	19c
Adult			0.9		1.4	2.1	2.4	
5.47			0.7		1.4	2.0	2.4	
6.14			0.9		1.8	2.1	3.3	
-mRNA <sup>3</sup>			0.4		0.7	1.3	2.5	

RNA was hybridised to plasmid DNA bound to nitrocellulose, eluted and translated in rabbit reticulocyte lysate (as described in materials and methods and legend to figure 21). (1) RNA used for translation hybridised to DNA as indicated. (2) RNA which had not hybridised to any cloned DNA at the end of the reaction was recovered from the hybridisation solution by ethanol precipitation and translated. (3) As a control to check for endogenous rabbit protein synthesis in the micrococcal nuclease treated lysate, a sample was incubated with radioactive amino acids in the absence of additional mRNA. (4) 1 $\mu$ l Aliquots were TCA precipitated at intervals during incubation and incorporation of label determined.

Figure 21      Fluorographs of translation products from RNA  
selected by hybridisation to X. borealis tadpole  
blood cell 9S cDNA clones

10µg plasmid DNA from X. borealis tadpole blood cell 9S cDNA clones was bound to small areas of nitrocellulose and hybridised to 5µg total Xbt9S poly A plus RNA in 100µl of 70% formamide, 0.4M NaCl, 0.01M PIPES pH7 at 50°C for 16h (modified from Ricciardi et al, 1979). Bound RNA was eluted from each filter separately and translated in nuclease treated rabbit reticulocyte lysate. 10µl of each of the [<sup>3</sup>H]-leucine and [<sup>3</sup>H]-histidine labelled translation products were separated on triton-acid-urea 10% polyacrylamide gels, run at 7mA constant current for 16h at 20°C, and fluorographed. Xbt9S poly A plus RNA was also translated in the reticulocyte lysate and run on gels as a control. 21a and 21b show the results of one experiment and 21c those of a separate experiment.

(21a) Fluorograph exposed for 24h. Products are derived from RNA bound to the following clone DNAs: track 1, 2.50; track 3, 6.14; track 4, 2.82; track 5, 1.43. Tracks 2 and 6 show 40 Kcpm and 80 Kcpm translation products from Xbt9S poly A plus RNA respectively. Track 7 shows endogenous rabbit protein synthesis in the absence of additional mRNA.

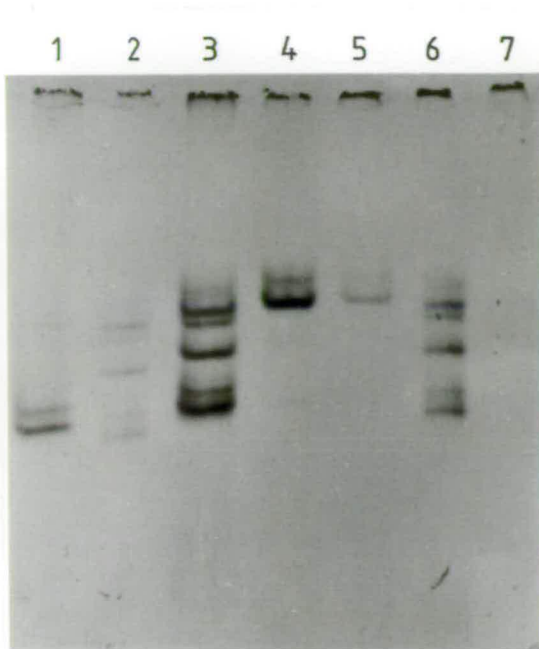
(21b) Fluorograph exposed for 24h at -70°C. Products are derived from RNA binding to the following clone DNAs: track 1, adult X. borealis globin clone; track 3, 2.22; track 5, 4.17; track 6, pAT 153. Tracks 2 and 7 show 80 Kcpm and 40 Kcpm respectively Xbt9S

poly A plus RNA translation products. Track 4 shows products from the "unbound" RNA sample, i.e. that RNA which had failed to hybridise by the end of the reaction.

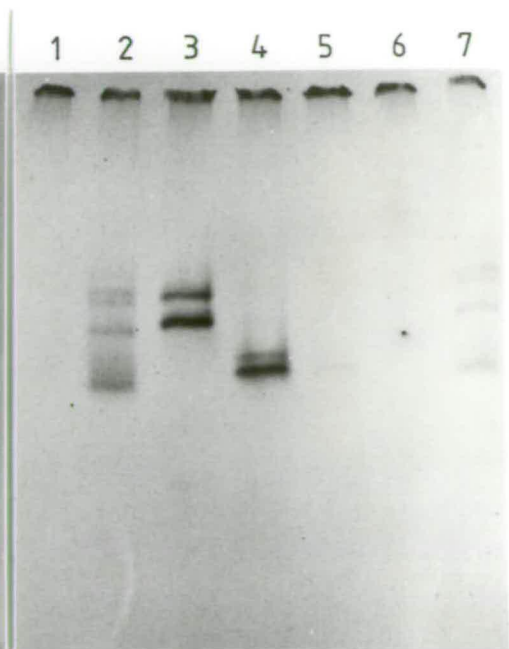
(21c) Fluorograph exposed for 6 weeks 4 days at  $-70^{\circ}\text{C}$  tracks 1, 2, 3, 5, 7, 8, 9 and 24h tracks 4 and 6. Tracks 4 and 6 were cut out of a photograph of the fluorograph exposed for 24h and inserted into relevant positions on a photograph of the 6 week 4 day exposed fluorograph. This was necessary since the control tracks (4 and 6) exposed more quickly than experimental tracks. Products are from RNA hybridising to the following DNA: track 3, 4.17; track 5, 5.47; track 7, 6.14; track 8, adult X. borealis globin clone; track 9, pAT 153. Track 1 shows endogenous rabbit protein synthesis, track 2 products of the "unbound" RNA sample and tracks 4 and 6, 45 Kcpm and 90 Kcpm Xbt9S poly A plus translation products.



a.



b.



c.

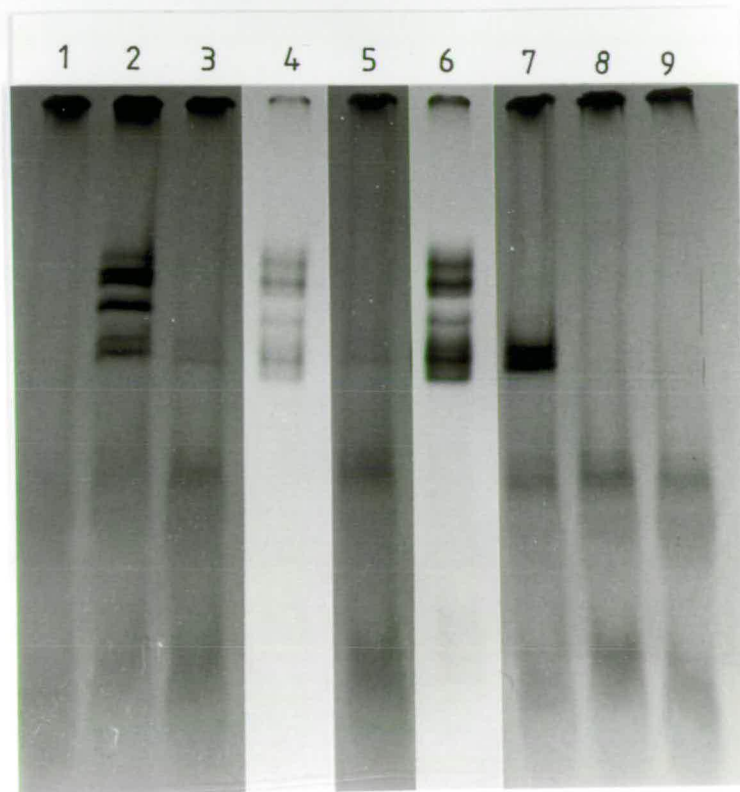


Figure 21b track 4 and 21c track 2 show translation products from "unbound" RNA i.e. that RNA which had failed to hybridise by the end of the reaction. Figures 21a and 21b are derived from one hybridisation-selection experiment and figure 21c from another. In the first experiment (figure 21a and b) all RNA coding for the slower migrating proteins (labelled 1a, 1b, 2 figure 1) has hybridised to DNA (figure 21a track 4, 2.82; track 5, 1.43; figure 21b track 3, 2.22) as shown by the absence of these proteins in translation products of the unbound RNA. Clones 2.50 (track 1, figure 21a), 6.14 (track 3, figure 21a) and 4.17 (track 5, figure 21b) hybridised RNA coding for the fastest running components (labelled 3, 3a, 3b, 3c figure 1) but a high proportion of this RNA remained unhybridised and stimulated translation of the proteins seen in track 4, figure 21b. When the experiment was repeated with clones 4.17, 5.47 and 6.14, all of which encode the fastest migrating proteins (figure 21c, tracks 3, 5, 7) there still remained some unhybridised RNA coding for band 3 proteins. There are two possible explanations for this: (1) there was insufficient complementary DNA on the filters to remove all this RNA; (2) there exists another class of clones which also code for band 3 proteins. These alternatives could be distinguished by rehybridising 6.14 DNA to the unbound RNA fraction. If saturation of the DNA was the problem, as suggested by the first explanation, then this should remove the remaining complementary RNA from the unbound RNA fraction. If, on the other hand, there is another class of clones as yet unidentified then reselection with 6.14 DNA would have no effect and translation products of the unbound RNA would be identical to those originally obtained. If the latter result was found then further analysis of the class of clones of which 6.14 is a member might yield two subsets, since restriction site data used to

separate the classes will not detect all differences between members. Alternatively there might be another class of clones amongst those so far not analysed.

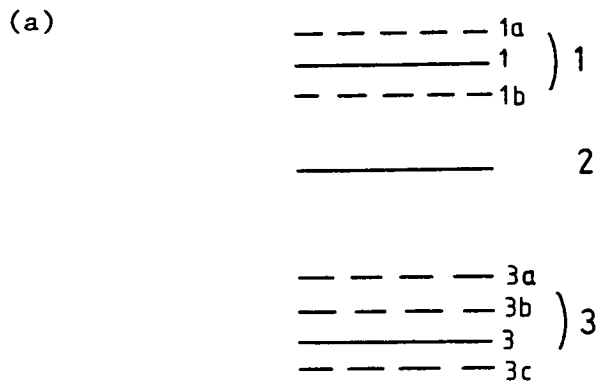
Figure 22 summarises the results presented in figure 21. Clones 1.43 and 2.82 code for proteins in the slowest migrating bands (labelled 1, figure 22) 2.82 apparently hybridised more RNA than 1.43 (figure 21a tracks 4 and 5), which is surprising since the reverse was the case in colony hybridisation experiments (Table 6). However, both 2.82 and 1.43 hybridise to the same RNA species, as predicted since they fall into the same class with respect to restriction enzyme sites (Table 6). 2.22 was the only clone coding for band 2 proteins (figure 21b, track 3) and it also hybridised to a lesser extent with RNA coding for band 1 proteins.

Clones 2.50, 4.17, 5.47 and 6.14 (figure 21a track 1, 21c tracks 3, 5 and 7) code for proteins in band 3. 4.17 and 5.47 appear to hybridise very little RNA, as judged by the extreme faintness of the bands seen on fluorographs of translation products from RNA hybridised to these clones. 6.14 and 2.50 apparently hybridise more RNA which suggests that they contain a major globin sequence whereas 4.17 and 5.47 code for minor globins. (Major and minor refer to relative abundances of each globin species in blood.)

### 5.3 Hybridisation of RNA bound to nitrocellulose

Filters containing varying amounts of ribosomal, Xbt9S poly A plus and Xbt9S poly A minus RNA arranged in a grid on the nitrocellulose were hybridised with either an  $\alpha$ - or a  $\beta$ -tadpole

FIGURE 22    Summary of hybrid-selection experiments



(b)

Clone/Band	1	2	3
1.43	++		
2.82	+++		
2.22	++	+++	
2.50			+++
6.14			++
5.47			+
4.17			+

See materials and methods and legend to figure 21 for experimental details.

(a) The pattern of protein bands from X. borealis tadpole RBCs electrophoresed through triton-acid-urea polyacrylamide gels is shown diagrammatically. 1, 2, 3 refer to major groups of bands (see for

example figure 1), no distinction is made between bands within a group in this summary.

(b) RNA was hybridised to plasmid DNA from the clones indicated, eluted and translated in rabbit reticulocyte lysate. Products comigrated with bands in group 1, 2 or 3 as labelled in figure 22a. The intensity of each signal on the fluorographs, shown in figure 21, varies and is scored thus:

+ = very faint band(s);

++ = intermediate;

+++ = very dark band(s).

globin cDNA probe. Autoradiographs of filters are shown in figure 23. The spots on the nitrocellulose were then cut out and counted in a scintillation counter, this data is presented in Table 9 and figure 24. The graph (figure 24) indicates that Xbt9S poly A plus RNA contains approximately seven times as much  $\beta$ -globin mRNA and four times as much  $\alpha$ -globin mRNA as the poly A minus fraction.

Xbt9S poly A plus and poly A minus RNAs were further analysed by Northern blots. RNA was electrophoresed through formaldehyde-agarose gels, transferred to nitrocellulose and hybridised separately to  $\alpha$ - and  $\beta$ -tadpole globin cDNA probes, the first probe used being washed off before the second one was hybridised. Autoradiographs of filters which have been hybridised to each probe are shown in figures 25 and 26. The autoradiographs have been photographed on the same scale and aligned, showing that each probe hybridises to a different RNA band. Comparison with the photograph (figure 25b) of ethidium bromide stained marker RNAs, run on the same gel as the RNA that the filter was made from, shows that the  $\alpha$ -globin probe hybridises to the slowest migrating RNA band (labelled 848, figure 25a) and the  $\beta$ -probe to the band running just ahead of this (labelled 680, figure 25a). The RNA species in these bands are approximately  $848 \pm 66$  and  $680 \pm 56$  nucleotides long, as discussed on page 85.

Neither probe hybridised to the fastest running components seen on the stained gel (figure 25b) and which can also be detected at a relatively high level, compared to the larger RNAs, in the poly A minus fraction (figure 12 and page 85).

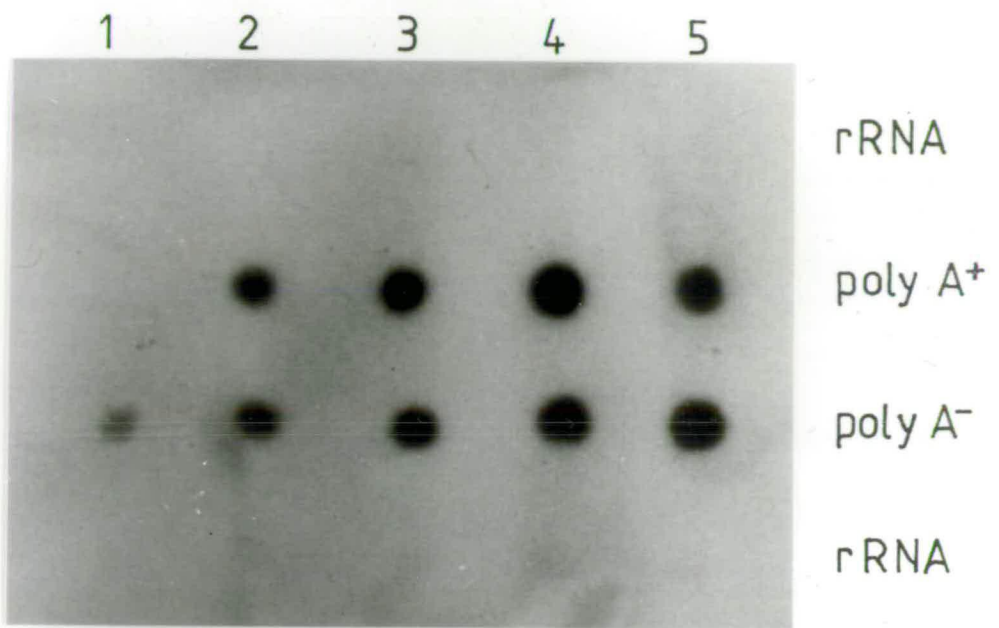
In an attempt to identify these smaller RNA species, the

Figure 23. Autoradiographs of nick-translated  $\alpha$ - and  $\beta$ -globin specific probes hybridised to Xbt9S poly A plus and poly A minus RNA.

RNA was spotted on to nitrocellulose as described in materials and methods and legend to Table 9. 0.2 $\mu$ g  $\alpha$  (clone 1.43) and  $\beta$  (clone 2.48) globin recombinant plasmid DNAs were nick-translated to specific activities of  $3.4 \times 10^8$  and  $3.6 \times 10^8$  cpm/ $\mu$ g respectively and hybridised to filters as described in materials and methods. Autoradiographs were exposed for 7h at  $-70^{\circ}\text{C}$  with a Kodak intensifying screen. Amounts of RNA in spots are as follows: top rRNA row and poly A plus RNA row on each filter, 1, 10ng; 2, 20ng; 3, 40ng; 4, 100ng; 5, 200ng; poly A minus RNA and lower rRNA rows, 1, 100ng; 2, 200ng; 3, 300ng; 4, 400ng; 5, 500ng.

- (a) Filter hybridised to  $\alpha$ -globin specific probe (1.43).
- (b) Filter hybridised to  $\beta$ -globin specific probe (2.48).

a.



b.

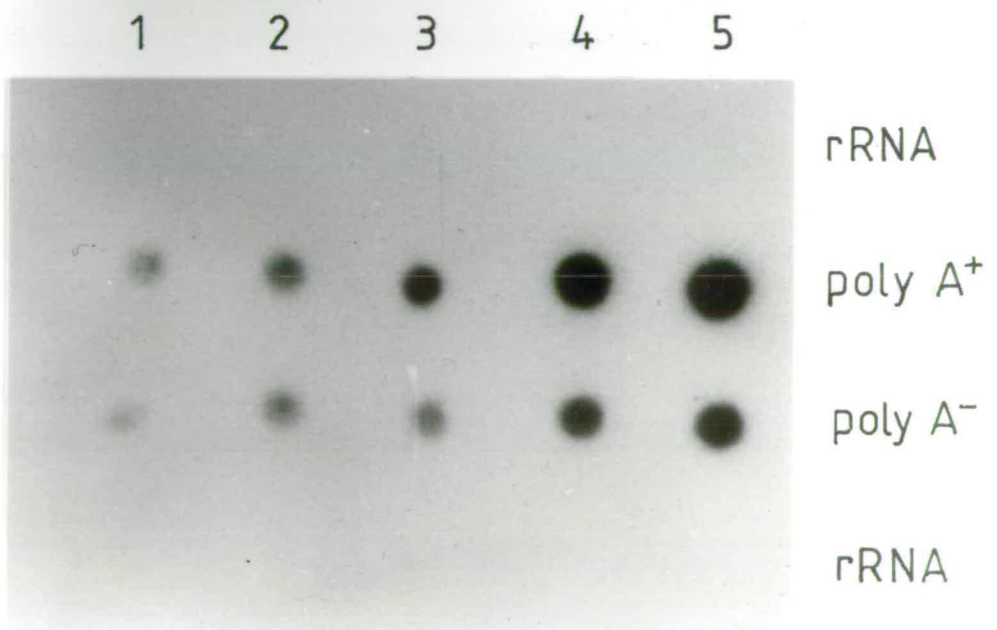




TABLE 9. Data from dot hybridisation analysis of Xbt9S poly A plus and poly A minus RNA hybridised to  $\alpha$ - and  $\beta$ -globin specific DNA probes.

RNA fraction	ng RNA in dot	cpm bound -background		linear regression analysis			
		$\alpha$	$\beta$	$r^1$	$\beta$	$m^2$ cpm/ng RNA	
		$\alpha$	$\beta$	$\alpha$	$\beta$	$\alpha$	$\beta$
A <sup>+</sup>	10	13	80	0.99 <sup>3</sup>	0.95	3.0	5.6
	20	267	123				
	40	348	206				
	100	514	286				
	200	365	1193				
A <sup>-</sup>	100	191	35	0.87	0.95	0.8	0.8
	200	228	80				
	300	258	97				
	400	304	223				
	500	566	350				

Varying amounts of Xbt9S poly A plus and poly A minus RNA were spotted on to nitrocellulose and hybridised to nick-translated plasmid DNAs from either an  $\alpha$ -globin (1.43) or a  $\beta$ -globin (2.48) clone, as described in materials and methods. As controls for non-specific background hybridisation the same amounts of ribosomal RNA (from X. borealis tadpole RBCs) were also spotted on to the filters. Specific activities of the probes were 1.43,  $3.4 \times 10^8$ , 2.48,  $3.6 \times 10^8$  cpm/ $\mu$ g and approximately 0.2 $\mu$ g DNA were used for each hybridisation in 20ml hybridisation buffer as described in materials and methods.

After washing the filters Xray film was exposed for 7h at -70°C with a Kodak intensifying screen (see figure 23). RNA spots were then cut out and radioactivity in each one determined by liquid scintillation counting. Results, corrected for background

hybridisation, are presented above and, to enable direct comparison of poly A plus and poly A minus RNA fractions, graphically in figure 24. There was considerable background on the filter hybridised to the  $\alpha$ -globin probe (figure 23a) and therefore the results presented above and in figure 24 are likely to be less accurate for the  $\alpha$ - than  $\beta$ -globin probe. The best fit straight line, shown in figure 24, was calculated by linear regression on a Texas Instruments SR56 calculator (see Appendix V for programme). 1,  $r$  = correlation coefficient; 2,  $m$  = slope of best fit straight line i.e. cpm bound per ng RNA; 3, only the values of cpm binding to the 20, 40 and 100 ng spots were used in the linear regression analysis of this hybridisation, since the other two values appear to be anomalous (see figure 24).

Figure 24. Dot hybridisation of Xbt9S poly A plus and poly A minus RNA with cloned  $\alpha$ - and  $\beta$ -globin specific probes

Experimental details are given in materials and methods and legend to Table 9. The best-fit straight lines were calculated by linear regression as described in the legend to Table 9 and Appendix V.

- ⊖—⊖  $\alpha$ -globin probe (1.43) hybridised to poly A plus RNA;
- ⊖—⊖  $\alpha$ -globin probe (1.43) hybridised to poly A minus RNA;
- ×—×  $\beta$ -globin probe (2.48) hybridised to poly A plus RNA;
- ⊖—⊖  $\beta$ -globin probe (2.48) hybridised to poly A minus RNA.

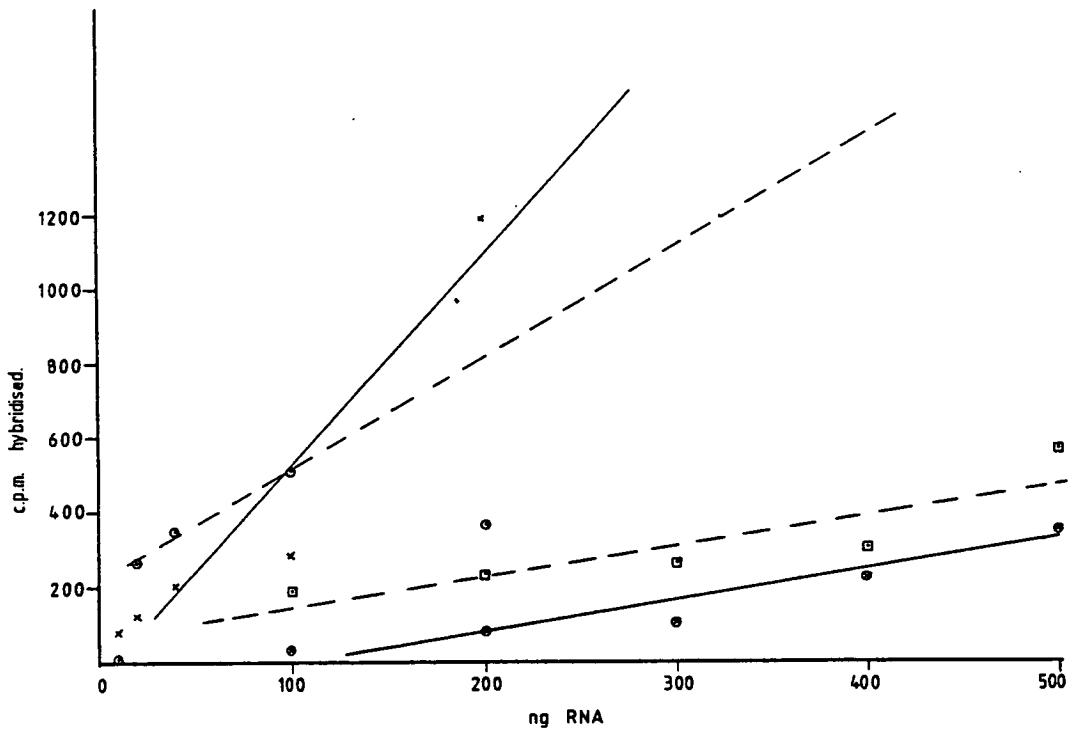


Figure 25. Analysis of Xbt9S poly A plus and poly A minus RNA  
by RNA gel blot hybridisations

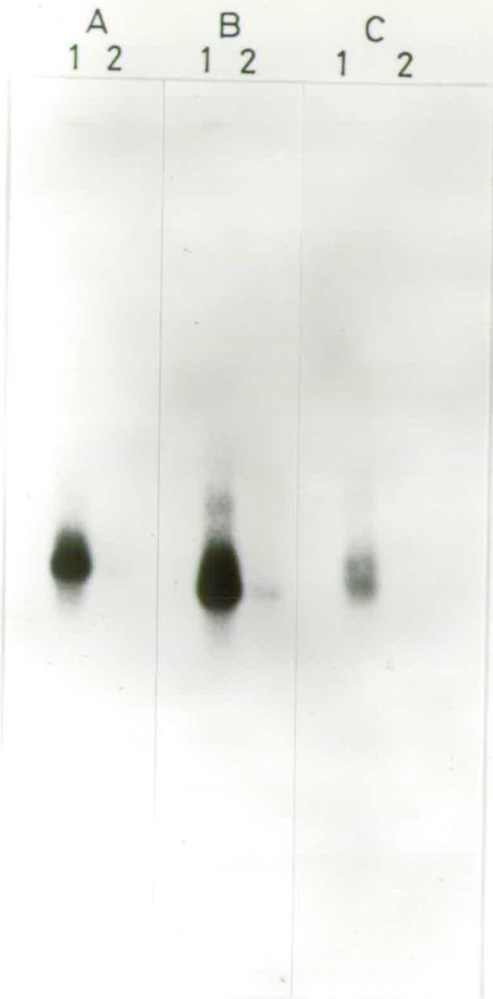
25a. Details of experimental procedures are given in materials and methods. 1 $\mu$ g Xbt9S poly A plus (track 1) and 10 $\mu$ g Xbt9S poly A minus (track 2) RNA were electrophoresed through a 1.5% agarose-formaldehyde gel, transferred to nitrocellulose and hybridised successively to nick translated cloned  $\beta$ - (2.48) and  $\alpha$ - (1.43) globin probes and then to total cDNA prepared from Xbt9S poly A plus RNA. Specific activities of probes were 1.43,  $3.4 \times 10^8$ ; 2.48,  $3.6 \times 10^8$ ; cDNA about  $5 \times 10^8$  cpm/ $\mu$ g. Each probe was washed off the filter before the next one was hybridised. Autoradiographs were exposed as follows: A, 1.43, 5h at  $-70^\circ\text{C}$  with a Kodak intensifying screen; B, 2.48, 7h at  $-70^\circ\text{C}$  with a Kodak intensifying screen; C, cDNA, 16h at room temperature without an intensifying screen. Photographs of autoradiographs were taken on the same scale and aligned to allow direct comparison of band positions.

A =  $\alpha$ -globin probe (1.43); B =  $\beta$ -globin probe (2.48);  
C = cDNA probe.

25b. 5 $\mu$ g 28S+18S RNA (track 1) and 5 $\mu$ g Xbt9S poly A plus RNA (track 2) were run as markers alongside RNA which was to be transferred. They were run on each side of the gel to determine whether the gel ran straight. That part of the gel containing markers was cut away from the rest prior to transfer, stained with 5 $\mu$ g/ml ethidium bromide in 0.1M ammonium acetate for 30 minutes, destained for 2h in 0.1M ammonium acetate and photographed under UV illumination. Marker tracks from one side of the gel only are shown here. 848 and 680

indicate sizes of the RNA bands hybridising  $\alpha$ - and  $\beta$ -globin probes respectively, estimated from this gel, that in figure 12a and one other (not shown) to be  $848 \pm 66$  and  $680 \pm 56$  nucleotides.

a.



b.

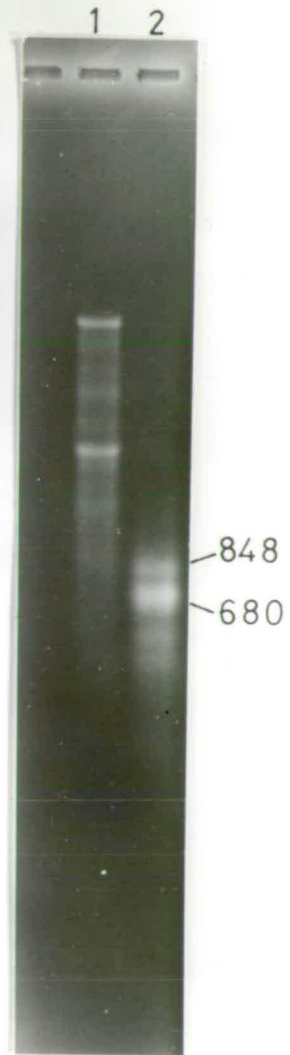
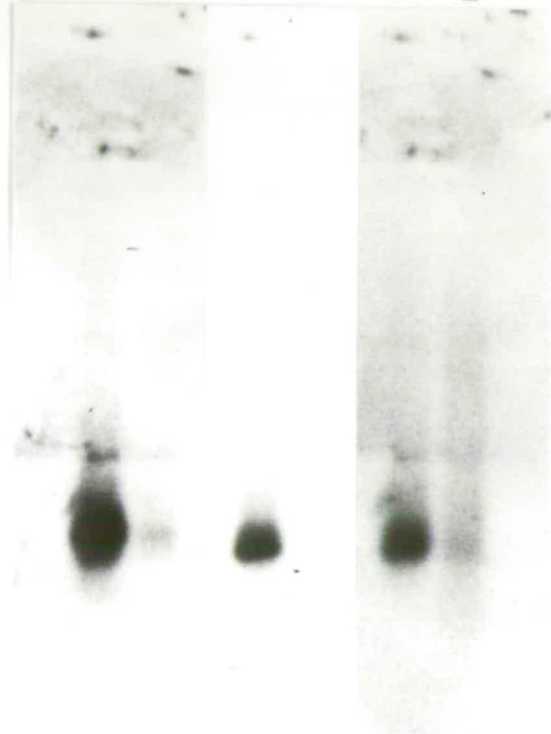


Figure 26. Analysis of Xbt9S poly A plus and poly A minus RNA  
by RNA gel blot hybridisation

See legend to figure 25a and materials and methods for details. The filter shown in this figure was first hybridised to nick-translated 1.43 plasmid DNA, then 2.48 and finally to an adult globin cDNA clone. 1.43 ( $\alpha$ ) and 2.48 ( $\beta$ ) probes were the same ones used for figure 25a hybridisations. They were boiled for 10 minutes prior to use in case they had renatured. Specific activity of the adult probe was  $2.1 \times 10^8$  cpm/ $\mu$ g. Autoradiographs were exposed at  $-70^\circ\text{C}$ , with Kodak intensifying screens for the following times: 1.43, 7h; 2.48, 5h; adult, 48h. Photographs were taken on the same scale and aligned to permit direct comparison of band positions. A =  $\alpha$ -globin probe (1.43); B =  $\beta$ -globin probe (2.48); C = adult probe. Track 1 = 1 $\mu$ g Xbt9S poly A plus RNA; track 2 = 10  $\mu$ g Xbt9S poly A minus RNA. 848 and 680 indicate the sizes of the mRNAs to which the probes are hybridising (see figure 25b and legend).



A            B            C  
1 2        1 2        1 2



848  
—  
680

filters were hybridised either to cDNA prepared from Xbt9S poly A plus RNA or to a X. borealis adult globin cDNA probe. Figure 25a shows that Xbt9S cDNA hybridises to both the  $\alpha$ - and  $\beta$ -globin RNA bands. The adult probe (figure 26) hybridises to a band of approximately the same size as the  $\beta$ -tadpole globin probe. Neither the total cDNA nor the adult globin cDNA probe showed any detectable hybridisation to the small RNA bands seen in figure 25b, even on overexposure of the autoradiographs.

#### 5.4 DISCUSSION

Hybridisation-selection of RNA homologous to cloned cDNA sequences and subsequent in vitro translation of bound RNA permits correlation of cloned sequences with specific protein bands on a gel. In this way it has been shown that X. borealis tadpole globin clones 1.43 and 2.82 select RNA coding for the slowest migrating group of proteins seen on triton-acid-urea polyacrylamide gels (labelled 1, figures 1 and 22). Clone 2.22 hybridised to RNA coding for proteins in groups 1 and 2 (figure 22), which suggests that they may be related. It is unclear from the autoradiographs (figures 21a and 21b) whether 2.22 is giving rise to exactly the same protein band as clones 1.43 and 2.82. In order to load sufficient radioactivity into each track of the gels it was necessary to use 10  $\mu$ l of lysate, causing endogeneous rabbit proteins to interfere with electrophoresis and slow down migration of X. borealis tadpole globins compared to those in control tracks. It is therefore difficult to correlate precisely which protein band within a group is being produced by a particular clone. Since the slowest migrating band shown by 2.22 (figure 21b track 3) runs just ahead of the slowest band in the

control track (figure 21b track 2) it is likely that clone 2.22 is hybridising RNA coding for proteins in the second slowest migrating band (labelled 1b figures 1 and 22). A similar comparison for clones 1.43 and 2.82 (figure 21a tracks 4 and 5) suggests that they are coding for the major protein in group 1 (labelled 1 figure 1). Assuming that these clones are hybridising different RNA species this explains why 1.43 and 2.82 show no cross hybridisation to RNA encoding band 2 protein.

Clones 2.50, 6.14, 4.17 and 5.47 code for proteins in the fastest migrating group (labelled 3, figures 1 and 22). 4.17 and 5.47 apparently hybridised for less RNA than did 2.50 and 6.14, as judged by the ability of the eluted RNA to stimulate incorporation of radioactive amino acids into proteins in rabbit reticulocyte lysate. This suggests that 5.47 and 4.17 select an RNA species which is only rarely represented in the total Xbt9S poly A plus RNA population, whereas RNA complementary to 2.50 and 6.14 cDNA is represented at a relatively high level. The cDNA sequences of these three clones are all apparently related, since they select RNAs coding for the same proteins.

Clones used in the hybridisation-selection experiments fall into four classes (1.43 + 2.82; 2.22; 2.50 + 6.14; 4.17 + 5.47) with respect to the restriction enzyme sites they contain (Table 6). As expected, clones from the same class select the same RNA species.

Comparison with the restriction enzyme data from X. laevis tadpole globin cDNA clones (Widmer et al, 1981) allows tentative identification of  $\alpha$ - and  $\beta$ -globin coding clones. 1.43 + 2.82 and

2.22 contain sequences similar to X. laevis tadpole major and minor  $\alpha$ -globin cDNA sequences respectively. 2.22 cannot be said to code for a minor  $\alpha$ -globin, as defined by the relative amount of each globin species in the blood, since figure 1 shows that the centre band (labelled 2, figure 1) is as well represented as the other two major bands (labelled 1 and 3, figure 1) in globins extracted from tadpole RBCs.

2.50 + 6.14 and 4.17 + 5.47 cDNA sequences are similar to those of X. laevis tadpole minor and major  $\beta$ -globin clones respectively. However, unlike the situation in X. laevis where clones with a similar restriction pattern to 2.50 and 6.14 code for minor and those like 4.17 and 5.47 for major globin species (Widmer et al, 1981) in X. borealis clones 2.50 and 6.14 appear to code for major and 4.17 and 5.47 for minor globins. Data presented here suggests that 2.50 and 6.14 code for RNA present at much higher levels in total Xbt9S poly A plus RNA than that encoded by 4.17 and 5.47. The number of clones isolated in each class of which 2.50 and 6.14 or 4.17 and 5.47 are members also supports this idea (see Table 6). 2.50 and 6.14 are two of nine similar clones isolated, whereas 5.47 and 4.17 are the only representatives of their type. However, so few clones have been analysed, (18, Table 6), that this result may not be significant.

In vitro translation of Xbt9S poly A plus and poly A minus RNA suggests that the poly A plus fraction contains about 10 to 40 times more coding RNA than the poly A minus fraction (see Chapter 3 for details). Hybridisation of X. borealis tadpole  $\alpha$ - and  $\beta$ -globin coding cDNA sequences to Xbt9S poly A plus and poly A minus RNA bound to nitrocellulose confirms this observation: figure 24 shows that

poly A plus RNA hybridises about 7 times as much  $\beta$ -globin cDNA as does poly A minus RNA. A similar result is also obtained with an  $\alpha$ -globin probe.

Electrophoretically separated Xbt9S poly A plus and poly A minus RNA transferred from formaldehyde - agarose gels to nitrocellulose and hybridised to X. borealis tadpole  $\alpha$ - and  $\beta$ -globin cDNA probes identifies bands containing  $\alpha$ - and  $\beta$ -globin mRNA. Figure 12a shows both poly A plus and poly A minus RNA fractions to be heterogeneous, poly A plus containing relatively more of the two slowest migrating components than poly A minus RNA. The smallest RNA species are approximately equally represented in both fractions (figure 12a). Figures 25 and 26 show that X. borealis tadpole  $\alpha$ - and  $\beta$ -globin cDNA probes hybridise to the two slowest migrating RNA bands, the  $\alpha$ -probe to one about  $848 \pm 66$  and the  $\beta$ -probe to one about  $680 \pm 56$  nucleotides. cDNA probes used to detect these bands were 720 (1.43,  $\alpha$ -probe) and 613 (2.48,  $\beta$ -probe) nucleotides long respectively (Table 6) and therefore contain almost full length copies of the globin messages, probably missing some of the 5' residues.

Ten times as much poly A minus RNA (10 $\mu$ g) compared to poly A plus RNA (1 $\mu$ g) had been run on these gels, since as mentioned above the poly A minus fraction was thought to contain less coding RNA than the poly A plus fraction. The intensity of hybridisation signals with both probes was significantly lower for the poly A minus RNA (figures 25a and 26) than for the poly A plus RNA providing further evidence that most X. borealis tadpole globin message molecules are polyadenylated.

Neither  $\alpha$  - nor  $\beta$ -globin probes hybridised to the smaller RNA bands (figures 12 and 25b) which are present in roughly equal amounts in poly A plus and poly A minus RNA (figure 12). A X. borealis adult globin cDNA probe (figure 26) and total cDNA prepared from Xbt9S poly A plus RNA (figure 25a) also failed to hybridise to these smaller RNA components. This result is paradoxical since molecules selected twice by binding to oligo dT cellulose should have poly A tracts of at least 15-20 nucleotides (Cabada et al, 1977) capable of priming reverse transcription. However, total cDNA hybridised only to RNA bands comigrating with molecules identified by hybridisation with cloned tadpole  $\alpha$  - and  $\beta$ -globin coding sequences (figure 25a). The small RNAs may bind to oligo dT cellulose but apparently not prime cDNA synthesis for a number of reasons: (1) the poly A is at or near the 5' end of the molecule; (2) the molecules contain poly U tracts not poly A and hybridise to oligo dT through the poly A of mRNA molecules already bound; (3) the small molecules have a collection of dispersed short oligo A tracts capable, by cooperation, of binding to oligo dT but incapable of priming reverse transcription. The relatively high proportion of small RNA in the poly A minus fraction also suggests that it is binding inefficiently to oligo dT cellulose. Repeated passage of the column flow through fractions over the oligo dT column to recover maximum poly A plus RNA might raise the amount of inefficiently bound RNA contaminating poly A plus RNA.

Figure 26 shows a X. borealis adult globin cDNA probe hybridised to Xbt9S poly A plus and poly A minus RNA. The autoradiograph was exposed for 48h with an intensifying screen at -70°C, almost ten times longer than necessary for tadpole probes (e.g. 1.43 figure 25a), indicating that adult globin mRNA is present

at a relatively low level in Xbt9S RNA. The adult probe hybridised to a band in about the same position as the  $\beta$ -globin tadpole probe, of approximate size  $680 \pm 56$  nucleotides. X. borealis adult globin mRNA thus appears to be about the same size as that from X. laevis, reported to be  $670 \pm 50$  nucleotides (Kay et al, 1980).

CHAPTER SIX

DETERMINATION OF XENOPUS BOREALIS TADPOLE  $\alpha$ - AND  $\beta$ -GLOBIN

cDNA SEQUENCES



## 6.1 INTRODUCTION

The Sanger sequencing method (Sanger et al, 1977; Sanger and Coulson, 1978) is based on the ability of DNA polymerase to copy accurately a single stranded (ss) DNA template. A short primer is annealed to the template and extended 5' to 3' by DNA polymerase Klenow fragment (which lacks 5' - 3' exonuclease activity) in the presence of the four deoxynucleoside triphosphates (dNTPs) (one of which is radioactively labelled) and each one of the 2', 3'-dideoxynucleoside triphosphates (ddNTPs) in turn. When a ddNTP is incorporated instead of a normal dNTP there is no 3' hydroxyl group from which polymerisation can continue. Thus a set of molecules are produced, all having common 5' ends but differing at their 3' ends. Inclusion of each ddNTP in turn gives molecules all ending at a specific nucleotide and parallel electrophoresis of all four reaction products on acrylamide gels (Sanger and Coulson, 1978) allows the sequence to be deduced.

Single stranded template is conveniently produced by subcloning into bacteriophage M13, an E. coli male specific single-stranded filamentous phage which has been genetically modified to produce suitable cloning vehicles (Messing et al, 1977, Messing et al, 1981). Initially the E. coli lac regulatory region and the genetic information for the  $\alpha$  peptide of  $\beta$ -galactosidase ( $\beta$ -D-galactoside galactohydrolase, E.C.3.2.1.23) were inserted into a non-essential region of the M13 genome (Messing et al, 1977). Derivatives of this have been produced containing an array of restriction sites within the  $\beta$ -galactosidase gene, which do not affect the functioning of this gene (Messing et al, 1981 M13mp7; Messing unpublished M13mp8 and

M13mp9). The sequence of M13mp8 in the region of the lac Z gene is shown in figure 27.

The lac Z gene in the recombinant phage, when induced by IPTG, produces the  $\alpha$ -peptide of  $\beta$ -galactosidase which cannot function alone. By growing the phage in a suitable host (e.g. E. coli JM101) which has its own  $\beta$ -galactosidase gene deleted but contains a defective  $\beta$ -galactosidase (missing residues 11-41) on a plasmid, a functional enzyme can be produced by complementation. Normal phages are identified by including a lactose analogue, BCIG, on the plate which is hydrolysed by  $\beta$ -galactosidase to a blue dye, bromochloroindole, giving rise to blue plaques. In contrast, when the lac Z gene is interrupted by cloned sequences inserted at any of the restriction sites (figure 27) a functional  $\beta$ -galactosidase cannot be produced and plaques formed are white. A very simple visual selection identifies phage containing inserted sequences.

Cloning into M13 is also a convenient way of obtaining separately both strands of a DNA sequence. Screening plaques with either RNA or cDNA probes identifies clones containing strands complementary to each probe. This enables both strands to be sequenced, providing confirmation of sequences obtained without having to resort to other sequencing methods.

## 6.2 RESULTS

cDNA sequences from two clones (1.43 and 2.50, thought to be  $\alpha$ - and  $\beta$ -globin sequences respectively, see Chapter 4) were excised by Pst I digestion and ligated with Pst I cut M13mp8 RF. Since small

Figure 27. Sequence of M13mp8 in the region of the lac Z gene

M13mp8 sequence was compiled by J. Messing. The oligonucleotide delineated /...../ was inserted into the lac Z gene, but does not affect  $\alpha$ -complementation. Restriction sites suitable for cloning into mp8 and primer annealing position are shown below the sequence.



fragments are preferentially included in ligation reactions there was no need to separate cDNA sequences (about 600 nucleotides) from plasmid (about 3.5Kb). Competent JM101 were transformed with aliquots of the ligation mixes and plated on indicator plates. A total of 15 white plaques were obtained from each, see Table 10 for details. They were screened by hybridisation to Xbt9S poly A plus RNA and cDNA made from the same RNA. Figure 28 shows results of RNA hybridisation and Table 11 gives details of clones hybridising to each probe. 12 1.43 and 7 2.50 recombinants hybridised to RNA, 2 and 5 respectively to cDNA and 1 and 3 gave no signal with either probe. Theoretically 50% of clones should hybridise to each probe, however the numbers involved here are so small that deviation from expectation is probably not significant. Clones giving no signal are most likely due to ligation of M13mp8 with pAT 153 or JM101 endogenous plasmid DNA copurifying with M13mp8 replicative form DNA, although it is possible that deletions in the lac Z gene may have occurred during the cloning procedure.

Templates were prepared from several white plaques which hybridised to RNA or cDNA probes. To check their quality each was used in a single sequencing reaction incorporating ddT and analysed on a gel run for about 4h. This produces a characteristic pattern of bands for each sequence and definition and darkness of bands distinguishes poor from good templates. A good template should give distinct, dark bands all the way up the gel and there should be a region of heavy labelling near the origin demonstrating that larger fragments, not separated on a short run, have been produced. It is not clear what causes the production of good or poor templates and it is therefore necessary to repeat the single track screening for every

TABLE 10. Preparation of M13mp8 recombinants

Clone	$\mu$ l ligation mix used to transform	No. blue plaques	No. white plaques	% white plaques
1.43	$\sim$ 0.25	39	3	7.1
	"	64	3	4.5
	"	24	0	0
	"	69	1	1.4
	"	114	7	5.8
2.50	$\sim$ 0.5	16	1	5.9
	"	204	2	1.0
	"	103	1	1.0
	"	286	2	0.7
	"	281	7	2.4
	"	164	0	0
	"	186	2	1.1

Pst I digested M13mp8 and plasmids from X. borealis tadpole globin cDNA clones were ligated as described in materials and methods. Initially 2 $\mu$ l of each ligation mix was used to transform 0.2ml competent JM101, as described in materials and methods. Amounts of ligation mix used in further transformations were adjusted so that there should have been between 1-200 plaques per plate, thus keeping individual plaques well separated. Transformed cells were plated on BBL plates in BBL top agar containing a lac inducer IPTG and indicator BClG. White plaques are those containing inserted sequences.

Figure 28. Autoradiograph of M13mp8 recombinant plaque  
hybridisation to Xbt9S poly A plus RNA.

White plaques, from transformation of JM101 with M13mp8 ligated to X. borealis tadpole RBC 9S cloned cDNA sequences, were grown up in grids on BBL plates and transferred to nitrocellulose (method modified from Benton and Davis, 1977). Filters were hybridised to kinase-labelled Xbt9S poly A plus RNA (Donis-Keller et al, 1977), specific activity approximately  $10^7$  cpm/ $\mu$ g or to cDNA to identify phage containing cDNA-like or RNA-like strands respectively. The autoradiograph of the RNA hybridisation shown was exposed for 8h at  $-70^{\circ}\text{C}$  with a Kodak intensifying screen. The two filters shown are duplicates.

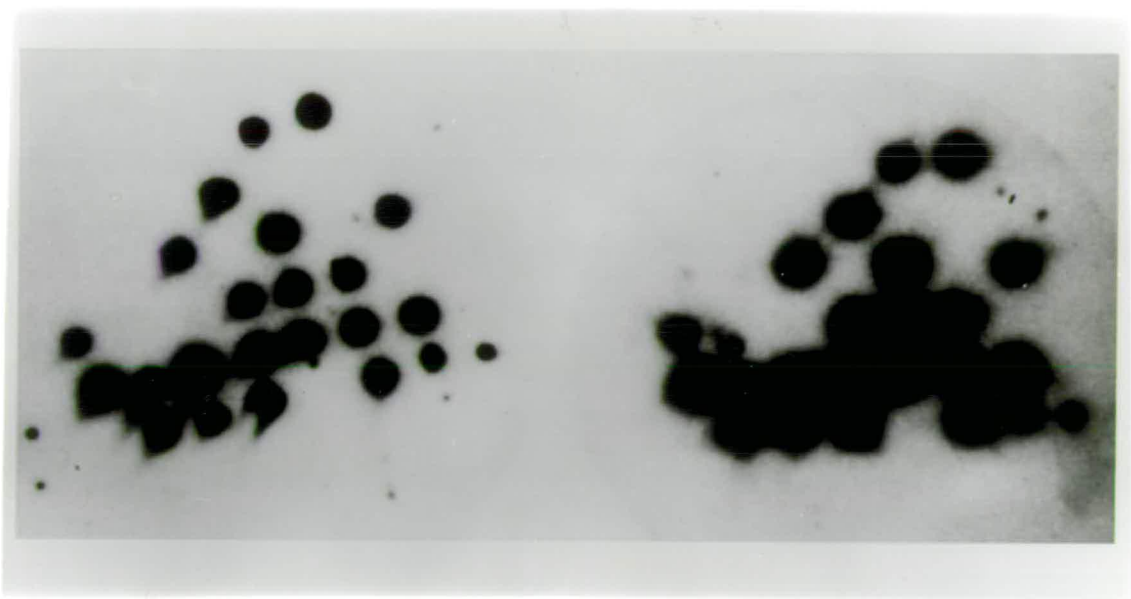




TABLE 11. Plaque hybridisation of M13mp8 recombinants

Clone	No. white plaques	No. no signal	No. cDNA signal	No. RNA signal
1.43	15	1	2	12
2.50	15	3	5	7

White plaques from transformations with ligated M13mp8 RF and Xbt9S cDNA clones were screened for the presence of X. borealis sequences by plaque hybridisation to Xbt9S poly A plus RNA or cDNA made from the same RNA. This data was derived from the autoradiograph shown in figure 28 (RNA hybridisation) and another similar one of the cDNA hybridisation.

template.

The best templates were picked and sequenced as described (page 74). Figure 29 is an example of results obtained. The left hand four tracks (figure 29) show M13mp8 sequence in the region of the lac Z gene, the others are from recombinants. As can be seen from figure 29 sequences of recombinant clones were much more difficult to read than those of mp8. There was considerable laddering in all tracks of the recombinants after the poly G stretches. At the junction of poly G and what was expected to be poly T or globin sequence (figure 30) the polymerase was apparently unable to distinguish between nucleotides and all four were randomly inserted. This is known to be a problem with DNA polymerase. In some areas, further up the gel, it was relatively easy to decide which nucleotide was correct since one particular band was much darker than the rest, however some positions are ambiguous.

Nucleotide sequences of clones 1.43 and 2.50 are shown in figures 31 and 32. Approximately 200 nucleotides were read from each template, thus the centre sequences from each clone are missing. Amino acid sequences, written above the nucleotides (figure 31 and 32), have been predicted from DNA sequences, since no protein data is available for X. borealis tadpole globins.

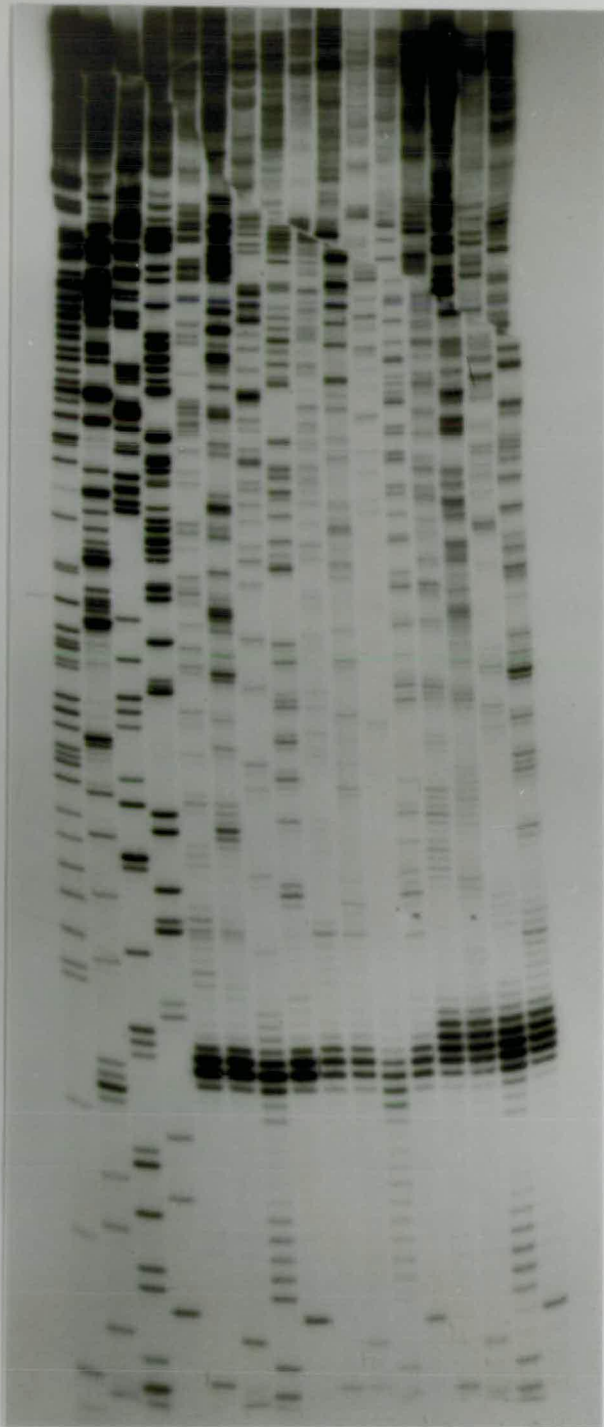
An open reading frame, starting at the first ATG (59 nucleotides from the 5' end) of 1.43 cDNA sequence, encodes 52 amino acids of an  $\alpha$ -globin polypeptide. Evidence for this is the conservation of amino acids (underlined in figure 31) as compared to other known  $\alpha$ -globins (Dayhoff, 1976). Comparison with partial

Figure 29. Autoradiograph of Sequencing Gel

Four templates were sequenced using chain-terminating inhibitors as described in materials and methods (Sanger et al, 1977; Sanger, 1980). Electrophoresis (Sanger and Coulson, 1978) was for about 4h at 1.2KV constant voltage, until the xylene cyanol marker dye had travelled about three quarters of the way down the gel. The gel was fixed for 10 minutes in 10% acetic acid, rinsed in distilled water and blotted dry on Whatman 3MM paper. The autoradiograph was exposed for 20h at room temperature.

T, C, G, A indicate chain termination at respective nucleotides.

1, M13mp8 sequence; - 2, Xbt9S clone 1.43 cDNA-like strand;  
3, Xbt9S clone 1.43 RNA-like strand; 4, Xbt9S clone 2.50 cDNA-like strand.



T C G A T C G A T C G A T C G A  
1 2 3 4

Figure 30. Subcloning into M13mp8 - expected sequences at the junctions of M13 and inserted sequences

DNA can be cloned into M13 in two orientations, one of which results in the RNA-like strand and the other the cDNA-like strand being present in viral ss DNA.

1. Cloned sequence and surrounding Pst I sites in plasmid pAT 153.
2. Recombinant plasmid (1) cut with Pst I generates a small fragment containing cDNA sequence plus pAT 153 (not shown).
3. Ligation of the sequence shown in (2) to Pst I cut M13mp8 RF gives rise to M13 recombinants containing cDNA sequences inserted in both orientations. M13 sequences are underlined, the top strand shown is viral ss DNA and therefore that which is sequenced.
4. Sequencing commences at a primer annealed 3' to the ss sequence shown. From the construction of the original plasmid clones the first few nucleotides of inserted DNA can be predicted as shown. Poly dG and poly dT tracts are not drawn in full.

FIGURE 30

1. ...CTGCAGGG RNA AACCCCTGCAG ...  
 ...GACGTC cDNA TTGGGACGTC ...

cut with Pst I

2. 5'-GGG RNA AACCCCTGCAG-3'  
 ACGTCC cDNA TTGGG

clone into Pst I cut M13mp8 RF

3. 5'-GACCTGCAGGG RNA AACCCCTGCAGGCCA  
CTGGACGTC cDNA TTGGGACGTCGGT

OR

5'-GACCTGCAGGGTT cDNA CCCTGCAGGCCA  
CTGGACGTCCA RNA GGGACGTCGGT

sequence

4. M13...GGGTT .... from an RNA-like template, 3' end of mRNA sequenced first.

M13...GGG..... from a cDNA-like template, 5' end of mRNA sequenced first.

Figures 31 and 32

DNA sequences of 5' and 3' regions of *X. borealis* tadpole globin  
cDNA clones 1.43 (figure 31) and 2.50 (figure 32)

DNA sequencing was as described in materials and methods. Sequences were determined from two separate experiments, but were not confirmed by sequencing the opposite DNA strand. Amino acid residues, predicted from nucleotide sequence, are shown above coding regions. Ambiguities in sequence are shown by bracketing alternatives, a question mark indicates that the alternative is less likely. Conserved amino acids (Dayhoff, 1976) are underlined (—), residues common to *X. laevis* adult and *X. borealis* tadpole  $\alpha$ - or  $\beta$ -globins are underlined (----). Restriction enzyme sites shown in figure 19 and Table 6 are indicated.

Figure 32. In common with most other  $\beta$ -globins (Dayhoff, 1976) *X. borealis* tadpole  $\beta$ -globin possesses one more amino acid residue at the N-terminus than does *X. laevis* adult  $\beta$ -globin (Williams et al, 1980).

FIGURE 31

G<sub>19</sub> A T T T G T C T T A A G T T T C C A C T T G A A G T C A A T  
10 20 30

C T A A A G C C T A C T A G C T C C C T C T T G T C A A G A T G/  
40 50 60

thr<sup>1</sup> leu thr asp ser asp lys ala ala ile<sup>10</sup> val  
A C T / C T G / A C C / G A C / A G T / G A T / A A G / G C T / G C A / A T T / G T T /  
70 80 90

ala leu trp gly lys ile ala pro gln<sup>20</sup> ala asn  
G C T / C T G / T G G / G G C / A A A / A T C / G C A / C C C / C A A / G C T / A A T /  
100 110 120

ala ile gly ala glu ala leu glu<sup>30</sup> arg leu phe  
G C C / A T T / G G A / G C C / G A G / G C T / T T G / G A C / A G A / C T T / T T C /  
130 140 150 160

leu ser tyr pro gln thr lys<sup>40</sup> thr tyr phe ser  
C T C / T C C / T A T / C C C / C A G / A C C / A A G / A C T / T A C / T T C / A G C /  
170 180 190

his phe asp val ser his<sup>50</sup> gly ser  
C A C / T T T / G A T / G T G / T C C / C A T / G G C / T C T / .....  
200 210

..... leu ala ser lys tyr arg STOP  
C T A / G C C / T C C / A A G / T A C / A G A / T A A / A G A A A G A A G G  
460 470 480

C A T C A G C A A G G G C A G C T C A C C A G C C T G G G G (A) A G  
490 500 510 520 (G)

A T C T A T G G A A A C C A T G C A C A A A A A T G C C A C T T T A  
G? 530 540 550

A C A A T G G T C A T G T T T A G A C C A A T A T G A A A C T G T  
560 570 580

A A C A A A T A T C A G C T G A G G T T A G A A G G T G T T A A C A  
590 600 Pvu II 610 620

T T T G A C A T G G T T T A T A G C C A G A A G C T G G T T A A T A T T  
630 640 650

G T T G (G) (A) A A A A A T C N (A) (A) (A) N<sub>2</sub> C<sub>12</sub>  
660 (C) (G) (G?) 670 (G) (G) (G) 680 (T) (T) (T)



FIGURE 32

$G_{21}$  A A (A) A G A G G C C T T T (T)(T)(T)(T) T (T) T (A) T (A) A (A)  
 1 (C) 10 (C) (C) (C) 20 (C) (C) (C)

A A T T (G) A G G (T)(C) T A A G C A G (C) A A T T C C C A T (T) T A  
 30 (A) (C) 40 50

(A) A G T (A) A C C A T G G T C / C A C / C T G / T C A / G C T / G A T / G A G /  
 (C) 60 70 PvuII A 80

lys ser ala<sup>10\*</sup> ile asn ala val trp ser / pro / gln arg / lys  
 A A A / T C C / G C C / A T T / A A C / G C C / G T (G) / T G G / (T) (C) A / A (G) G /  
 90 100 110

val ser val<sup>20</sup> glu asn asp gly his asp ala leu  
 G T (C) / A G C / G T T / G A A / A T / G A T / G G C / C A T / G A T / G C C / C T G /  
 120 (A) 130 140 150

thr arg<sup>30</sup> leu leu val val tyr pro trp thr gln  
 A C C / A G G / C T G / C T G / G T T / G T G / T A C / C C C / T G G / A C C / C A G /  
 160 170 180

arg<sup>40</sup> tyr phe ser thr phe gly asn leu ser asn / lys<sup>50</sup>  
 A G G / T A T / T T C / A G C / A C C / T T T / G G A / A A C / C T T / T C C / A A .....  
 190 200 210

ser leu leu ser pro gly cys leu gly<sup>130</sup> lys  
 ..... T / T C A / C T C / C T A / A G T / C C A / G G C / T G C / C T G / G G A / A A A /  
 430 440 450

ile arg ala leu gly trp leu ala ala<sup>140</sup> leu ser / arg  
 A T T / C G A / G C G / C T G / G G C / T G G / T T G / G C T / G C T / C T T / A G (C) /  
 Taq I 460 470 480 490

his gly tyr phe STOP  
 C A T / G G C / T A C / T T C / T A A A G G G T T T T G G T C C C A T G T A  
 500 510 520 C? T?

G A G C C C A A (C) A T C T G C T C G C C T G T G (G) C (C) A A G (A)  
 530 (A) 540 550 (A) (A) (G)

C A A C T C T C A C A A T A (A) G T G C A T T T C (C) T A T T G (G) C  
 560 (G) 570 580 (T) (T)

A A T (A) A A G G T T T C T T G T C A A A C T A<sub>12</sub> C<sub>2</sub> N<sub>6</sub> C<sub>14</sub>  
 (T) 590 600 610

sequence data from a X. laevis adult  $\alpha$ -globin cDNA clone (Partington and Baralle, 1981) shows additional amino acids to be conserved between X. laevis adult and X. borealis tadpole  $\alpha$ -globins (dotted underlined figure 31).

The 3' non coding region of 1.43 is approximately 190 nucleotides long and does not appear to contain the proposed polyadenylation signal AAUAAA found in most eukaryote messages (Proudfoot and Brownlee, 1976). The nearest related sequence to this (double underlined position 590-595, figure 31) is AACAAA but this is almost 80 bases, compared to 15-20 in most mRNAs, away from the start of the poly A. The best sequence in the expected region is AATATT at position 654-659 (figure 31).

2.50 cDNA sequence has a 5' non-coding region of 63 nucleotides (figure 32). There was considerable laddering in all tracks resulting in many ambiguities in the sequence of the 5' non-coding region, as shown in figures 29 and 32.

2.50 coding sequence starts at the first ATG codon and homology with other known  $\beta$ -globins (Dayhoff, 1976) shows this clone codes for a  $\beta$ -globin polypeptide. Compared to X. laevis adult  $\beta$ -globin (Williams et al, 1980) there is one more amino acid at the N-terminus (at some position prior to that indicated \* figure 32) and the X. borealis tadpole  $\beta$ -globin sequence is therefore similar to those from most other species (Dayhoff, 1976). Some amino acids which are normally conserved in  $\beta$ -globins have been changed in the tadpole sequence, notably at amino-acid positions 26 (glu - asp), 122 (phe - ser), 127 (gln - gly), 137 (val - trp) and 146 (his - phe). However,

the almost complete homology between positions 27 and 48 leaves no doubt that this clone codes for a  $\beta$ -globin.

The 5' non-translated region of clone 2.50 is 104 nucleotides long, excluding the poly A tail. The conserved AATAAA sequence is found 16 nucleotides upstream from the start of the poly A (double underlined figure 32).

### 6.3 DISCUSSION

DNA sequencing provides direct evidence that cDNA clones made from X. borealis tadpole blood cell 9S RNA contain sequences coding for globin polypeptides. The amino acid sequences of 22 other species (Dayhoff, 1976) and DNA sequences of X. laevis adult  $\alpha$ - and  $\beta$ -globin cDNA clones (Partington and Baralle, 1981, Williams et al, 1980) are known. Comparison of nucleotide and predicted amino acid sequences from two Xbt9S cDNA clones confirms that an  $\alpha$ - and a  $\beta$ -globin-coding clone have been isolated (1.43 and 2.50 respectively). Other clones falling into the same restriction classes (Table 6) presumably also code for the same  $\alpha$ - and  $\beta$ -globins.

DNA sequences obtained were ambiguous in several positions and none have been confirmed by sequencing the opposite strand. Amino acids predicted from coding regions sequenced correspond on the whole with conserved amino acids found in globins from other species (Dayhoff, 1976). It therefore seems likely that these sequences are correct. There is no comparable way of determining whether the non-translated regions have been sequenced correctly and it is therefore

important that they should be checked.

Ambiguities in the DNA sequence are thought to arise because DNA polymerase is unable to copy faithfully homopolymer stretches of ss DNA. Thus at junctions between poly G tails and cDNA sequences (figure 30) there were always very prominent bands across all four tracks (figure 29). For the next few nucleotides it was then difficult to determine the sequence. This problem should be totally overcome by subcloning suitable cDNA restriction fragments into M13mp8 and mp9. Sequencing would not then proceed through homopolymer tracts and by using both M13mp8 and mp9 restriction fragments are automatically cloned in both orientations. (M13mp8 (figure 27) and M13mp9 have the same oligonucleotide containing unique restriction sites inserted in either orientation into the lac Z gene.) Subcloning smaller fragments would also enable centre regions of the cDNA to be sequenced.

The 5' untranslated regions of 1.43 and 2.50 cDNA sequences are 59 and 63 nucleotides long respectively. Since the clones were constructed by priming cDNA synthesis at the 3' poly A of the mRNA the 5' untranslated regions may not be complete. S1 nuclease was not used during the cloning procedure, cDNA-RNA hybrids were annealed directly to pAT 153 minimising the loss of 5' sequences. RNA gel blot hybridisations (page 63) suggest that mRNAs complementary to 1.43 and 2.50 cDNAs are  $848 \pm 66$  nucleotides and  $680 \pm 56$  nucleotides long respectively, compared to 720 and 640 nucleotide cDNA sequences and these clones are therefore likely to be almost full length.

The coding regions of both 1.43 and 2.50 are deficient in C-G

doublets (Table 12), as is typical of most eukaryote mRNAs (Russel et al, 1976). Out of approximately 130 codons only two have the form CGX or XCG.

In general agreement with data from other systems codon selection in X. borealis tadpole globin mRNAs appears to be non-random, although obviously the numbers involved here are too small for results to be statistically significant. Kafatos et al (1977) noted a set of 16 strongly favoured and 8 more or less avoided codons in human and rabbit  $\beta$ -globin mRNAs. When the comparison is extended to include chicken  $\beta$ -globin cDNA (Richards et al, 1979), X. laevis adult  $\beta$ -globin cDNA (Williams et al, 1980) and the limited data available for X. borealis tadpole  $\beta$ -globin cDNA (Table 12) only four codons are similarly selected in all five species. Codons for X. borealis tadpole  $\alpha$ - and  $\beta$ -globins appear to be similarly selected. There is evidence to suggest that non-random selection of codons is genome-specific and dependent on relative frequencies of iso-acceptor tRNA species (Grantham et al, 1980).

Several amino acids conserved in most other species (Dayhoff, 1976) have been changed in the X. borealis  $\beta$ -globin cDNA sequence presented here. Of these substitutions the most interesting is the replacement of the C-terminal histidine residue by phenylalanine. In mammalian haemoglobins this residue is responsible for a major part of the Bohr effect as explained below.

Mammalian haemoglobins exhibit what is known as the "alkaline Bohr effect". In the presence of high levels of  $H^+$  and  $CO_2$ , for instance in the capillaries of metabolically active tissues, oxygen

TABLE 12. Codon usage in *X. borealis* tadpole  $\alpha$ - and  $\beta$ -globin mRNAs.

		2nd				
		U	C	A	G	
U	phe 1+1		ser 0+1	tyr 1+1	cys 1+0	U
	phe 2+2 <sup>1</sup>		ser 2+3	tyr 2+2	cys 1+0	C
	leu		ser 3+0 <sup>1</sup>	stop	stop	A
	leu 1+1		ser	stop	trp 3+1	G
C	leu 2+1		pro	his 2+0	arg	U
	leu 2+1		pro 1+2	his 1+1	arg	C
	leu 1+2		pro 2+0 <sup>1</sup>	gln 1+1 <sup>1</sup>	arg 1+0	A
	leu 6+2		pro	gln 1+1	arg	G
1st						3rd
A	ile 2+2		thr 0+1	asn 2+1 <sup>1</sup>	ser 1+1	U
	ile 0+1		thr 3+2	asn 2+0	ser 3+1 <sup>1</sup>	C
	ile		thr	lys 2+1	arg 0+2	A
	met		thr	lys 1+3 <sup>1</sup>	arg 4+0 <sup>2</sup>	G
G	val 3+1		<u>ala 3+4</u>	asp 3+2 <sup>1</sup>	gly	U
	val 2+0 <sup>1</sup>		<u>ala 3+3</u>	asp 0+1	gly 4+2	C
	val 1+0 <sup>1</sup>		<u>ala 0+2</u>	glu 1+0	gly 2+1	A
	val 2+1 <sup>1</sup>		<u>ala 1+0</u>	glu 1+2	gly	G

The frequency of codon usage has been determined from the partial sequences of *X. borealis* tadpole  $\alpha$ - and  $\beta$ -globin cDNA clones shown in figures 31 $\frac{1}{2}$  and 32. Values for  $\alpha$ -globin sequences are given first, those for  $\beta$ -globin sequences second. In some cases there was ambiguity in the  $\alpha$ -globin sequence and all possible codons have been included here, 1 and 2 indicate one or two codons respectively were ambiguous. Underlined codons are those that are similarly selected in human, rabbit, chicken and *X. laevis* adult  $\beta$ -globin mRNAs (see text for references).

release from haemoglobin is promoted. The reciprocal effect occurs in the alveolar capillaries of lungs, where oxygen concentration is high. The Bohr effect arises because all vertebrate haemoglobins are in an equilibrium between the deoxy (or tense T) and the oxy (or relaxed R) structures. The T structure has a low affinity for oxygen and a high affinity for  $H^+$  and  $CO_2$ , the R structure has reversed affinities.

One of the amino acid residues involved in the Bohr effect is the  $\beta$  chain C-terminal histidine, which is highly conserved (Dayhoff, 1976). The imidazole group of this histidine forms a salt bridge with Asp or Glu (94) or Glu (90) on the same  $\beta$  chain, in the T structure of all haemoglobins that exhibit a Bohr effect (Perutz and Brunori, 1982). The pK of the histidine is thus raised and protons are bound. In the R structure the carboxyl group of the C-terminal histidine forms a salt bridge with Lys (144) of the same  $\beta$  chain and its imidazole group is free.

In Xenopus laevis, which exhibits a weak Bohr effect (Perutz and Brunori, 1982) Lys (144) of the adult major  $\beta$  globin chain is replaced by glycine (Williams et al, 1980) which cannot bind the carboxyl group of the C-terminal histidine. This reduces the energy needed to move the C-terminal histidine towards Asp 94 on going from the R to T structure, and hence oxygen can be released more readily.

Haemoglobins of Rana catesbeiana tadpoles show a zero or reversed Bohr effect in the absence of organic phosphates, under conditions where mammalian haemoglobins exhibit an alkaline Bohr effect (Watt and Riggs, 1975). Substitutions within the globin

chains whilst maintaining the conserved  $\beta$  chain C-terminal histidine are also responsible for this, as is the case in all amphibian haemoglobins whose amino acid sequences are known (Perutz and Brunori, 1982).

In *X. borealis* tadpole haemoglobin it appears that the Bohr effect has been reduced by an alternative mechanism. Here the  $\beta$  chain C-terminal histidine has been replaced by phenylalanine which has no side chain and can not therefore participate in the Bohr effect.

Substitution of phenylalanine for the C-terminal histidine in the  $\beta$  chains of *X. laevis* tadpole globins (J. Williams in press) and in two fish species (unpublished J. Williams personal communication) has also been noted.

The 3' untranslated region of 1.43 appears not to contain the conserved AAUAAA sequence which is found in most other eukaryote mRNAs (Proudfoot and Brownlee, 1976) upstream from the polyadenylation site. Its precise location ranges from 11 nucleotides (silk fibroin mRNA, Hagenbuchle et al, 1979) to 30 nucleotides (adenovirus type 2 IVa mRNA, Alestrom et al, 1980) upstream, and it is thought to be involved with splicing primary transcripts and poly A addition. Work with SV40 late mRNAs (Fitzgerald and Schenk, 1981) suggests AAUAAA is essential for poly A addition, but that other sequences located between it and the coding region are also required. Some eukaryote mRNAs not containing this AAUAAA sequence are, however, polyadenylated e.g. chicken lysozyme mRNA (Jung et al, 1980) and mouse pancreatic  $\alpha$ -amylase mRNA (Hagenbuchle et al, 1980) which



substitute an AUUAAA sequence. Histone genes do not appear to contain any related sequence, but most histone mRNAs are not normal poly A plus messages (Hentschel and Birnstiel, 1981). A cDNA sequence made from Xenopus oocyte poly A plus H4 RNA does not contain any related sequence (Zernik et al, 1980). Yeast histone mRNA is normally polyadenylated and has been shown to contain this sequence (M. Smith unpublished, reported in Hentschel and Birnstiel, 1981), but the 3' non-coding regions of yeast are very A-T rich (Zaret and Sherman, 1982) and chance alone would frequently give rise to AAUAAA or related sequences. It is therefore not clear whether all mRNAs are polyadenylated by the same mechanism.

The lack of an AAUAAA sequence in 1.43 and related clones should be confirmed by sequencing independently isolated clones from the same class as 1.43, since the A - T transitions at positions 658 and 659 (figure 31) may have been introduced during cloning. It is unlikely that a sequencing error has been made since there was no difficulty reading sequences from relevant areas of the autoradiographs.

CHAPTER SEVEN

DISCUSSION

## 7.1 X. borealis tadpole globins

Globins isolated from RBCs of stage 57 or 58 X. borealis tadpoles can be separated into three major and five minor species with no evidence of individual polymorphism on triton-acid-urea polyacrylamide gels (figures 1 and 2). In contrast X. laevis tadpole globins show polymorphism of at least one component and consist of four electrophoretically separable species that do not comigrate with X. borealis tadpole globins (Hentschel et al, 1979; Hosbach et al, 1982). Some differences in mobility of X. borealis tadpole globins could be due to amino acid modifications, for example N-terminal acetylation is known to occur in  $\beta$ -globin chains of Rana catesbeiana tadpoles (Watt and Riggs, 1975). X. laevis adult globins, consisting of two major and four minor species (Hentschel et al, 1979) have been assigned to at least four different genes (Hentschel et al, 1979; Kay et al, 1980) and it is probable that all six species are encoded by separate genes. Four different cDNA clones have been isolated from X. laevis tadpole globin mRNA (Widmer et al, 1981) indicating that there is one gene for each of the four tadpole globins seen on polyacrylamide gels. By analogy it seems most likely that there are globin genes coding for each of the X. borealis tadpole globin species. This is supported by the isolation of four different types of cDNA clone from X. borealis tadpole globin mRNA (Table 6) which were shown to code for proteins in the three major bands and a minor component in the fastest migrating group of bands (e.g. figures 1 and 2, see Chapter 5).

Proteins from X. borealis tadpole RBCs separated on polyacrylamide gels have been identified as  $\alpha$ - or  $\beta$ -like globins by

hybridisation of cloned  $\alpha$ - and  $\beta$ -globin cDNA probes to Xbt9S poly A plus mRNA and translation of the bound RNA in rabbit reticulocyte lysate (hybrid-selection, see Chapter 5). The two slowest migrating groups (labelled 1 and 2 figures 1 and 2) are  $\alpha$ -globins and the fastest migrating bands (labelled 3, figures 1 and 2) are  $\beta$ -globins. This is the opposite arrangement to that in X. laevis where tadpole  $\beta$ -globins migrate more slowly than  $\alpha$ -globins (J. Williams, personal communication) however, triton-X 100 was not included in polyacrylamide gels of X. laevis globins and therefore differential binding of triton to  $\alpha$ - and  $\beta$ -globins could perhaps explain the observed variation in migration. Hosbach et al (1982) show X. laevis tadpole globins separated on triton-acid-urea polyacrylamide gels, but they do not say which bands correspond to  $\alpha$ - and which to  $\beta$ -globins.

Results presented here suggest that X. borealis tadpoles possess two major  $\alpha$ - globin species and only one major  $\beta$ -globin species. Analysis of cDNA clones has identified a second type of  $\beta$ -globin sequence but hybrid-selection experiments suggest that it is only rarely represented in Xbt9S poly A plus RNA (page 103 figure 21). Globin chains associate as  $\alpha_2 \beta_2$  tetramers in haemoglobin of higher vertebrates and the ratio of  $\alpha$ - and  $\beta$ -like globins in blood is known to be critical, any imbalance resulting in diseases known as thalassemys (for reviews of human haemoglobins see Weatherall and Clegg, 1979; Maniatis et al, 1980). It is therefore interesting to note that only one major  $\beta$ -globin band can be seen in globins isolated from tadpole RBCs whereas there are two prominent  $\alpha$ -globin bands (figure 21). cDNA sequences coding for some minor components have yet to be identified and it is therefore possible that there are

several minor  $\beta$ -globins in X. borealis tadpole blood which together compensate for the additional major  $\alpha$ -globin species.

Electrophoresis of Xbt9S poly A plus RNA on formaldehyde-agarose gels and hybridisation to cloned  $\alpha$ - or  $\beta$ -tadpole-globin cDNA probes shows the RNA to consist of two major components, of approximately equal distribution (figures 12, 25 and 26), the larger ( $848 \pm 68$  nucleotides) being  $\alpha$ - and the smaller one ( $680 \pm 56$  nucleotides)  $\beta$ -globin mRNA.

In some vertebrates there are two switches in haemoglobin type during development; for example humans synthesise embryonic haemoglobins up to 8 weeks of development, these are then replaced by foetal haemoglobin which contains the adult  $\alpha$ -globin chains and foetal  $\beta$ -like chains. Just prior to birth adult  $\beta$ - and  $\delta$ -globins are synthesised and by 6 months after birth adult haemoglobins ( $\alpha_2 \beta_2$  and  $\alpha_2 \delta_2$ ) comprise about 99% of the total haemoglobin (for reviews see Weatherall and Clegg, 1979; Maniatis et al, 1980). Since globins were isolated from stage 57-59 X. borealis tadpoles which are already undergoing metamorphosis it remains possible that very early tadpoles synthesise different globins. This has been investigated in X. laevis (J. Williams, personal communication) and it was found that the same gene which was active at stage 45 was also expressed in very early tadpoles. No such investigation was undertaken with X. borealis tadpoles.

## 7.2 Onset of X. borealis adult globin synthesis

The onset of adult globin gene expression during metamorphosis

of X. borealis tadpoles was investigated in order to determine the most suitable stage from which to isolate tadpole globin mRNA. Early work, using immunofluorescence and column chromatography to characterise haemoglobins suggested that the transition from tadpole to adult haemoglobin synthesis takes place after completion of metamorphosis (Jurd and Maclean, 1969, 1970, 1974; Maclean and Jurd, 1971a). Hentschel et al (1979) could first detect adult globins, by staining polyacrylamide gels, in blood of X. laevis frogs three to four weeks after metamorphosis. Using antisera specific for X. laevis tadpole or adult globins Just et al (1980) showed that very little adult haemoglobin is present in tadpole blood cells up to stage 65, after which there is a sharp increase coincident with a rapid decrease in the amount of tadpole haemoglobin. By analysing incorporation of [<sup>3</sup>H]-amino-acids in vivo into haemoglobins Just et al (1977) were able to detect an increase in adult and corresponding decrease in larval haemoglobin synthesis beginning around stage 58.

Tadpole RBCs in culture synthesise haemoglobin (shown by Moss and Ingram, 1965 for Rana catesbeiana; Maclean et al, 1969 for X. laevis) and this was used as an assay to detect adult globin synthesis in X. borealis tadpole RBCs. In agreement with the data of Just et al (1977) X. borealis adult globin synthesis was first detectable at about stage 57-58 and thereafter tadpole globin synthesis decreased and adult globin synthesis increased (Chapter 3, figures 5, 6 and 7).

Larval RBCs of Rana pipiens, Rana catesbeiana, and X. laevis are morphologically distinct from their adult counterparts

(Hollyfield, 1966a; McCutcheon, 1936; Jurd and Maclean, 1970). Nuclei of Rana pipiens tadpole RBCs are round, whereas those of adult RBCs are elliptical in shape (Hollyfield, 1966a; Benbassat, 1970) and the cytoplasm of tadpole RBCs contains a larger amount of reticular substance (Benbassat, 1970). These differences suggest that tadpole RBCs are more immature than adult RBCs and synthesise more RNA and protein. This has been demonstrated more directly by labelling blood cells from X. laevis or Rana catesbeiana tadpoles in culture with radioactive amino acids (Maclean et al, 1969; Moss and Ingram, 1965). Thus, unlike the situation in adult X. borealis where globin mRNA can only be isolated from blood of anaemic animals, in tadpoles it can be isolated directly from circulating RBCs. This is advantageous since there are conflicting reports regarding the effects of anaemia on globin gene expression. In chicken (Richards and Wells, 1980; Salser et al, 1979) and sheep (Wood et al, 1979) new globin species can be detected during recovery from anaemia. When anaemia is induced in adult Rana pipiens, Rana catesbeiana (Meints and Forehand, 1977) or X. laevis (Maclean and Jurd, 1971a) there is a partial reversion to tadpole haemoglobin synthesis. In all three species there is also a marked increase in the synthesis of one of the adult components (Meints and Forehand, 1977; Maclean and Jurd, 1971; reviewed in Broyles, 1981). Hentschel et al (1979) reported the appearance of a novel globin in immature erythroid cells of adult X. laevis produced during recovery from severe anaemia. However no change in the pattern of globin synthesis by anaemic adult X. borealis RBCs has been detected. Since it is possible to isolate RNA from physiologically normal tadpoles, the problem of deciding whether globin sequences detected are expressed during normal tadpole development does not arise.

Tadpole globin synthesis was shown to predominate at stages 57-59 (Chapter 3, figure 6) and these stages were therefore used in the preparation of blood cell 9S RNA. The choice of stage to use is a compromise between using earlier animals in which adult globin synthesis is not detectable and using later animals which are larger and therefore yield more RNA per animal but which also synthesise detectable quantities of adult globins. Perlman et al (1977) showed by hybridisation of adult X. laevis globin cDNA to tadpole globin mRNA that anaemic X. laevis tadpole blood cell RNA contains adult globin mRNA. Just et al (1977) concluded from incorporation of [<sup>3</sup>H]-amino acids in vivo into haemoglobins that all X. laevis larval stages synthesis small amounts of adult haemoglobin.

RNA gel blot hybridisation of a X. borealis adult globin cDNA probe to Xbt9S poly A plus RNA has shown that RNA isolated from stage 57-59 X. borealis tadpole RBCs does indeed contain a small proportion of adult globin mRNA (figure 26). The actual amount was not determined, but since the specific activities of tadpole and adult probes were approximately equal and the autoradiograph of the adult hybridisation had to be exposed for almost ten times longer than those of tadpole hybridisations (figures 25 and 26) it is clear that there is only a small proportion of adult globin mRNA in Xbt9S RNA.

### 7.3 Comparison of Xbt9S poly A plus and poly A minus RNA

9S RNA isolated from stage 57-59 X. borealis tadpoles was shown, by in vitro translation in rabbit reticulocyte lysate, to code for all proteins normally isolated from tadpole RBCs (Chapter 3, figure 13). Comparison of translation products from Xbt9S poly A



plus and poly A minus RNA (figure 13) showed that both RNA populations contain mRNA coding for all the different globin species, but the poly A plus fraction contains about 10 to 40 times more coding RNA than the poly A minus fraction. This is supported by dot blot experiments described in Chapter 5 and figures 23 and 24. Poly A plus RNA hybridised approximately seven times as much  $\beta$ -globin cDNA as poly A minus RNA (Table 9, figures 23 and 24). A similar result was also obtained when  $\alpha$ -globin cDNA was hybridised to poly A plus and poly A minus RNA. RNA gel blot hybridisations with cloned  $\alpha$ - and  $\beta$ -globin cDNA probes of Xbt9S poly A plus and poly A minus RNA separated on formaldehyde-agarose gels and transferred to nitrocellulose also demonstrate that most X. borealis tadpole globin message is polyadenylated (Chapter 5, figures 25 and 26).

#### 7.4 Preparation and analysis of X. borealis tadpole globin cDNA clones

cDNA clones were prepared from Xbt9S poly A plus RNA as described in Chapter 4 (figure 15). Since the RNA was expected to contain some adult sequences, as discussed above, clones were screened by colony hybridisation with X. borealis adult anaemic blood cell 9S poly A plus RNA. By analogy to X. laevis (Perlman et al, 1977) this RNA was assumed not to cross hybridise with tadpole globin sequences and therefore any clone hybridising probably codes for adult globins. Approximately 10% (Table 4) of the cDNA clones analysed hybridised to adult globin mRNA, in agreement with the data from RNA gel blot hybridisations (figure 26) which demonstrate that there is a low level of adult globin message in Xbt9S poly A plus RNA.

Colony hybridisation with adult 9S mRNA allowed clones containing adult globin sequences to be eliminated. Further colony hybridisations with Xbt9S poly A plus RNA (Chapter 4, figure 16) identified those containing tadpole specific sequences. As discussed in Chapter 4 it was hoped that signal intensity and cDNA insert size would be directly related. Figure 18 and Table 5 show this relationship, but although in general clones with larger inserts do give more intense signals it would not be advisable to select clones for further study on this basis, since those containing rarer sequences are likely to give weak signals regardless of insert length.

Clones which hybridised to Xbt9S poly A plus RNA to any extent were therefore randomly chosen for further analysis. Insert sizes were determined by Pst I digestion (figure 17) and 18 clones containing cDNA sequences over 600 nucleotides in length were selected. Limited restriction analysis (figure 19, Table 6) sorted the clones into four groups which, by comparison with the restriction maps of X. laevis tadpole globin cDNA sequences (Widmer et al, 1981) could be tentatively identified as  $\alpha$ - or  $\beta$ -globin cDNA sequences. This was confirmed for the two largest groups by partially sequencing a member from each class (figure 31 and 32, Chapter 6) and comparing the predicted amino acid sequences with those known for other vertebrate globins (Dayhoff, 1976). Since the restriction analysis was very limited there may be some differences between members of each class (Table 6) which have not been detected.

DNA sequences of one  $\alpha$ - and one  $\beta$ -globin cDNA clone are presented and discussed in Chapter 6. The sequences are designated

as  $\alpha$ - or  $\beta$ -globin coding by comparison of conserved amino acids in other vertebrates (Dayhoff, 1976). As expected (Kozak, 1978; Baralle and Brownlee, 1978) translation starts at the first available AUG codon assuming there are none in the 5' sequence which has not been cloned. It would be possible to estimate the length of 5' sequences missing from the cDNA by primer extension studies. However, the cDNA sequences are likely to be almost full length copies of the mRNA for several reasons: (1) X. borealis tadpole  $\alpha$ - and  $\beta$ -globin mRNAs have been estimated to be  $848 \pm 68$  and  $680 \pm 56$  nucleotides long respectively (figures 12, 25; Chapter 3). 1.43 ( $\alpha$ ) and 2.50 ( $\beta$ ) cDNA sequences are approximately 720 and 640 nucleotides long as estimated from agarose gel electrophoresis of Pst I digested plasmids; (2) Assuming X. borealis tadpole  $\alpha$ - and  $\beta$ -globins, in common with most other species, consist of 141 and 146 amino acid residues respectively the DNA sequencing studies (figures 31 and 32) predict that the  $\alpha$ -globin sequence is 680 and the  $\beta$ -globin sequence 610 nucleotides long; (3) the 5' non-coding regions of human and rabbit  $\beta$ -globin mRNA are 56 nucleotides long (Baralle, 1977; Lockard and Raj-Bhandary, 1976) and the X. borealis tadpole  $\beta$ -globin clone sequenced here (figure 32) has 63 nucleotides of 5' untranslated sequence.

In common with other amphibian haemoglobins (Perutz and Brunori, 1982) the Bohr effect in X. borealis tadpole haemoglobin appears to have been reduced, facilitating dissociation of oxygen from oxyhaemoglobin. This is achieved in X. borealis tadpole haemoglobin by substitution of phenylalanine for the C-terminal histidine on the  $\beta$ -globin chain, in contrast to other amphibians where substitutions of internal  $\beta$ -chain residues prevent interactions

with the C-terminal histidine (Perutz and Brunori, 1982).

## 7.5 Future experiments

The hybrid-selection experiments described in Chapter 5 suggest that a major class of clones might not yet have been identified. This could be tested by further hybrid-selection experiments as discussed in section 5.2 and if it is the case restriction analysis of other clones should yield the missing class. Several minor classes are apparently not represented in the 18 clones so far analysed, since they fall into four classes with respect to restriction enzyme sites and 3 major and 5 minor globin species are detectable on triton-acid-urea polyacrylamide gels. It is possible that the classes shown in Table 6 would be subdivided by further restriction analysis, alternatively other different clones so far not detected might exist. It has been assumed, by analogy to X. laevis (Henschel et al, 1979; Widmer et al, 1981) that each globin species seen on polyacrylamide gels is encoded by a different gene and that amino acid modifications are not responsible for the mobility differences.

The tadpole globin cDNA clones could be used as probes for the expression of globin genes during very early development of X. borealis. Total RNA from tadpoles could be spotted onto nitrocellulose or transferred to nitrocellulose from formaldehyde-agarose gels and hybridised to  $\alpha$ - and  $\beta$ -globin cDNA probes to detect early expression. It is possible that different globin genes might be expressed during very early development, this occurs for example in humans, and the cDNA probes made from stage 57-59 tadpole blood

cell RNA might not therefore cross hybridise to the first globin sequences expressed. Electrophoretic analysis of globins from very early tadpoles, obtained by tail amputation and collection of blood into a solution of Barth X and heparin might resolve this problem, although a large number of animals would be required to provide sufficient sample.

A X. borealis genomic DNA library has been prepared (N. Hunter unpublished) and the globin cDNA clones could be used to isolate corresponding genomic globin sequences. The organisation of X. borealis globin genes could then be examined and compared to X. laevis (see section 1.2.c, Patient et al, 1980). Analysis of conserved sequences within non-coding regions of globin gene clusters from Xenopus and mammalian and avian systems might then provide information about sequences important for regulation of globin gene expression during development.

## REFERENCES

- Aggarwal, S.J. and Riggs, A. (1969), *J. Biol. Chem.* 244, 2372-2383
- Aleström, P., Akusjärvi, G., Perricaudet, M., Mathews, M.B.,  
Klessing, D.F. and Pettersson, U. (1980), *Cell* 19, 671-681
- Alfageme, C.R., Zweidler, A., Mahowald, A. and Cohen, L.H. (1974),  
*J. Biol. Chem.* 249, 3729-3736
- Baglioni, C. and Sparks, C.E. (1963), *Develop. Biol.* 8, 272-285
- Baralle, F. (1977), *Cell* 10, 549-558
- Baralle, F.E. and Brownlee, G.G. (1978), *Nature* 274, 84-87
- Barth, L.G. and Barth, L.G. (1959), *J. Embryol. exp. Morph.* 7,  
210-222
- Battaglia, P. and Melli, M. (1977), *Develop. Biol.* 60, 337-350
- Benbassat, J. (1970), *Develop. Biol.* 21, 557-583
- Benbassat, J. (1974), *J. Cell Sci.* 16, 143-156
- Benton, W.D. and Davis, R.W. (1977), *Science* 196, 180-182
- Berger, S.L. and Birkenmeier, C.S. (1979) *Biochemistry* 18, 5143-5149
- Birnboim, H.C. and Doly, J. (1979), *Nucl. Acids Res.* 7, 1513-1523
- Bisbee, C.A., Baker, M.A., Wilson, A.C., Hadji-Azimi, I. and  
Fischberg, M. (1977), *Science* 195, 785-787
- Bishop, J.O., Beckmann, J.S., Campo, M.S., Hastie, N.D.,  
Izquierdo, M. and Perlman, S. (1975), *Phil. Trans. R. Soc.*  
*Lond. B.* 272, 147-157
- Bishop, J.O., Rosbash, M., Evans, D. (1974), *J. Mol. Biol.* 85, 75-86
- Bodmer, W.F. (1979), *CIBA Foundation Symp.* 66, 205-229
- Boyer, H.W. and Roulland-Dussoix, D. (1969), *J. Mol. Biol.* 41,  
459-479
- Broyles, R.H. (1981), pp 461-490 in: *Metamorphosis, a problem in  
Developmental Biology*. 2nd edition (L.I. Gilbert and  
E. Frieden, eds.) Plenum Press, New York and London.
- Broyles, R.H. and Deutsch, M.J. (1975), *Science* 190, 471-473
- Broyles, R.H. and Frieden, E. (1973), *Nature New Biol.* 241, 207-209
- Busslinger, M., Portmann, R. and Birnstiel, M.L. (1979), *Nucl. Acids  
Res.* 6, 2997-3008

- Cabada, M.O., Darnbrough, C., Ford, P.J. and Turner, P.C. (1977)  
*Develop. Biol.* 57, 427-439
- Cline, M.J. and Waldmann, T.A. (1962), *Amer. J. Physiol.* 203, 401-403
- Cole, R. and Tarbutt, R. (1973), p365 in: *The Cell Cycle in Development and Differentiation.* (Eds. Balls and Billet)  
 Cambridge University Press
- Dagert, M. and Ehrlich, S.D. (1979), *Gene* 6, 23-28
- Dayhoff, M.O. (1972), *Atlas of protein sequence and structure*,  
 Volume 5
- Dayhoff, M.O. (1976), *Atlas of protein sequence and structure*,  
 Volume 5, supplement 2
- de Graaf, A.R. (1957), *J. Exp. Biol.* 34, 173-176
- Donis-Keller, H., Maxam, A.M. and Gilbert, W. (1977), *Nucl. Acids Res.* 4, 2527-2538
- Dorn, A.R. and Broyles, R.H. (1982), *Proc. Natl. Acad. Sci. USA* 79,  
 5592-5596
- Efstratiadis, A., Kafatos, F.C., Maxam, A.M. and Maniatis, T. (1976),  
*Cell* 7, 279-288
- Efstratiadis, A., Posakony, J.W., Maniatis, T., Lawn, R.M.,  
 O'Connell, C., Spritz, R.A., DeRiel, J.K., Forget, B.G.,  
 Weissman, S.M., Slightom, J.L., Blechl, A.L., Smithies, O.,  
 Baralle, F.E., Shoulders, C.C. and Proudfoot, N.J. (1980),  
*Cell* 21, 653-668
- Ewer, D.W. (1959), *Nature* 183, 271
- Fey, H., Pfister, H., Messerli, J., Sturzenegger, N. and  
 Grolimund, F. (1976), *Zentralbl. Veterinaermed. Reihe B.* 23,  
 269-300
- Fitzgerald, M. and Shenk, T. (1981), *Cell* 24, 251-260
- Forman, L.J. and Just, J.J. (1976), *Develop. Biol.* 50, 537-540
- Foxon, G.E.H. (1964), pp151-209 in: *Physiology of the amphibia*  
 (J.A. Moore, ed.), Academic Press, New York
- Frieden, E. (1968), pp349-398 in: *Metamorphosis: A problem in Developmental Biology* (W. Etkin and L.I. Gilbert, eds.),  
 Appleton-Century-Crofts, New York
- Gluackson-Waelsch, S. (1960), p89 in: *Symposium on Mammalian Genetics and Reproduction.* USA Oak Ridge Nat. Library

- Grantham, R., Gautier, C., Gouy, M., Mercier, R. and Pave, A. (1980)  
Nucl. Acids Res. 8, 49-62
- Hagenbüchle, O., Bovey, R. and Young, R.A. (1980), Cell 21, 179-187
- Hagenbüchle, O., Krikeles, M.S. and Sprague, K.U. (1979), J. Biol.  
Chem. 254, 7157-7169
- Hamada, K., Sakai, Y., Shukuya, R. and Kaziro, K. (1964), J. Biochem.  
55, 636-642
- Hanahan, D. and Meselson, M. (1980), Gene 10, 63-67
- Hentschel, C.C. and Birnstiel, M.L. (1981), Cell 25, 301-313
- Hentschel, C.C., Kay, R.M. and Williams, J.G. (1979), Develop. Biol.  
72, 350-363
- Herner, A.E. and Frieden, E. (1961), Arch. Biochem. Biophys. 95,  
25-35
- Hollyfield, J.G. (1966a), J. Morphol. 119, 1-6
- Hollyfield, J.G. (1966b), Develop. Biol. 14, 461-480
- Hosbach, H.A., Widmer, H.J., Andrews, A-C. and Weber, R. (1982),  
pp115-125 in: Embryonic Development, Part A : Genetic Aspects.  
Alan R. Liss Inc.
- Humphries, P., Old, R., Coggins, L.W., MacShane, T., Watson, C. and  
Paul, J. (1978), Nucl. Acids Res. 5, 905-924
- Jahn, C.L., Hutchinson, C.A.III, Phillips, S.J., Weaver, S.,  
Haigwood, N.L., Voliva, C.F. and Edgell, M.H. (1980), Cell 21,  
159-168
- Jeffreys, A.J., Wilson, V., Wood, D., Simons, J.P., Kay, R.M. and  
Williams, J.G. (1980), Cell 21, 555-564
- Jordan, H.E. and Spiedel, C.C. (1923), J. Exp. Med. 38, 529-542
- Jung, A., Sippel, A.E., Grez, M. and Schütz, G. (1980), Proc. Natl.  
Acad. Sci. USA 77, 5759-5763
- Jurd, R.D. and Maclean, N. (1969), Experientia 25, 626-628
- Jurd, R.D. and Maclean, N. (1970), J. Embryol. exp. Morphol. 23,  
299-309
- Jurd, R.D. and Maclean, N. (1974), J. Microsc. 100, 213-217
- Just, J.J. and Atkinson, B.G. (1972), J. Exp. Zool. 182, 271-280
- Just, J.J., Schwager, J. and Weber, R. (1977), Wilhelm Roux's Arch.  
183, 307-323



- Just, J.J., Schwager, J., Weber, R., Fey, H. and Pfister, H. (1980),  
*Wilhelm Roux's Arch.* 188, 75-80
- Kafatos, F.C., Efstratiadis, A., Forget, B.G. and Weissman, S.M.  
 (1977), *Proc. Natl. Acad. Sci. USA* 74, 5618-5622
- Kay, R.M., Harris, R., Patient, R.K. and Williams, J.G. (1980), *Nucl.*  
*Acids Res.* 8, 2691-2707
- Kellet, G.L. (1971), *Nature New Biol.* 234, 189-191
- Kindle, K.L. and Firtel, R.A. (1978), *Cell* 15, 763-778
- Kistler, A. and Weber, R. (1975), *Mol. Cell. Endocrinol* , 261-288
- Kollros, J.J. (1961), *Am. Zoologist* 1, 107-114
- Kozak, M. (1978), *Cell* 15, 1109-1123
- Lacy, E., Hardison, R.C., Quon, D. and Maniatis, T. (1979), *Cell* 18,  
 1273-1283
- Laemmli, U.K. (1970), *Nature* 227, 680-685
- Laird, C.D. and McCarthy, B.J. (1968), *Genetics* 60, 303-312
- Laskey, R.A. and Mills, A.D. (1975), *Eur. J. Biochem.* 56, 335-341
- Leder, A., Swan, D., Ruddle, F., D'Eustachio, P. and Leder, P.  
 (1981), *Nature* 293, 196-200
- Lehrach, H., Diamond, D., Wozney, J.M. and Boedtke, H. (1977),  
*Biochemistry* 16, 4743-4751
- Lobban, P.E. and Kaiser, A.D. (1973), *J. Mol. Biol.* 78, 453-471
- Lockard, R.E. and Raj-Bhandary, U.L., (1976), *Cell* 9, 747-760
- Maclean, N., Brooks, G.T. and Jurd, R.D. (1969), *Comp. Biochem.*  
*Physiol.* 30, 825-834
- Maclean, N., Hilder, V.A., Baynes, Y.A. (1973), *Cell differentiation*  
2, 261-269
- Maclean, N. and Jurd, R.D. (1971a), *J. Cell Sci.* 9, 509-528
- Maclean, N. and Jurd, R.D. (1971b), *Comp. Biochem. Physiol.* 40B,  
 751-755
- Maclean, N. and Jurd, R.D. (1972), *Biol. Rev.* 47, 393-437
- Maclean, N. and Turner, S. (1976), *J. Embryol. exp. Morph.* 35,  
 261-266
- McCutcheon, F.H. (1936), *J. Cell. Comp. Physiol.* 8, 63-81

- Maniatis, G.M. and Ingram, V.M. (1971a), J. Cell Biol. 49, 372-379
- Maniatis, G.M. and Ingram, V.M. (1971b), J. Cell Biol. 49, 380-389
- Maniatis, G.M. and Ingram, V.M. (1971c), J. Cell Biol. 49, 390-404
- Maniatis, G.M., Steiner, L.A. and Ingram, V.M. (1969), Fed. Proc. 28,  
819
- Maniatis, T., Fritsch, E.F., Lauer, J. and Lawn, R.M. (1980), Ann.  
Rev. Genet. 14, 145-178
- Meints, R.H. and Forehand, C. (1977), Comp. Biochem. Physiol. 58A,  
265-268
- Messing, J., Crea, R. and Seeburg, P.H. (1981), Nucl. Acids Res. 9,  
309-321
- Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P.H.  
(1977), Proc. Natl. Acad. Sci. USA 74, 3642-3646
- Moss, B. and Ingram, V.M. (1965), Proc. Natl. Acad. Sci. USA 54,  
967-974
- Moss, B. and Ingram, V.M. (1968a), J. Mol. Biol. 32, 481-492
- Moss, B. and Ingram, V.M. (1968b), J. Mol. Biol. 32, 493-504
- Nieuwkoop, P.D. and Faber, J. (1956), Normal table of Xenopus laevis.  
(Daudin), North Holland, Amsterdam.
- Ouchterlony, O.E. (1958), Prog. Allergy 5, 1-78; 6, 30-154
- Panyim, S. and Chalkey, R. (1969), Arch. Biochem. Biophys. 130,  
337-346
- Partington, G.A. and Baralle, F.E. (1981), J. Mol. Biol. 145,  
463-470
- Paterson, B.M., Roberts, B.E. and Knuff, C. (1977), Proc. Natl. Acad.  
Sci. USA 73, 3414-3422
- Patient, R.K., Elkington, J.A., Kay, R.M. and Williams, J.G. (1980)  
Cell 21, 565-573
- Pelham, H.R.B. and Jackson, R.J. (1976), Eur. J. Biochem. 67,  
247-256
- Penman, S. (1966), J. Mol. Biol. 17, 117-130
- Perlman, S.M., Ford, P.J. and Rosbash, M.M. (1977), Proc. Natl. Acad.  
Sci. USA, 74, 3835-3839
- Perutz, M.F. and Brunori, M. (1982), Nature, 299, 421-426

- Proudfoot, N.J. and Brownlee, G.G. (1976), *Nature* 263, 211-214
- Radloff, R., Bauer, W. and Vinograd, J. (1967), *Proc. Natl. Acad. Sci. USA*, 57, 1514-1521
- Reichlin, M., Bucci, E., Wyman, J., Antonini, E. and Rossi-Fanelli, A. (1965), *J. Mol. Biol.* 11, 775-784
- Ricciardi, R.P., Miller, J.S., Roberts, B.E. (1979), *Proc. Natl. Acad. Sci. USA*, 76, 4927-4931
- Richards, R.I., Shine, J., Ullrich, A., Wells, J.R.E. and Goodman, H.M. (1979), *Nucl. Acids Res.* 7, 1137-1146
- Richards, R.I. and Wells, J.R.E. (1980), *J. Biol. Chem.* 255, 9306-9311
- Richardson, C., Capello, J., Cochran, M.D., Armentrout, R.W. and Brown, R.D. (1980), *Develop. Biol.* 78, 161-172
- Rigby, P.W., Dieckmann, M., Rhodes, C. and Berg, P. (1977), *J. Mol. Biol.* 113, 237-251
- Riggs, A. (1951), *J. Gen. Physiol.* 35, 23-40
- Roninson, I.B. and Ingram, V.M. (1982), *Cell* 28, 515-521
- Rosenberg, M. (1970), *Proc. Natl. Acad. Sci. USA*, 67, 32-36
- Roychoudhury, R., Jay, E. and Wu, R. (1976), *Nucl. Acids Res.* 3, 863-877
- Russel, G.J., Walker, P.M.B., Elton, R.A. and Subak-Sharpe, J.H. (1976), *J. Mol. Biol.* 108, 1-15
- Ruud, J.T. (1954), *Nature*, 173, 848-50
- Ryback, B. (1960), *Nature*, 185, 777-778
- Salser, W.A., Cunnings, I., Liu, A., Strommer, J., Padayatty, J. and Clarke, P. (1979), pp621-645 in: *Cellular and Molecular Regulation of Haemoglobin Switching*". (G. Stamatoyannopoulos and A.W. Nienhuis, eds.), Grune and Stratton, New York.
- Sanger, F. (1980), *DNA sequencing manual from the Cambridge course*, December 1980
- Sanger, F. and Coulson, A.R. (1978), *Febs. Lett.* 87, 107-110
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977), *Proc. Natl. Acad. Sci. USA*, 74, 5463-5467

- Saucier, W.J. (1976), Red blood cell differentiation in the bullfrog, Rana catesbeiana: The haemoglobin transition and erythropoietic sites during and following metamorphosis. M.S. thesis. University of Wisconsin, Milwaukee.
- Southern, E.M. (1979), Anal. Biochem. 100, 319-323
- Stratton, L.P. and Frieden, E. (1967), Nature 216, 932-934
- Taylor, A.C. and Kollros, J.J. (1946), Anat. Record 94, 7-23
- Thomas, N. and Maclean, N. (1975), J. Cell Sci. 19, 509-520
- Thomas, P.S. (1980), Proc. Natl. Acad. Sci. USA, 77, 5201-5205
- Tompkins, R. and Townsend, J.K. (1977), J. exp. Zool. 200, 191-196
- Wahli, W., Dawid, I.B., Wyler, T., Jaggi, R.B., Weber, R. and Ryffel, G.U. (1979), Cell 16, 535-549
- Watt, W.K. and Riggs, A. (1975), J. Biol. Chem. 250, 5934-5944
- Weatherall, D.J. and Clegg, J.B. (1979), Cell 16, 467-479
- Weaver, C.A., Gordon, D.F. and Kemper, B. (1981), Proc. Natl. Acad. Sci. USA, 78, 4073-4077
- Westley, B., Wyler, T., Ryffel, G. and Weber, R. (1981), Nucl. Acids Res. 9, 3557-3574
- Widmer, H.J., Andres, A-C., Niessing, J., Hosbach, H.A. and Weber, R. (1981), Develop. Biol. 88, 325-332
- Williams, J.G., Kay, R.M. and Patient, R.K. (1980), Nucl. Acids Res. 8, 4247-4258
- Wood, K.O. and Lee, J.C. (1976), Nucl. Acids Res. 3, 1961-1971
- Wood, W.G., Nash, J., Weatherall, D.J., Robinson, J.S. and Harrison, F.A. (1979), pp153-167 in: Cellular and Molecular Regulation of Haemoglobin Switching. (G. Stamatoyannopoulos and A.W. Nienhuis, eds.), Grune and Stratton, New York.
- Zain, S., Sambrook, J., Roberts, R.J., Keller, W., Fried, M., and Dunn, A.R. (1979), Cell 16, 851-861
- Zaret, K.S. and Sherman, F. (1982), Cell 28, 563-573
- Zernik, M., Heintz, N., Boime, I. and Roeder, R.G. (1980), Cell 22, 807-815

APPENDICES

## APPENDIX I

### Calculation of amount of cDNA synthesised

#### 1. cpm incorporated

1  $\mu$ l sample TCA precipitated from 250  $\mu$ l reaction = 500 cpm  
Background = 100 cpm  
Total cpm incorporated in reaction (500-100) x 250 =  $10^5$  cpm

#### 2. Specific activity [ $\alpha^{32}$ P]dCTP in reaction

20  $\mu$ l 20mM dCTP added to final volume 250  $\mu$ l

. . 400 nmoles dCTP in reaction

10 Ci [ $\alpha^{32}$ P]dCTP in reaction

. . specific activity is 10  $\mu$ Ci/400 nmoles (since dCTP from label is negligible).

$$= 0.025 \text{ Ci/nmole}$$

$$\equiv 5 \times 10^4 \text{ cpm/nmole, assuming}$$

$$1 \mu\text{Ci } [\alpha^{32}\text{P}]d\text{CTP} \equiv 2 \times 10^6 \text{ cpm}$$

#### 3. cDNA synthesis

$$\text{nmoles dC incorporated} = \frac{\text{cpm incorporated}}{\text{specific activity}}$$

$$= 10^5 / (5 \times 10^4) \text{ nmoles}$$

$$= 2 \text{ nmoles}$$

$$1 \text{ nmole dC} \equiv 340 \text{ ng}$$

$$\begin{aligned} & \cdot \\ & \cdot \cdot 2 \times 340 \text{ ng dC incorporated} \\ & = 680 \text{ ng} \end{aligned}$$

$\mu\text{g}$  cDNA synthesised

$$\begin{aligned} & = \text{ng dC incorporated} \times 10^{-3} / \text{mole \%dG in RNA} \\ & = (680 \times 10^{-3}) / (0.22) \\ & = 3.1 \mu\text{g} \end{aligned}$$

$$\% \text{ cDNA synthesis} = 31$$

APPENDIX II

A. Calculation of number of residues added to 3' termini in cDNA dC tailing reaction.

1. Number of 3' termini available for tailing

1 nmole double stranded DNA  $\equiv$  680 ng nucleotide

680 ng  $\equiv$   $(6 \times 10^{14}) / (650)$  molecules, where 650 is the estimated average length of globin mRNA.

$$\begin{aligned} \therefore .0.5\mu\text{g} &\equiv (6 \times 10^{14} / 650) \times (0.5 / 0.68) \text{ molecules} \\ &= 6.8 \times 10^{11} \text{ molecules} \\ &= 1.3 \times 10^{12} \text{ 3' termini} \end{aligned}$$

2. Specific activity of  $[\alpha^{32}\text{P}]\text{dCTP}$  in reaction

10  $\mu\text{Ci}$   $[\alpha^{32}\text{P}]\text{dCTP}$

2  $\mu\text{l}$  2mM dCTP added

= 4 nmoles

$\therefore$  specific activity = 10  $\mu\text{Ci}$  / 4 nmoles, since dCTP contributed by label is negligible.

$$= 2.5 \mu\text{Ci/nmole}$$

$$\equiv 5 \times 10^6 \text{ cpm/nmole,}$$

assuming 1  $\mu\text{Ci}$   $[\alpha^{32}\text{P}]\text{dCTP} \equiv 2 \times 10^6 \text{ cpm}$

3. cpm expected for 1 residue at each 3' terminus

$$(5 \times 10^6) / (6 \times 10^{14}) \times 1.3 \times 10^{12} = 10^4 \text{ cpm}$$



4. Actual number of residues added

1  $\mu$ l sample from 50  $\mu$ l reaction TCA precipitated

. .  $5.6 \times 10^5$  cpm incorporated

. .  $(5.6 \times 10^5 / 10^4)$  residues added = 56

B. Calculation of number of residues added to 3' termini in pAT 153 dG-tailing reaction

1. Number of 3' termini available for tailing

1 nmole double stranded DNA  $\equiv$  680 ng nucleotide

680 ng  $\equiv$   $(6 \times 10^{14} / 3657)$  molecules, where 3657 is the length of pAT 153 linear molecules.

. .  $20 \mu\text{g} = (6 \times 10^{14} / 3657) \times (20 / 0.68)$  molecules  
=  $4.8 \times 10^{12}$  molecules  
=  $9.6 \times 10^{12}$  3' termini

2. Specific activity of [ $^3\text{H}$ ]dGTP in reaction

10  $\mu\text{Ci}$  [ $^3\text{H}$ ]dGTP in reaction, specific activity 11.7  $\mu\text{Ci}/\text{nmole}$

4  $\mu\text{l}$  of 2mM dGTP added

= 8 nmoles

. . specific activity = 10  $\mu\text{Ci}/9$  nmoles

= 1.1  $\mu\text{Ci}/\text{nmole}$

=  $9.7 \times 10^5$  cpm/nmole

assuming  $^3\text{H}$  counting 40% efficient and 1  $\mu\text{Ci} \equiv 2.22 \times 10^6$  dpm

3. cpm expected for 1 residue at each 3' terminus

$$(9.7 \times 10^5) / (6 \times 10^{14}) \times 9.6 \times 10^{12} = 1.6 \times 10^4 \text{ cpm}$$

4. Actual number of residues added

1  $\mu$ l from 100  $\mu$ l reaction TCA precipitated

. . 4.3  $\times 10^5$  cpm incorporated

. . (4.3  $\times 10^5$ ) / (1.6  $\times 10^4$ ) residues added = 27

APPENDIX III

Calculation of amounts of tailed cDNA/RNA  
in sucrose gradient fractions

Total counts loaded on to gradient = 565 850

Total nucleic acid on gradient = 0.5 $\mu$ g

Fractions Cerenkov counted, 50% efficient

Therefore assume 0.5 $\mu$ g cDNA/RNA  $\equiv$  280 000 cpm

0.05 $\mu$ g cDNA/RNA required per annealing reaction

i.e.  $280\ 000 \times 0.05/0.5 = 2.8 \times 10^4$  cpm

. . . 0.025 $\mu$ g cDNA/RNA  $\equiv 1.4 \times 10^4$  cpm

APPENDIX IV

Determination of restriction fragment sizes

Programme for Texas Instruments SR-56 calculator. (Modified from method given by Southern 1979.)

Loc.	Code	Key	Loc.	Code	Key	Loc.	Code	Key
00	74	-	35	93	+/-	69	34	RCL
01	34	RCL	36	84	+	70	09	9
02	02	2	37	34	RCL	71	94	=
03	94	=	38	05	5	72	33	STO
04	33	STO	39	94	=	73	08	8
05	09	9	40	54	÷	74	41	RS
06	64	X	41	52	(	75	74	-
07	52	(	42	01	1	76	52	(
08	34	RCL	43	74	-	77	34	RCL
09	05	5	44	34	RCL	78	06	6
10	74	-	45	06	6	79	64	X
11	34	RCL	46	94	=	80	34	RCL
12	03	3	47	33	STO	81	08	8
13	94	=	48	07	7	82	94	=
14	54	÷	49	41	RS	83	33	STO
15	52	(	50	74	-	84	09	9
16	34	RCL	51	34	RCL	85	41	RS
17	02	2	52	07	7	*86	74	-
18	74	-	53	94	=	87	34	RCL
19	34	RCL	54	20	*1/x	88	07	7
20	04	4	55	33	STO	89	94	=
21	94	=	56	06	6	90	20	*1/x
22	54	÷	57	74	-	91	64	X
23	52	(	58	52	(	92	34	RCL
24	34	RCL	59	34	RCL	93	08	8
25	03	3	60	03	3	94	84	+
26	74	-	61	74	-	95	34	RCL
27	34	RCL	62	34	RCL	96	09	9
28	01	1	63	07	7	97	94	=
29	94	=	64	53	)	98	41	RS
30	33	STO	65	20	*1/x	99	42	RST
31	06	6	66	94	=			
32	64	X	67	20	*1/x			
33	34	RCL	68	64	X			
34	01	1						

Registers

0	$L_1$	5	$m_3$	where $L_1$ = length of fragment 1 in nucleotides and $m_1$ = mobility of fragment 1.
1	$m_1$	6	used	
2	$L_2$	7	$m$	
3	$m_2$	8	$K_1^0$	
4	$L_3$	9	$K_2$	

Check programme using  $L_1 = 4340$ ,  $m_1 = 26.1$ ,  $L_2 = 2260$ ,  
 $m_2 = 44.2$ ,  $L_3 = 602$ ,  $m_3 = 85.5$  when  $m_0 \sim -5.79$ ,  $K_1 \sim 183218$   
and  $K_2 \sim -1405$ .

User Instructions

Step	Procedure	Enter	Press	Display
1	Enter data points $L_1$ $m_1$ $L_2$ $m_2$ $L_3$ $m_3$ in registers 0-5 respectively			
2	Reset program counter to 00		RST	
3	Compute $m_0$		RCLO RS	$m_0$
4	Compute $k_1$		RCLI RS	$k_1$
5	Compute $k_2$		RCLO RS	$k_2$
6	Compute unknown $L_u$ from known $m$ $m$		GTO 86 RS	$L_u$

APPENDIX V

Linear regression programme for Texas Instruments SR56 calculator

Programme

Loc	Code	Key	Loc	Code	Key	Loc	Code	Key
00	33	STO	33	05	5	65	08	8
01	00	0	34	34	RCL	66	93	+/-
02	35	SUM	35	02	2	67	84	+
03	01	1	36	33	STO	68	26	*f(n)
04	64	X	37	06	6	69	01	Mean
05	32	$x \geq t$	38	34	RCL	70	33	STO
06	33	STO	39	04	4	71	02	2
07	03	3						
08	94	=	40	54	$\div$	72	26	*f(n)
09	35	SUM	41	34	RCL	73	00	S.Dev
10	04	4	42	07	7	74	20	*1/x
11	34	RCL	43	74	-	75	64	X
12	00	0 <sub>2</sub>	44	34	RCL	76	34	RCL
13	43	x	45	08	8	77	09	9
14	35	SUM	46	64	X	78	64	X
15	02	2	47	26	*f(n)	79	34	RCL
16	34	RCL	48	01	Mean	80	03	3
17	03	3	49	54	$\div$	81	94	=
18	26	*f(n)	50	34	RCL	82	41	R/S
19	04	$\Sigma+$	51	09	9 <sub>2</sub>	83	64	X
20	41	R/S	52	43	x <sup>2</sup>	84	34	RCL
21	42	RST	53	54	$\div$	85	03	3
22	26	*f(n)	54	52	(	86	84	+
23	01	Mean	55	01	1	87	34	RCL
24	33	STO	56	74	-	88	02	2
25	08	8	57	34	RCL	89	94	=
26	26	*f(n)	58	07	7	90	41	R/S
27	00	S.Dev.	59	20	*1/x	91	74	-
28	33	STO	60	94	=	92	34	RCL
29	09	9	61	33	STO	93	02	2
30	34	RCL	62	03	3	94	94	=
31	01	1	63	64	X	95	54	$\div$
32	33	STO	64	34	RCL	96	34	RCL
						97	03	3
						98	94	=
						99	41	R/S

Registers

0	Used	5	$\Sigma$
1	Used	6	$\Sigma y^2$
2	b	7	N
3	m	8	$\bar{x}$
4	$\Sigma xy$	9	$\sigma x^2$

User Instructions

Step	Procedure	Enter	Press	Display
1	Enter program			
2	Initialize		*CMs RST	
3	Enter data points Repeat Step 3 for all values	x; y;	$x > t$ R/S	i
4	Calculate variables Correlation coefficient Slope Intercept		GTO 22 R/S RCL 3 RCL 2	r m b
5	Calculate new y points Repeat Step 5 as many times as needed	x	GTO 83 R/S	y'
6	Calculate new x points	y	GTO 91 R/S	x'
7	Return to Step 2 for new problem			





Therefore the % hybridised at time, t

$$= 1 - \frac{R_t}{R_0} \times 100$$
$$= 1 - \frac{1}{e^{khR_0 t}} \times 100$$

The following values are obtained for a sequence present at a level of 1% of the total RNA, as mentioned above:

---

t(sec.)	% hybridised
1	0.1
10	0.99
100	10
1 000	65
10 000	99.99

---

Therefore, after 2.8h any sequence which is present in the total RNA at a level of 1% or above should have all hybridised.