Molecular cloning of three major sequence species from Rainbow trout protamine mRNA

J.R.Jenkins\*<sup>†</sup>, J.O.Bishop<sup>+</sup> and P.H.W.Butterworth<sup>++</sup>

\*Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2E 3PX, <sup>+</sup>Department of Genetics, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JN, and <sup>++</sup>Department of Biochemistry, University College, London, WC1E 6BT, UK

Received 9 July 1979

#### ABSTRACT

Double stranded cDNA molecules complementary to purified Rainbow trout protamine mRNA have been cloned in the bacterial plasmid pBR322. In order to circumvent the problems associated with a heterogeneous cDNA probe when identifying recombinants, a comparative hybridisation technique was used which can resolve between closely related cloned sequences. Using this technique, selected recombinants were shown to carry sequences corresponding to separate major fractions of protamine mRNA. Partial nucleotide sequences of the inserts in two clones confirms this conclusion.

#### INTRODUCTION

The protamines are a family of small, highly basic proteins synthesised during the late stages of spermatogenesis in the Rainbow trout (<u>Salmo</u><u>gairdnerii</u>). They displace the histones and play a unique role in packaging the sperm DNA (for review see (1)). Three protamine amino acid sequences have been determined (2), consisting of 32-33 residues, approximately two-thirds of which are arginine. In order to study the control of protamine gene expression and the genomic organisation and divergence of this closely related family of proteins, bacterial plasmids have been constructed which contain cDNA sequences complementary to major fractions of protamine mRNA.

Protamine mRNA from Rainbow trout testis was extensively purified by repeated cycles of oligo dT cellulose chromatography and electrophoresis in native and denaturing polyacrylamide gels. Double stranded (ds) DNA complementary to this purified mRNA was synthesised as described by Maniatis et al. (3) and this ds DNA inserted into the Pst 1 site of pBR322 using an oligo dG-dC tailing procedure. Restriction analysis of recombinants shows that both full length and/or near full length cDNA has been cloned and that three plasmids designated pTP4, pTP8 and pTP12, together represent the bulk of the input cDNA sequences. In order to circumvent the problems associated with cDNA probes containing a mixture of abundant sequences, a hybridisation

C Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

# **Nucleic Acids Research**

procedure was used to resolve between different cloned sequences capable of cross hybridisation.

## MATERIALS AND METHODS

# Trout testis

Rainbow trout testes were collected at the Dantrout Ltd. Factory, Brande Denmark in October 1975. Fresh testes were rinsed briefly in distilled water, frozen in solid CO<sub>2</sub> and stored at  $-70^{\circ}$ .

# Preparation of total polysomal RNA

All buffer solutions were treated with diethylpyrocarbonate followed by autoclaving. Glassware was siliconised, rinsed thoroughly with double distilled water and heated overnight at 160<sup>0</sup>.

Thawed trout testes were homogenised in 2 volumes of ice cold buffer containing 50 mM Tris.HCl,pH 7.6, 5 mM MgA $\bar{c}$ , 25 mM KCl, 250 mM sucrose, 50  $\mu$ g/ml cycloheximide and 50  $\mu$ g/ml polyvinyl sulphate at setting 80 for 1 min in a Sorvall omnimix blender.

The homogenate was centrifuged for 30 minutes at 12,000 rpm in a Sorvall 8 x 50 angle rotor at  $5^{\circ}$ , the supernatant decanted and Nonidet P40 added to 1%. The supernatant was then layered on 2 ml cushions of 1M sucrose in the same buffer and centrifuged for 3 hrs at 40,000 rpm in a Beckman SW40 rotor at  $4^{\circ}$ . After centrifugation the polysomal pellets were taken up in 20 mM HEPES pH 7.5 and 1 volume of 0.1M NaAc pH 6.0 added. The solution was made 0.5% in SDS and RNA was extracted at room temperature twice with equal volumes of Phenol:Chloroform,1:1 and once with two volumes of Chloroform, followed by precipitation with ethanol.

#### Isolation of total polysomal polyadenylated RNA

The RNA was taken up in 10 mM Tris.HCl,pH 7.5 and the buffer adjusted to 0.5M NaCl,20 mM Tris. HCl,pH 7,1mM EDTA,1% SDS by the addition of 10X stock buffer. The RNA was enriched in polyadenylated sequences by two cycles of oligo dT cellulose chromatography followed by ethanol precipitation.

# Isolation of protamine mRNA

Polyadenylated RNA was taken up in 20 mM Tris.HCl, pH 7.5, 40 mM NaCl, 10 mM EDTA,1% SDS and fractionated on 10-40% neutral sucrose gradients in this buffer, by centrifugation for 18 hrs at 40,000 rpm in a Beckman SW40 rotor at 25<sup>0</sup>. Protamine mRNA was recovered from gradient fractions by a single cycle of oligo dT cellulose chromatography followed by ethanol precipitation.

The protamine mRNA was taken up in 36 mM Tris pH 7.6.30 mM  $NaH_2PO_4$ , 1 mM  $Na_2$  EDTA, 0.2% SDS, and fractionated in 7.5% acrylamide 0.19% bisacrylamide gels in this buffer. The gels were scanned at 265 mM in a Joyce Loebel spectrophotometer. Gel slices were macerated by forcing them through a hypodermic needle, and the mRNA eluted by mild agitation overnight in 10 mM Tris.HCl,pH 7.6,0.2M NaCl, 1mM Na2EDTA, at 4<sup>0</sup>. The NaCl concentration was then increased to 0.5M and SDS added to 1%. Protamine mRNA was recovered by a cycle of oligo dT chromatography and ethanol precipitated. This mRNA was taken up in 20 mM diethylbarbituric acid, pH 7.5,20 mM NaCl, 8 mM EDTA 98% formamide, heated to 70<sup>0</sup> for 1 minute, and fractionated on 7.5% acrylamide 0.19% bisacrylamide gels in this buffer. These gels were cast in distilled water and the buffer diffused in by two sequential washes in the appropriate solution. Protamine mRNA was recovered as before.

# Preparation and incubation of wheat germ cell free translation system

Wheat germ S-30 prepared essentially by the method of Roberts and Paterson (4) was a gift from Dr. J. Sampson. Incubations, in a total volume of 60 µl, contained 20 µl wheat germ S-30, 25 mM HEPES. KOH,pH 7.5, 80 mM KCl, 1.5mM MgAc, 200 µM spermidine, 1mM DTT, 4mM ATP, 0.4 mM GTP, 0.4mM cold amino acids, and 15 µM  $C^{14}$  arginine or  $C^{14}$  leucine (318mCi/mmole or 155mCI/mmole (Amersham)).Samples were spotted onto 2cm Whatman No. 1 filters, and boiled for 15 minutes in 5% trichloracetic acid, 0.25% sodium tungstate pH 2.0 (TCA/W04). They were then rinsed once in this solution, once in ethanol and once in diethyl ether. After drying, radioactivity was determined using a toluene based scintillant.

# Preparation of high specific activity complementary H<sup>3</sup>-cDNA

High specific activity cDNA, complementary to gel purified protamine mRNA was prepared as described by Bishop and Freeman (5). Specific activity was  $10^7$  cpm/µg.

# **RNA** excess hybridisations

Hybridisations were performed as described (5)

# Synthesis of double stranded cDNA

Double stranded cDNA (ds cDNA) was synthesized essentially as described by Maniatis et al. (3), with minor modifications.

Full length single stranded cDNA was isolated on a 7.5% acrylamide denaturing gel. After Sl treatment to cleave hairpin loops, the ds cDNA was fractionated on a 10-40% neutral sucrose gradient. The high molecular weight fractions from this gradient were precipitated with ethanol and stored at  $-20^{\circ}$ .

## Homopolymer tailing

Oligo dG tails approximately 15 nucleotides in length were added to Pst l cleaved pBR322 in an incubation containing 50 mM HEPES .KOH, pH 7.4, 6 mM MgCl<sub>2</sub>, 5 mM DTT,  $H^3$ -GTP (Amersham), specific activity 800 mCi/mmole and terminal transferase, at 37<sup>o</sup>.

Oligo dC tails 20-30 nucleotides in length were added to ds cDNA in an incubation containing 50 mM HEPES .KOH, pH 7.4, 1 mM CaCl<sub>2</sub>, 0.1 mM  $\beta$ -mercaptoethanol, 0.5 mM dCTP (Amersham) specific activity 800 mCi/mmole, and terminal transferase, at 37<sup>0</sup>.

#### Annealing of molecular hybrids

0.11  $\mu$ g dG-tailed pBR322 was annealed with an equimolar amount of dCtailed ds cDNA in 50  $\mu$ l of a buffer containing 10 mM Tris.HCl, pH 8.0, 0.1M NaCl, 1 mM EDTA and was incubated for 2 hrs at 60<sup>0</sup>, then allowed to cool slowly to room temperature.

# Transformation and colony selection

All manipulations were carried out in a Category II containment laboratory as recommended by the UK Genetic Manipulation Advisory Group. Molecular hybrids were used to transform <u>E.coli</u>  $\chi$ 1776 using the procedures described by Inoue and Curtis (6). Recombinant colonies were selected for tetracycline resistence, ampicillin sensitivity (Tet<sup>+</sup>amp<sup>-</sup>).

#### Isolation of plasmid DNA

Plasmid DNA was isolated by chloramphenicol amplification and lysis with Triton X-100 as described by Clewell and Helinski (7).

#### Purification of ds cDNA inserts

Pst 1 endonuclease-cleaved plasmid DNA was fractionated on 10-40%

neutral sucrose gradients in 20 mM Tris.HCl, pH 7.5, 40 mM NaCl, 1 mM EDTA 1% SDS by centrifugation for 23 hrs at 40,000 rmp in a Beckman SW40 rotor at 25<sup>0</sup>. Insert DNA was recovered from gradient fractions by ethanol precipitation.

# Analytical agarose gel electrophoresis

Restriction digests of plasmid DNA were electrophoresed in 2% agarose vertical slab gels containing 89 mM boric acid, 89 mM Tris, 2.5 mM  $Na_2EDTA$ . Gels were stained with EthBr (1 µg/ml) and photographed under UV light.

# Insert excess hybridisation

Insert DNA was hybridised to high specific activity protamine cDNA at  $70^{\circ}$  in 16 mM Tris.HCl, pH 7.5, 0.36 M NaCl. Insert DNA (400 ng) and H<sup>3</sup> cDNA (2 ng) in 50 µl 10 mM Tris.HCl, pH 7.5 was heated in stoppered and siliconised glass tubes for 3 min at  $100^{\circ}$  to denature the insert DNA duplex. The tubes were then transferred to  $70^{\circ}$  and a 10 µl sample was removed for Sl digestion. 10 µl of 1.8 M NaCl, 0.1 M Tris. HCl, pH 7.5 was added to the hybridisation mixture which was then overlaid with paraffin oil and incubated at  $70^{\circ}$  for 1 hr, after which a second 10 µl sample was removed for Sl digestion. Samples were digested with Sl nuclease in 70 mM NaCl, 30 mM NaAc, pH 4.5, 1 mM ZnSO<sub>4</sub> at  $50^{\circ}$  for 40 minutes. Where inserts were hybridised to cDNA in pairs, the conditions were identical except that 200 ng of each insert was present to give a 50-fold molar excess of each.

#### RESULTS

# Purification of protamine mRNA

When polyadenylated polysomal RNA is fractionated on neutral sucrose gradients, the protamine mRNAs sediment as a distinct peak at about 5.7S (8, and J.R. Jenkins, unpublished results). Fractionation on native (Fig. 1A) and denaturing (Fig. 1B) acrylamide gels further purifies this 5.7S RNA and partially resolves in into three and two major subspecies respectively. When these polyA + mRNAs are added to a wheat germ cell free system they stimulate the incorporation of  $C^{14}$  arginine into trichloracetic acid/sodium tungstate (TCA/WO<sub>4</sub>) insoluble material (Table 1).

#### In vitro translation

From published amino acid sequences (2) it is known that rainbow trout protamines are rich in arginine but do not contain any leucine residues. The



Figure 1. Acrylamide gel electrophoresis of protamine mRNA (A) 7.5% non-denaturing gel (B) 7.5% denaturing gel. Gels were run as described in Materials and Methods.

relative stimulation of a  $C^{14}$  arginine and  $C^{14}$  leucine incorporation observed when trout testis mRNA fractions are added to the wheat germ cell free system therefore serves as a convenient way of estimating the purity of protamine mRNA (9), since contaminating messengers are almost certain to stimulate

<u>TABLE 1</u> Translation of trout testis mRNA in the wheat germ cell free system. 60  $\mu$ l incubations were carried out as described in Materials and Methods, and contained 0.1  $\mu$ g of gel purified protamine mRNA and 1.7  $\mu$ g total polysomal polyA + mRNA respectively.

	cpm C <sup>14</sup> -arginine	cpm C <sup>14</sup> -leucine
endogenous synthesis	313	66
+ protamine mRNA	7913	31
Stimulation	25X	-
endogenous synthesis	65	14
+ polysomal polyA + mRNA	1211	191
Stimulation	19X	14X

leucine incorporation. When denaturing gel purified protamine mRNA is translated in the wheat germ cell free system the incorporation of  $C^{14}$ -arginine into TCA/WO<sub>4</sub> insoluble material is stimulated more than 25-fold, while  $C^{14}$ -leucine incorporation is slightly inhibited (Table 1). In contrast, total polyadenylated RNA from trout testis stimulates  $C^{14}$ -arginine incorporation some 19-fold, and  $C^{14}$ -leucine incorporation more than 14-fold (Table 1). These results suggest that the protamine mRNA has been substantially

purified. The slight inhibition of endogenous  $C^{14}$  leucine incorporation by purified protamine mRNA (Table 1) is probably a result of this RNA competing with endogenous messengers for components of protein synthesis. Protamine mRNA at similar concentrations also inhibits the endogenous incorporation of  $C^{14}$  phenylalanine in this system (J.R. Jenkins, unpublished data).

# Sequence complexity analysis of protamine mRNA

By the criteria of physical size and the ability to selectively stimulate arginine incorporation in the wheat germ cell free system the protamine mRNA is essentially pure. However, the RNA preparation could still be contaminated with a mixture of polyadenylated RNAs which are translationally inactive but nevertheless act as efficient templates for reverse transcription and so be represented in the final clones. In addition, it is known that there are several protamines (2) of slightly differing protein sequence, so that it is almost certain that cDNA synthesized using "purified" protamine mRNA as template will be heterogeneous. To determine the sequence composition of the protamine mRNA used in this work, a sequence complexity analysis was performed by the methods of Bishop and Freeman (5). The results (Fig. 2) show that 72% of the cDNA hybridises with a Rot} of 4.4 x  $10^{-4}$ . Using rabbit  $\alpha$ -chain globin data as a kinetic standard (10) and assuming a protamine mRNA/cDNA chain length of 284 bases, the data indicates that this major fraction of the mRNA is composed of 3-4 sequence species. This is in general agreement with other estimates for the sequence complexity of protamine mRNAs (9,11) and work which suggests that there may be a maximum of six unique protamine genes per haploid genome (9). However, caution must be used in interpreting this complexity data since it is likely that a significant level of cross hybridisation takes place between the mRNA sequences of these slightly dissimilar proteins. Such cross hybridisation would tend to both raise the apparent complexity of the mRNA



Figure 2. Hybridisation kinetics of protamine mRNA with its cDNA copy. Rot curve analysis was performed as described in Materials and Methods

fraction (12) and depress the apparent fraction of mRNA hybridising at low Rot.

With these reservations it appears that the purified protamine mRNA is suitable starting material for molecular cloning.

# Cloning of ds cDNA derived from protamine mRNA

Protamine ds cDNA was cloned in the bacterial plasmid pBR322. This plasmid contains a single Pst 1 cleavage site, located in the coding region of the  $\beta$ -lactamase gene (13) (Fig. 3). Transformants which contain DNA sequences inserted within this Pst 1 site can be detected by screening for a tetracycline resistant, ampicillin sensitive phenotype (Tet<sup>+</sup>amp<sup>-</sup>). However, this screening procedure does suffer from the slight disadvantage that some recombinants containing DNA within the Pst 1 site score as ampicillin resistant (amp<sup>+</sup>) and others as ampicillin partially resistant (amp<sup>±</sup>)



Figure 3. Strategy used for cloning protamine cDNA.  $Ap^r$ ,  $\beta$ -lactamase gene; Ic<sup>r</sup>, tetracycline resistence gene; Eco Rl, Pst l, restiction enzyme cleavage sites; ds cDNA, double stranded cDNA.

(J.R. Jenkins, unpublished data). This is probably due to transcription proceeding through the inserted DNA sequences to give a modified but functional or partially functional  $\beta$ -lactamase protein. ds DNA was annealed to vector plasmid DNA using oligo dG-dC tailing as described in Materials and Methods. An advantage of this method is that the outcome of tailing a Pst 1 cleaved 3' end with dG residues is the regeneration of a complete Pst 1 recognition site on that strand of DNA (Fig. 3). When Pst 1 -cleaved oligo dG-tailed pBR322 is annealed with dC-tailed ds cDNA and transfected into a host strain such as  $\chi$ 1776, in vivo repair results in a Pst 1 cleavage site at either side of the insert DNA (Fig 3). This allows cloned DNA sequences to be separated from pBR322 DNA by a single Pst 1 digestion followed by size fractionation.

# Size of the inserted sequences

12 recombinants were grown up in small scale liquid culture and the plasmid DNA isolated as described (Materials and Methods). Following cleavage with Pst 1 endonuclease, plasmid DNA was fractionated on 2% agarose gels with an appropriate molecular weight standard (Fig. 4) and the size range of the inserts calculated (Table 2). These recombinants contain inserts of up to 310 base pairs and allowing up to 40 base pairs per insert for oligo dG-dC linkers and a size range of 272-300 bases for protamine mRNA (Fig. 1), it appears that full length or near full length cDNA sequences have been cloned.

The recombinants containing the largest inserts, designated pTP4, pTP6,pTP8,and pTP12 were chosen for further study.

#### Hybridisation analysis of inserted sequences

It has been shown that the protamine mRNA starting material appears to contain 3-4 major sequence species (Fig. 2) which presumably correspond to various species of protamine. Since the amino acid sequences of these



Figure 4. Analysis of cDNA insert size on 2% agarose gels. Plasmid DNA was cleaved with Pst 1 endonuclease and run on agarose gels as described in Materials and Methods. Tracks (from right to left): 1-8, plasmids pTP1-pTP8; track 9, R1 partial digest of mouse satellite DNA; tracks 11-14, plasmids pTP9-pTP12. The mouse satellite monomer comprises 243 base pairs. Plasmid Insert Size (base pairs) pTP4 310 pTP6 285 pTP8 245

220

pTP12

<u>TABLE 2</u>. Size of insert DNA carried by selected recombinants. Size values were obtained by reference to a log linear plot constructed using mouse satellite partial RI digest as a reference (see Fig. 4).

proteins are so similar (2) it might be expected that their mRNA sequences would possess sufficient homology for some cross hybridisation to occur. This expectation is supported by the work of Sakai et al. (9) who found evidence for cross hybridisation in both mRNA-cDNA and cDNA - total DNA reactions, and more recently by the direct determination of several protamine mRNA sequences (14, and Figure 5 below) which shows conclusively that these messengers contain extensive regions of common sequence. From this it can be predicted that recombinants containing substantial lengths of protamine mRNA derived sequence will (a) confer S1 resistance under stringent conditions on a significant (say >20%) fraction of protamine cDNA and (b) contain regions of homology with other similar but not identical recombinants. The results of the hybridisation analysis (Table 3) show all four recombinants tested fulfil both these criteria.

Plasmids pTP4, pTP6, pTP8 and pTP12 were cleaved with Pst 1 endonuclease and after phenol-chloroform extraction the insert fragments were isolated on neutral 10-40% sucrose gradients (the fragment isolation relates to other work and is unnecessary for the hybridisation procedure). The gradients were scanned at 265 nM and the appropriate fractions retained. This purified DNA was then hybridised with  $H^3$ -cDNA complementary to protamine mRNA under conditions where the insert fragment was in estimated 100-fold plus strand molar excess over cDNA (see Materials and Methods).

As a control, protamine mRNA was hybridised to cDNA under conditions in which mRNA was in estimated 100-fold molar excess. To test the possibility that other cloning products such as oligo dG-dC linkers were conferring S1 resistence on cDNA a further control was included in which total Pst 1-cleaved  $p\lambda7$  DNA was hybridised with cDNA under conditions where plasmid plus strand was in estimated 100-fold molar excess over cDNA.  $p\lambda7$ is a pBR322-derived recombinant containing a  $\lambda$  HindII fragment within the

CTS	TAG		TAG	CTS		
Arg	CGT		CGT	Burg		
, Brit	AGG (		AGG .	, Brit		e v
Buy	AGG		AGG	, Cut		fro he ). an
hing	CGC		ບິວ	, Bri		• 14) (2 14)
Buy	cec	•	660	. R10		enc nes by (
£19	960	٠	GGA	R15		ted ther ther
GLY	299	•	AGA	Arg		d s rot cat
вuy	AGG		AGG	Buy		aci le pu
в	ပ	4	CGC	Ang	Prod 4	e ed s
Å	ទ	٩	CGT	Ang	8 R1	aci w t ariar
bry	CGT		CGT	Arg	9 	inbo des ons
Arg	090	•	CGA	Arg	Stra	am am siti
Sær	TCC		TCC	Ser	Ser	de e ans ans
Val	GTG		GTG	Val	Val	anco nco tair tair
Arg	AGG		AGG	Arg	8 Burg	d ei duei auei Base
Pro	22		222	Pro	Pro 11a	ereart.
Arg	CGC		ວຍວ	Arg		cide un v de un v
Arg	CGT	•	292	Arg	Bury	10 e c i L B a c i
Arg	292		ວຍວ	Arg	Val Ile	d a c t a d a c t a
Arg	AGG		AGG	Arg	oud	d ar lso a a
Arg	292		282	Arg	Arg	A) tia TP4 n a
$_{lle}$	ATC	•	GTC	Val	Ser 3	par ( (p) ca
Arg	CGT	٠	CCT	Pro		P8. B) ene rts
Arg	ပ္စင္ပ	•	SGA	Ang	Arg	pT ) g nse nse
Ser	AGC	•	AGA	Arg	Pro	e d ctd c
A1a	000	•	TCC	Ser		ins ins th; th
Arg	AGA	:	TCC	Ser		nse of
Arg	090					a or 8 j the ence
r Arg	AGA					the pTF of
o Arg	C AG					s se
ž	о С					ion of i lete
<b>f</b> Nizi	ATC					v seq on p on p on p
						se degi o degi
-	F		:	5	Ω.	i this of the second
	2			-	5	din din lete
						또볁 2 골 볼 융

Sequence	Counts p	er 5 min	% S1 resistance		% S1 resistance	
	T <sub>0</sub> ± S1	T <sub>75</sub> ± S1	т <sub>о</sub>	T75	control	
P4	246 3846	690 2775	6.4	24.8	35.4	
p6	134 3441	718 2808	3.9	25.5	36.6	
p8	160 3440	639 2487	4.65	25.7	36.9	
p12	147 3218	525 2217	4.5	23.7	33.6	
р <b>λ</b> 7	83 3001	84 2340	2.8	3.6	0	
mRNA	143 3385	1481 2334	4.2	63.5	100	
p4 + p6	102 3604	465 2099	2.8	22.2	31	
p4 + p8	102 3741	794 2016	2.7	37.7	57	
p6 + p8	92 · 3198	878 2067	2.8	42.5	65	
p4 + p12	109 3309	676 2121	3.3	31.9	47.3	
p8 + p12	130 3527	769 1997	3.6	38.5	58.2	

<u>TABLE 3</u>. Hybridisation analysis of cDNA fragments carried in four recombinants. Samples were processed as described in Materials and Methods.

Pst 1 site and was constructed by the oligo dG-dC tailing procedure (J.R. Jenkins, unpublished work). In the same way different insert fragments were hybridised in pairs with cDNA under conditions where each insert was in estimated 50-fold plus-strand molar excess over cDNA.

The results (Table 3) show the following: firstly, in every case inserts contain sequences homologous to more than 20% of protamine cDNA; secondly, the sequences contained in inserts from pTP4 and pTP6 are probably nearly identical as evidenced by the lack of additive protection from S1 nuclease of cDNA when the two inserts are present. However, the sequences carried may not be identical since the pairs pTP4 + pTP6 and pTP6 + pTP8 give different values for S1 resistence of cDNA; thirdly, the sequences contained in the inserts from pTP4, pTP8 and pTP12 are all dissimilar, since any pair from this group shows additive protection of cDNA. Finally, the inserts from pTP4, pTP8 and pTP12 have regions of homology with each other, as the additive protection conferred on the cDNA by any pair from this group is less than the sum of the protection conferred by the two individual inserts in isolation.

The inserts contained in pTP4 and pTP8 have been sequenced and the complete data has now been published elsewhere (14). Figure 5 shows the nucleotide and encoded amino acid sequences from the coding regions of these inserts together with the generalised amino acid sequence of trout protamines. This data proves unambiguously that the cloned sequences are derived from protamine mRNAs and comparison of the nucleotide sequences shows extensive but incomplete homology, thereby confirming the results of the insert excess hybridisation.

## DISCUSSION

Three plasmids pTP4, pTP8 and pTP12 have been shown by the insert excess hybridisation procedure to contain sequences complementary to three major fractions of protamine mRNA (Table 3). One explanation for the apparent homology between these inserts is that they contain partial but overlapping sequences derived from different regions of the same mRNA. However, a comparison of insert and mRNA size rules this out. Allowing for 40-60 base pairs of oligo dG-dC linkers, the insert sizes in each possible combination would result in a far greater degree of homology than observed. For example pTP4 and pTP8, if derived from the 5' and 3' ends of a 300 base mRNA would share some 60% of uninterrupted common sequence, compared with the 15.3% observed. However, it is important to stress that while the procedure makes it possible to resolve between related sequences, the conditions of stringency used mean that the SI resistence values obtained are conservative. A number of these cloned cDNA frangments have now been completely sequenced (14). All fragments so far examined contain regions coding for protamines and the sequence data in every case supports the results of insert excess hybridisation. This hybridisation procedure is therefore a very useful preliminary step when analysing recombiants derived from cDNA containing several abundant species, especially in the case of protamine mRNA derived cDNA where abundant species cross hybridise extensively.

#### ACKNOWLEDGEMENTS

This work was supported by the Medical Research Council and the Science Research Council of Great Britain.

<sup>†</sup>To whom correspondence should be addressed.

# REFERENCES

- Dixon, G.H. (1972) Karolinska Symposia on Research Methods in 1. Reproductive Endocrinology, 5th Symposium, 130-154 Ando, T. and Watanabe, S. (1969) Int. J. Protein Res. 1, 221-224 Maniatis, T., Kee, S.G., Efstratiadis, A. and Kafatos, F.C. (1976)
- 2.
- 3. Cell 8, 163-182
- Roberts, B.E. and Paterson, B.M. (1973) Proc. Nalt. Acad. Sci. U.S.A. 4. 70, 2330-2334
- Bishop. J.O. and Freeman, K.B. (1973) Cold Spring Harbor Symp. Quant. 5. Biol. <u>38</u>, 707-716
- Inoue, M. and Curtis, R. (1977) Manuscript in preparation. 6.
- 7.
- 8.
- Clewell, D.B. and Helinski, D.R. (1970) Biochem. 9, 4428-4440 Gedamu, L. and Dixon, G.H. (1976) J. Biol. Chem. 251, 1455-1463 Sakai, M., Fuji-Kuriyama, Y. and Muramatsu, M. (1978) Biochem. <u>17</u> 9. 5510-5515
- Bishop, J.O., Beckmann, J.S., Campo, M.S., Hastie, N.D.? Izquierdo, and Perlman, S. (1975) Phil. Trans. Roy. Soc. Lond. B. 272, 147-157. Levy, B.W. and Dixon, G.H. (1977) J. Biol. Chem. 252, 8062-8065 Μ. 10.
- 11.
- Bishop. J.O. (1972) Karolinska Symposia on Research Methods in 12. Reproductive Endocrinology, 5th Symposium, 247-273 Sutcliffe, J.G. (1978) Proc. Natl. Acad. Sci. U.S.A. <u>75</u>, 3737-3741
- 13.
- Jenkins, J.R. (1979) Nature 279, 809-811 14.