16. Mays, T. D., Smith, J., Welch, R., Defini, C. & Macrina, F. L. *Antimicrob. Agents Chemother.* **21**, 110–118 (1982).
17. Tally, F. P., Shimell, M. J., Carson, G. R. & Malamy, M. H. in *Molecular Biology, Pathogenicity and Ecology of Bacterial Plasmids* (eds Levy, S., Clowes, R. & Koenig, E.) 51–59 (Plenum, New York, 1981).
18. Smith, C. J., Welch, R. A. & Macrina, F. L. *J. Bact.* **151**, 281–287 (1982).
19. Nugent, M. E. *J. gen. Microbiol.* **126**, 305–310 (1981).
20. Stuy, J. H. *J. Bact.* **142**, 925–930 (1980).
21. Jacob, A. E. & Hobbs, S. J. *J. Bact.* **117**, 360–372 (1974).
22. Tomich, P. K., An, F. Y. & Clewell, D. B. *J. Bact.* **141**, 1366–1374 (1980).
23. Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683–3687 (1979).
24. Dunny, G. M. & Clewell, D. B. *J. Bact.* **124**, 784–790 (1975).
25. Meyers, J. A., Sanchez, D., Elwell, L. P. & Falkow, S. *J. Bact.* **127**, 1529–1538 (1976).

Regulatory sites for *his3* gene expression in yeast

Kevin Struhl*

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The expression levels of many genes are regulated in response to particular environmental or developmental cues. A simple regulated gene can conceptually be divided into three elements: structural components that encode the gene product, promoter elements that are essential for gene expression, and regulatory elements responsible for changing the level of expression after a specific stimulus[1,2]. When cells of the yeast *Saccharomyces cerevisiae* are subjected to amino acid starvation, *his3* and many other amino acid biosynthetic genes are expressed threefold above the basal level[3,4]. Previously, I isolated mutations mapping outside the structural gene that severely reduce or eliminate *his3* expression[5,6]. These mutations define two distinct *his3* promoter elements located 115–155 and 32–52 base pairs (bp) from the site of transcriptional initiation. Here I describe 18 small deletion mutations, some of which express *his3* at the basal level but are unable to increase the level of expression in the appropriate physiological conditions. These define two regulatory sites located 32–41 and 80–100 bp upstream from the site of transcriptional initiation, and they strongly suggest that these regions are necessary for the positive regulation of *his3* expression. I consider the results in terms of models for the general control of amino acid biosynthesis.

his3 encodes imidazoleglycerol phosphate dehydratase (IGPD), an enzyme involved in histidine biosynthesis. Its expression is regulated coordinately with the genes involved in amino acid biosynthesis. When yeast cells are starved of histidine, methionine or tryptophan by adding appropriate inhibitors to the growth medium, the levels of IGPD increase about threefold (Table 1). In addition, *his3* expression is controlled in the expected manner by the *aas2* and *tra3* genes.

It is probable that *his3* is regulated solely as a function of general amino acid biosynthesis: first, IGPD levels do not depend on the presence or absence of histidine in the growth medium (Table 1). Second, growth conditions resulting in the selective induction of histidine biosynthetic enzymes have never been found. Third, mutations selected for defective regulation of histidine enzymes always affect other amino acid biosynthetic enzymes also[3]. These experiments, which fail to demonstrate histidine-specific regulation in yeast, are identical to those that clearly establish such control in *Salmonella typhimurium*.

The mapping positions of 18 mutations of the cloned *his3* gene are presented in Table 2. Each mutation deletes sequences (ranging between 9 and 66 bp) within the untranscribed region located 6–138 nucleotides upstream from the 5′ end of the gene. Using standard techniques, the mutant genes were introduced back into yeast cells so that the resulting strains contained one copy of the cloned DNA per cell at the normal *his3* chromosomal location (for experimental details see Fig. 1 legend). Strains harbouring these cloned mutant alleles were always grown in medium containing histidine so that *his3* expression was gratuitous for cell growth; inducing conditions were achieved by starving the cells of tryptophan. The IGPD levels are shown in Table 2. Two strains were examined for each *his3* allele. One strain contained all the original transforming DNA sequences; the other lacked the vector sequences and is equivalent to replacement of the original *his3*-532 mutation by the allele of interest (see Fig. 1 legend). The IGPD levels for both strains of any given mutant were always the same. Of the 18 mutations tested, 9 regulate *his3* expression properly. However, strains with any of the nine mutations produce IGP constitutively, that is, the enzyme levels are the same in normal conditions and during amino acid starvation. Thus, these nine mutations have the expected properties of *his3* regulatory mutations.

his3 regulatory sites are defined by comparing the DNA sequences of mutations that either regulate or fail to regulate *his3* expression. The simplest interpretation of the data assumes that: (1) a subset of the sequences deleted by a regulatory mutation is essential for normal regulation; and (2) sequences deleted by correctly regulated derivatives are unimportant. Such analysis, when applied to all the derivatives, leads to the simple and internally consistent description of *his3* regulatory elements illustrated in Fig. 2 and discussed below. It is difficult to disprove alternative explanations for the phenotype of any given mutation, and more complex models can be advanced to explain the data.

The extent of *his3*-Δ22 indicates a regulatory site that includes a sequence between 32 and 41 bp upstream from the start of transcription (nucleotides −32 to −41). Deletion of this region accounts for the regulatory defects of five other mutations (*his3*-Δ20, Δ21, Δ23, Δ24 and Δ25). As yet, the DNA sequences sufficient to constitute this regulatory site are not well mapped. However, sequences further upstream than −52

Table 1 Regulation of the *his3* gene

Strain	Genotype	Complete	−His	−Trp	+AT	+Eth
KY29	*trp1*-289 *his3*$^+$	1.0	1.0	3.0	3.0	3.2
KY137	*trp1*-289 *his3*-532	<0.1	NG	<0.1	NG	NT
KY107	*aas2*-5039 *his3*$^+$	1.0	1.1	NT	1.5	1.3
KY108	*tra3*-1 *his3*$^+$	3.0	3.0	NT	3.5	3.4

The yeast strains used were: KY29 (*mat-a ura3*-52 *trp1*-289) isolated by M. Thomas as strain M1-2B; KY137 (*mat-a ade2*-1 *ura3*-1 *ura3*-2 *trp1*-289 *his3*-532 *can1*$^-$) isolated by S. Scherer; KY107 (*mat-a aas2*-5039) and KY108 (*mat-a tra3*-1) both obtained from G. Fink[3]. Because *tra3*-1 renders strains temperature-sensitive for growth, all cultures were incubated at 27 °C with shaking. Complete medium contained 2% glucose, 0.67% yeast nitrogen base without amino acids (Difco), and 0.1 mg ml^{-1} each of histidine, tryptophan, uracil and adenine. In these conditions, the doubling time was 2 h for all strains except KY108, which doubled every 3 h. Strains incubated in complete medium lacking histidine (−His) grew at the same rate, except for KY137 which did not grow (NG). Cells grown in either of these non-starvation conditions were collected in the middle of exponential growth (A_{600} = 2 or ~2 × 10^7 cells ml^{-1}). To achieve amino acid starvation, cells were first grown in complete medium until A_{600} was ~1.0, concentrated by centrifugation, washed once with distilled water, and then resuspended in complete medium that either lacked tryptophan (−Trp), lacked histidine but contained 10 mM aminotriazole (+AT), or contained 20 μM ethionine (+Eth). The resuspended cells were incubated for a further 6–9 h; the cells undergo an average of one more cell division before their growth is arrested as a result of Trp, His or Met starvation. To determine the levels of IGPD activity, cells were collected, washed twice with water, permeabilized with 1.5% chloroform at 37 °C for 20 min, pelleted by centrifugation, and resuspended in 0.1 M triethanolamine, *p*H 7.7. The enzyme assay was performed as described previously[12] with minor modifications (to be published elsewhere). The enzyme levels are normalized to both the number of cells assayed and to the amount of absorbance at 290 nM released after chloroform treatment. The activity levels shown are relative to those of wild-type yeast grown in complete medium (defined as 1.0). For enzyme levels around 1.0, the error is ~10%.

* Present address: Department of Biological Chemistry, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02115, USA.

0028–0836/82/460284—04$01.00

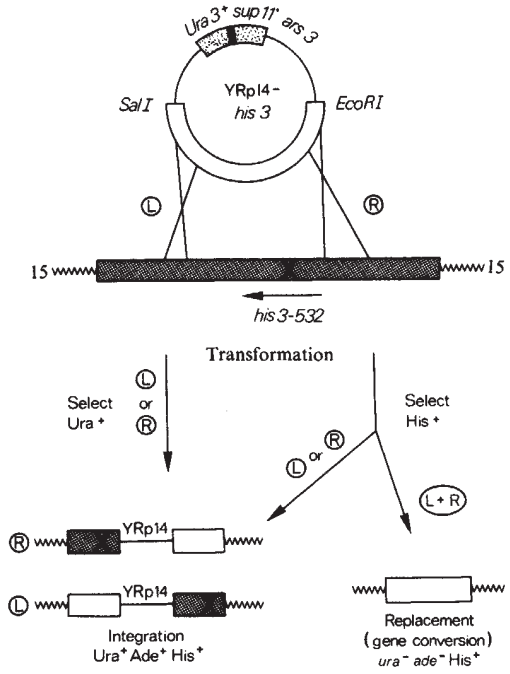

Fig. 1 Determining the phenotypes of cloned *his3* derivatives. The wild-type *his3* DNA used in these experiments (Sc2605) is a 6.1-kilobase pair (kb) fragment generated by *Eco*RI and *Sal*I cleavage[12]. In addition to the entire *his3* locus, it contains 2.9 kb flanking the 3′ end and 2.5 kb flanking the 5′ end. The isolation, haracterization and DNA sequence determination of the mutant derivatives of Sc2605 will be described elsewhere. The upper part of the figure shows the structure of *his3* hybrid molecules. All *his3*-containing DNA fragments (open boxes) were cloned into YRp14 (solid line with dotted boxes; see ref. 16), a vector containing the *ura3* gene, the ochre suppressing version of a tyrosine-inserting tRNA (*SUP11*), and *ars3*, a sequence that inefficiently causes hybrid molecules to replicate autonomously. YRp14 hybrid DNAs were introduced into KY137, a strain containing the ochre suppressible allele *ade2*-1, nonrevertible alleles *ura3*-1,2 and *his3*-532, and the revertible *trp1*-289 allele. Chromosome 15 of this strain is indicated by a wavy line below the YRp14–*his3* hybrid molecule. The shaded box indicates the 6.1-kb region homologous to the cloned *his3* derivatives; X marks the location of the *his3*-532 mutation[13]. Ura^+ or His^+ transformants were selected after 5 days of growth at 30 °C. All cloned *his3* fragments produced an equal number of His^+ and Ura^+ transformants for a given amount of DNA, except for Sc2884 (*his3*-Δ38) which produced no His^+ colonies. Essentially all such transformants were mitotically stable for the marker selected in the initial transformation event, indicating that the transforming sequences recombined with genomic DNA. Transformants containing autonomously replicating hybrid molecules were obtained only after ≥10 days. This is probably due to the inefficiency of *ars3*[16] and to lethality associated with multiple copies per cell of tyrosine-inserting suppressor tRNAs[14]. The figure shows recombination with the *his3* locus of chromosome 15. The L cross-over occurs to the left of the *his3*-532, while the R cross-over occurs to the right of this allele. The resulting structures are shown at the bottom. The YRp14 vector sequences which include *ura3*$^+$ *sup11*° *ars3* are indicated by a solid line. Almost all the Ura^+ transformants were also His^+ Ade^+ while the His^+ transformants fell into two classes. The class of His^+ transformant that was also Ura^+ Ade^+ is genotypically equivalent to the Ura^+ transformants. In this case, the entire YRp14 hybrid molecule has integrated into the genome. As most of the homology between the transforming DNA and the genome is *his3* DNA, integration probably occurs at the *his3* locus; this was true for all cases tested explicitly. The transformants that arise by integration are due to a single L or R cross-over. The other class of His^+ transformant is *ade*$^-$ *ura*$^-$ and results from replacement of the *his3*-532 allele by the transforming allele, probably via gene conversion[15]. This class is equivalent to the double (L+R) cross-over. Both classes of transformants should have one copy of the transforming DNA per cell. For the integration class, multiple copies of the *SUP11*-containing transforming DNA per haploid genome should be lethal and for the replacement class, it is difficult to envisage mechanisms that could result in multiple copies of *his3* and none of *ura3*. Indeed, hybridization analysis supports the above contentions. It is possible that cross-overs occur between the *his3*-532 and the cloned *his3* alleles—this would generate either the wild-type allele and/or the double mutant. However, as these alleles are very close (less than 250 bp; see ref. 13) recombination between them is rare[5]. In this regard, note that the phenotype of each allele was tested in two independent transformants.

or downstream from −32 are unlikely to be involved because both *his3*-Δ26 (which deletes nucleotides −53 to −80) and *his3*-Δ18 (nucleotides −32 to −21 deleted) are regulated properly.

The structures of the remaining deletion mutations suggest a second regulatory site: *his3*-Δ28 (nucleotides −106 to −53 deleted) and *his3*-Δ29 (nucleotides −106 to −73 deleted) fail to regulate *his3* expression. However, the related pair that differ only in the upstream breakpoint (*his3*-Δ26 lacking −80 to −53 and *his3*-Δ27 lacking −80 to −73) are regulated properly. By subtraction, these results strongly suggest a second regulatory site that includes at least a subset of the sequences between −80 and −106. The upstream boundary of this site is defined further by deletions 32–35. These mutations remove sequences between −101 and −137, but do not prevent *his3* regulation. Some of them do seem, however, to have reduced basal levels of expression.

Thus, there seem to be at least two regulatory elements for *his3* expression separated by at least 30 bp. The downstream element maps between −32 and −52 and includes a region localized between −32 and −41, the upstream element maps somewhere between −80 and −100.

It is worth comparing the sequence of the *his3* regulatory regions defined here[4,7] with the nucleotide sequence of *his4*[8], a gene that is co-regulated with *his3*. The *his3* sequence between nucleotides −44 and −31, TATATAAAGTAATG, strongly resembles the *his4* sequence TATATAATAGA-TATG, which is located 46–60 bp before the mRNA start point[8]. These sequences contain 12 out of 14 identical base pairs if a single base pair is allowed to 'loop out'. With regard to the upstream regulatory element, the *his3* sequence between −97 and −90, ATGACTCT, shares 7 out of 8 base pairs with the *his4* sequence ATGACTAT located between −86 and −79. Furthermore, highly conserved variants of this sequence are found at *his3* sequences −139 to −132 (ATGCCTCG) and −178 to −171 (ATTACTCT) and *his4* sequences −137 to −130 (GTGACTCA) and −181 to −174 (GTGACTCC). This striking sequence homology supports the view that these regions are important in *his3* regulation.

Either of two simple models could account for the increased *his3* expression in conditions of amino acid starvation. In one, the gene interacts with positive regulatory factors that stimulate expression above the basal level; in the other model, the *his3* gene interacts with negative regulatory factors that maintain the basal level, and *his3* induction occurs by the inactivation of these factors. Proof of a positive and/or negative mode of *his3* regulation depends on the properties of purified regulatory factors with respect to the appropriate target DNAs. Nevertheless, evidence for the nature of *his3* regulation can be inferred from the phenotypes of regulatory mutations at the *his3* locus.

By definition, mutations at a positive regulatory site would result in basal levels of *his3* expression in all growth conditions. Indeed, five mutations (deletions 20, 22, 23, 24 and 28) have such a phenotype—this strongly suggests that at least some of the sequences deleted by these mutations comprise a positive regulatory site. The other four regulatory mutations result in constitutive expression below the basal level; thus it is difficult to assess their phenotype in terms of positive or negative control.

As for the possibility of negative control, so far I have obtained no *his3* mutations that result in constitutive *his3* expression at the induced level; these would be expected if the gene contains a site of negative regulation analogous to the

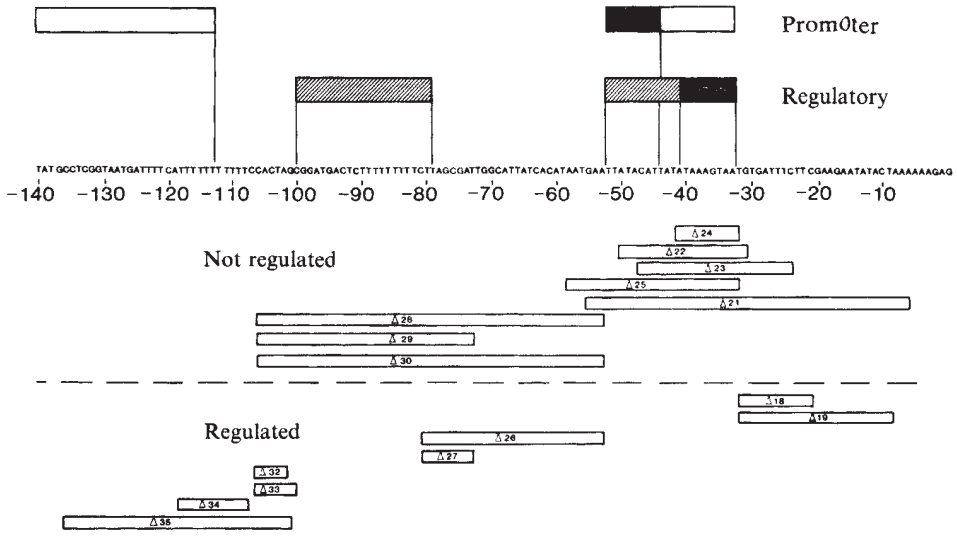

Fig. 2 *his3* promoter and regulatory elements. The open boxes on the top line indicate the extent of promoter elements defined previously[5,6] and the shaded boxes on the second line indicate the location of positive regulatory elements described here (the black regions within these boxes indicate sequences that are essential but not necessarily sufficient for promoter or regulatory function). These elements are placed with respect to the coding strand of the 5′ flanking region (nucleotides +1 to −140 defined with respect to the start of transcriptional initiation). The mRNA start was previously defined to an accuracy of ±1 bp[7]. The location of the deletion mutations used in this study are indicated below the DNA sequence. Deletions above the broken line do not allow for properly regulated *his3* expression, while the deletions below the line are regulated correctly.

operator of the *Escherichia coli* lactose operon. Nevertheless, it is possible that the five mutations expressing *his3* constitutively at the basal level could delete a negative regulatory site and simultaneously reduce promoter activity threefold. However, the principal advantage of deletion analysis is the ability to eliminate DNA sequences and hence genetic functions. For example, nine mutations described here eliminate regulatory behaviour, and 30 mutations (described in refs 5, 6) eliminate detectable *his3* expression. Thus, it seems unlikely that the five critical deletions fortuitously result in the same quantitatively minor defect, particularly considering the fact that nine deletions having full or partially reduced *his3* levels in normal conditions regulate the gene properly. Therefore, the mutants strongly support a positive control mechanism for *his3* expression as opposed to a negative one.

Table 2 Structures and phenotypes of *his3* derivatives

Fragment	*his3* allele	Upstream end point	Junction sequence	Downstream end point	Basal level	Induced level
Sc2605	Wild type	—	—	—	1.0	3.0
Sc2854	Δ18	−32	RI	−21	0.5	1.5
Sc2857	Δ19	−32	RI	−32	0.7	2.0
Sc3101	Δ20	−43		−31	1.1	1.0
Sc3102	Δ21	−56		−6	0.2	0.2
Sc3110	Δ22	−50		−31	1.2	1.3
Sc3111	Δ23	−47		−24	0.8	1.0
Sc3112	Δ24	−41	CC	−32	1.1	1.4
Sc3113	Δ25	−58	TCC	−32	0.2	0.2
Sc3121	Δ26	−80	RI	−53	0.9	2.7
Sc3122	Δ27	−80	RI	−73	0.6	1.5
Sc3125	Δ28	−106	RI	−53	1.1	1.0
Sc3126	Δ29	−106	RI	−73	0.2	0.3
Sc3129	Δ30	−119	RI	−53	0.2	0.1
Sc3138	Δ31	−106	RI	−109	1.0	3.1
Sc3159	Δ32	−106		−102	0.8	2.9
Sc3160	Δ33	−106		−100	1.2	3.2
Sc3161	Δ34	−118		−107	0.2	0.6
Sc3165	Δ35	−136		−101	0.5	1.2
Sc2884	Δ38	−80	RI	−32	<0.1	<0.1

All Sc fragments (first column) are derivatives of Sc2605 (see Fig. 1 legend). The *his3* deletion allele numbers and the end points with respect to the start of transcription (determined elsewhere by direct DNA sequencing) are indicated (columns 2 and 3). In some cases, there are junction nucleotides at the site of deletion (RI, *Eco*RI octanucleotide linker, 5′-GGAATTCC-3′). For each *his3* allele tested, 1 Ura^+ transformant (containing one integrated copy of the transforming DNA) and 1 His^+ transformant (in which the *his3*-532 allele is replaced with the transforming *his3* allele) were grown (see Fig. 1 legend for details). These always gave similar results; the values shown represent the averages of the two strains. To obtain the basal IGPD levels for a given mutation, strains were grown in complete medium. To obtain enzyme levels in starvation conditions, tryptophan was removed from the medium (see Table 1 legend for the methodology and reliability of the assay).

The data presented here suggest that the *his3* regulatory region consists of two separable positive control sites, each of which can be destroyed without apparently impairing promoter function. The conclusions represent the simplest formal description of the data, but have no bearing on the molecular mechanism by which *his3* expression is modulated in response to amino acid starvation. As induction of IGPD activity is accompanied by an increase in *his3* mRNA[9], it is likely that regulatory sites are involved in transcriptional initiation; this, however, has not been proved.

The key question is the nature of the 'positive regulatory factor'. Some possibilities are: (1) a protein product(s) that binds specifically to the sequence(s) defined as regulatory elements; (2) a conformational or covalent alteration in yeast RNA polymerase II that senses the regulatory sequences; (3) a structural feature of chromatin; and (4) a structural change in the *his3* transcript that affects its stability or translatability. Biochemical analysis of strains harbouring 'regulation-defective' *his3* alleles should be useful for distinguishing among these and other molecular models. Preliminary results with wild-type cells suggest that the chromatin structure and the mRNA structure are unchanged during amino acid starvation (ref. 4 and K.S., in preparation).

In considering molecular models for *his3* regulation, it is worth comparing the results obtained here with those of an analogous situation in *E. coli*. Amino acid starvation of *E. coli* results in increased expression of amino acid biosynthetic genes and decreased expression of those encoding stable RNA species such as the tyrosine-inserting tRNA. A mutation of this tRNA gene that changes sequences between the downstream promoter element (the Pribnow box) and the start of transcription retains promoter function but is regulated incorrectly[10]. Further, *in vitro* transcription of this gene using purified components suggests that regulation of the tyrosine tRNA promoter is mediated by the interaction of *E. coli* RNA polymerase with guanosine tetraphosphate[11]. It is striking that the *his3* regulatory site between −32 and −41 also maps between (or overlaps with) the TATA box region and the mRNA start. Whether this means that regulation in yeast occurs by a mechanism similar to that used by *E. coli* remains to be determined.

I thank Andrew Travers and Marvin Wickens for fruitful discussions. This work was supported by the Jane Coffin Childs Foundation and the National Institute of General Medical Sciences (grant GM 30186).

Received 26 May; accepted 30 September 1982.

1. Jacob, F., Ulmann, A. & Monod, J. *C.R. hebd. Seanc. Acad. Sci., Paris* **258**, 3125–3139 (1964).
2. Jacob, F. & Jacob, J. *J. molec. Biol.* **3**, 318–356 (1961).
3. Wolfner, M., Yep, D., Messenguy, F. & Fink. G. R. *J. molec. Biol.* **96**, 273–290 (1975).
4. Struhl, K. *Cold Spring Harb. Symp. quant. Biol.* **47** (Ch. 62 in the press).

5. Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1432–1436 (1981).
6. Struhl, K. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
7. Struhl, K. & Davis, R. W. *J. molec. Biol.* **152,** 553–568 (1981).
8. Donohue, T., Farabaugh, P. & Fink, G. *Gene* (in the press).
9. Struhl, K. & Davis, R. W. *J. molec. Biol.* **152,** 535–552 (1981).
10. Travers, A. A. *J. Bact.* **141,** 973–976 (1980).
11. Travers, A. A. *J. molec. Biol.* **141,** 91–97 1980).
12. Struhl, K. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **74,** 5255–5259 (1977).
13. Struhl, K. & Davis, R. W. *J. molec. Biol.* **136,** 309–332 (1980).
14. Cox, B. S. *Heredity* **26,** 211–232 (1971).
15. Scherer, S. & Davis, R. W. *Science* **209,** 1380–1384 (1980).
16. Struhl, K. thesis, Stanford Univ. (1979).

Isolation of a mouse pseudo tRNA gene encoding CCA—a possible example of reverse flow of genetic information

J. G. Reilly*, R. Ogden† & J. J. Rossi‡

* Division of Biology, City of Hope Research Institute, Duarte, California 91010, USA
† Agouron Institute, 505 Coast Boulevard South, La Jolla, California 92037, USA
‡ Department of Molecular Genetics, City of Hope Research Institute, Duarte, California 91010, USA

Transfer RNA genes offer an interesting system for studies of gene regulation. From experiments conducted with tRNA genes from eukaryotes, it seems that transcriptional regulatory regions are contained within the mature coding sequence of the DNA[1–4]. In tRNA genes of some yeast, *Xenopus* and *Drosophila*, intervening sequences located within the anticodon loop have been identified[5–8], thus adding a new dimension to tRNA expression. In contrast to the lower eukaryotes, relatively few mammalian tRNA genes have been studied. The genes obtained so far have been identified via cloned heterologous fragments, or labelled tRNA probes[9–11]. We sought to use synthetic oligodeoxyribonucleotide probes to isolate mouse phenylalanyl tRNA genes. We report here the isolation of a segment of DNA that is completely homologous to our 19-mer probe, as well as 38 nucleotides corresponding to the 3′ end of tRNA^{Phe}. The 5′ end of the gene is not present in our 1.7 kilobase clone. In addition, this pseudo gene has an encoded CCA. Since CCA is added post-transcriptionally to all known eukaryotic tRNAs for which genes have been identified, we believe this pseudogene to be representative of RNA encoded information going back into DNA.

We chose to isolate a mouse gene coding for tRNA^{Phe} by using as a probe a synthetic oligodeoxyribonucleotide 19-mer complementary to a portion of the known mammalian tRNA^{Phe} sequence[12,13] (see Fig. 1*a*). The strategy utilized was the screening of a partially digested mouse DNA library inserted at the *Hin*dIII site of pBR322. This plasmid library was used to transform *Escherichia coli* and selection was made for ampicillin resistant colonies. The colonies were screened for the tRNA^{Phe} gene containing inserts as described in Fig. 1 legend. Figure 1*b*, *c* shows autoradiographs of colony hybridizations with the synthetic probe. Four putative positives were obtained by the colony screening, two of which gave positive hybridization results on rescreening. These two clones appeared to be siblings as judged by DNA restriction (not shown). One of these plasmids, designated as pMTF3, was chosen for further studies.

Figure 2*a* presents ethidium bromide stains of restriction endonuclease digestions of pMTF3. Digestion with *Hin*dIII yielded a 3.4 kilobase pair (kbp) vector band, and 1.2 and 1.7 kbp insert bands. It is not known if the two *Hin*dIII inserts were due to ligation of partially digested DNA or were the result of ligation of two separate fragments. To test which of the two *Hin*dIII fragments contained the sequence homologous to the 19-mer probe, the DNA in the gel was transferred to nitrocellulose[14] and then hybridized with the ^{32}P-labelled synthetic 19-mer probe. These results demonstrated that only the 1.7 kbp *Hin*dIII fragment contained a sequence homologous to the probe (Fig. 2*a*). Digestion of pMTF3 with *Ava*I (CPyCGPuG) gave three stained fragments that had summated fragment lengths equal to undigested pMTF3. No hybridization with the 19-mer probe was seen in the *Ava*I digest (Fig. 2*a*). This same lack of hybridization was observed with *Hpa*II (CCGG) and *Sma*I (CCCGGG) digests (data not presented). This failure to detect hybridization with the 19-mer probe in the *Ava*I, *Hpa*II and *Sma*I digests is due to the presence of these sites in the middle of the region of complementarity to the 19-mer probe (Fig. 1).

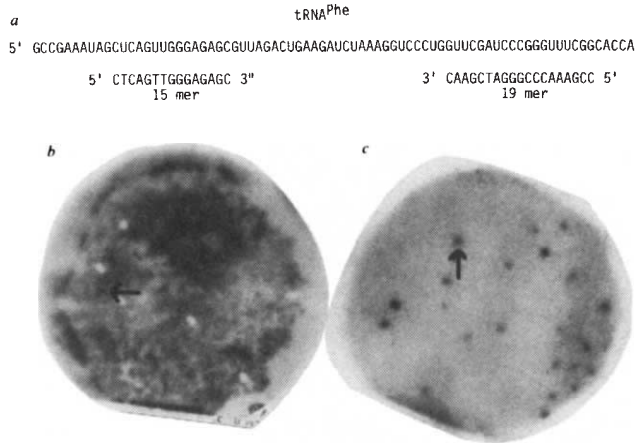

Fig. 1 Colony screening with a synthetic oligonucleotide complementary to tRNA^{Phe}. *a*, The sequence of mammalian tRNA^{Phe}[12,13] (above) with the 19 base oligodeoxyribonucleotide complementary to the 3′ region of the tRNA, and a 15 base oligonucleotide containing the same sequence as the dihydrouridine stem and loop region (below). Both oligodeoxyribonucleotides were synthesized using the phosphotriester method (38). *b*, Examples of colony screening using the synthetic 19 mer as a probe. Approximately 5 μg of mouse BALB/c liver DNA (partially digested with *Hin*dIII) was ligated with 1 μg of pBR322 DNA digested with *Hin*dIII and treated with alkaline phosphatase. This DNA mix was used to transform *E. coli* strain MC1061[39] from which approximately 1×10^5 transformants were obtained, about 90% of which were tetracycline sensitive. This represents approximately 15–20% coverage of the genome. These colonies were first replicated to dry, sterile nitrocellulose filters. The colonies on nitrocellulose were used to make replicas onto Whatman 541 paper[40]. The nitrocellulose replicas were incubated on L-agar plates containing ampicillin (30 μg ml^{-1}) and used as masters for retrieving positively hybridizing colonies. The Whatman 541 replicas were treated according to Gergen *et. al.*[40], and hybridized with 75 pmol (500×10^6 c.p.m.) ^{32}P-polynucleotide kinase labelled 19-mer. The hybridizations were carried out overnight at 45 °C in 3–5 ml per filter of 6×SET (30 mM Tris-HCl pH 7.5, 0.9 M NaCl, 6 mM EDTA), 5×Denhardt's medium, 0.5% Nonidet P40 following 2–3 hours of prehybridization in the same buffer. The filters were washed in 6×SET at progressively higher temperatures, beginning with 20 °C, until background hybridization was satisfactorily eliminated[41]. The final wash was at 45 °C. After washing, the filters were autoradiographed at −70°C on Kodak XRP-1 film with a DuPont Lightning Plus intensifying screen. *b*, A small area (approximately 1 cm) corresponding to the region of a positively hybridizing colony was scraped, diluted, replated and rescreened with the 19-mer probe. *c*, Positively hybridizing, isolated single colonies from the rescreening were used for DNA isolations.

Using standard double and triple restriction endonuclease digestions, a restriction site map for pMTF3 was constructed (Fig. 2*b*). Since the *Sma*I site (Figs 1*a*, 2*b*) is in the centre of the region of complementarity with the 19-mer probe, the DNA sequences including and surrounding this site were determined by the chemical cleavage method[15].

The DNA sequencing results are summarized in Fig. 3. From these data we find a region of perfect homology to the probe. Comparison to the known tRNA^{Phe} sequence (Fig. 3) demonstrates that 38 contiguous nucleotides of the tRNA^{Phe} 3′ region are included in the cloned segment. This homology diverges within the region corresponding to the tRNA^{Phe} anticodon loop.

0028-0836/82/460287—03$01.00