Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 501–504, January 1983
Genetics

DNA sequence polymorphism in an androgen-regulated gene is associated with alteration in the encoded RNAs

(DNA polymorphism/RNA polymorphism/hormonally regulated genes/recombinant inbred mouse strains)

ROSEMARY W. ELLIOTT AND FRANKLIN G. BERGER

Department of Molecular Biology, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, New York 14263

Communicated by Elizabeth Russell, October 21, 1982

ABSTRACT **We have used plasmid pMK908, whose cDNA insert corresponds to an androgen-inducible RNA from mouse kidney, as a probe to study both the complementary genomic DNA and the encoded RNA sequences in several inbred strains of mice. A polymorphism in the 908 structural gene, revealed by Southern blotting of *Hin*dIII-generated DNA fragments, was found to map near the *Gpi-1* and *Tam-1* loci on chromosome 7. The 908 structural gene has been termed *RP2*. The 908 RNAs constitute a sequence-related group displaying extensive size heterogeneity. This heterogeneity, reflected in the size distribution of the RNA on electrophoretic blots, is controlled by a genetic site that is tightly linked to *RP2*. Thus, associated with polymorphism at *RP2* is a change in the molecular size pattern of encoded transcripts.**

The mouse kidney offers several advantages in molecular studies of steroid-regulated genes (1, 2). Previously, we isolated a cDNA-containing plasmid, pMK908, that hybridizes to two kidney-specific testosterone-inducible RNAs, the 908 RNAs (3). These RNAs are induced approximately 7- to 10-fold in female mice after testosterone administration *in vivo*, and make up about 0.5% of the poly(A)-containing RNA in kidney after testosterone induction (3). At least one of the 908 RNAs encodes a 43,000-dalton polypeptide *in vitro* (3). Study of these RNAs should, therefore, be useful in addressing questions of tissue specificity and hormonal regulation of gene expression.

We have been using pMK908 to probe the structure and regulation of the corresponding sequences in the DNA and RNA from several inbred mouse strains. In this paper, we report the identification of polymorphism in the genomic DNA that hybridizes to pMK908 and have used this polymorphism to map the 908 structural gene, *RP2*, to chromosome 7. In addition, we present evidence that the 908 RNAs are actually a heterogeneous group of RNAs. The pattern of expression of these RNAs is controlled by a genetic determinant that is tightly linked to *RP2*.

MATERIALS AND METHODS

Animals. Inbred and wild-derived strains were either purchased from The Jackson Laboratory or provided by Verne Chapman of Roswell Park Memorial Institute. Recombinant inbred (RI) strains of the BXD set (4) and the AKXL set (5) were obtained from Benjamin Taylor of The Jackson Laboratory. Animals were used at 8–12 weeks of age. Testosterone treatment was by subcutaneous applications of pellets (6) for 2–7 days.

Isolation of Nucleic Acids. DNA was isolated from liver nuclei by the method of Jeffries and Flavell (7). Total RNA was extracted from kidneys by the guanidine hydrochloride protocol of Cox (8), as described by Labarca and Paigen (9).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

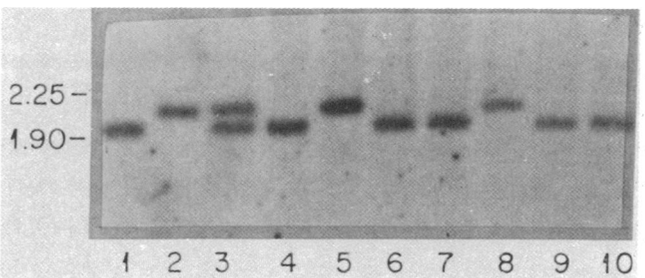

FIG. 1. Southern blot analysis of 908 DNA. Liver nuclear DNA was digested with endonuclease *Hin*dIII, subjected to agarose gel electrophoresis, blotted onto nitrocellulose, hybridized to ^{32}P-labeled pMK908 DNA, and observed by autoradiography. DNA was from strain C57BL/6J (lane 1), strain DBA/2J (lane 2), an F_1 hybrid derived from C57BL/6J and DBA/2J (lane 3), and the RI strains BXD-1, -2, -5, -11, -12, -13, and -14 (lanes 4–10, respectively).

DNA Blotting Analysis. DNA samples (15 μg) were treated with restriction endonuclease *Hin*dIII (Bethesda Research Laboratories) for 2–4 hr under conditions recommended by the supplier. Digested DNA was electrophoresed through a 0.8% agarose gel at 27 V for 17 hr in 40 mM Tris acetate, pH 7.8, containing 1 mM Na_2EDTA. After electrophoresis, DNA fragments were blotted onto nitrocellulose (10) and hybridized to pMK908 DNA that had been labeled to a specific activity of 1–2 $\times$ 10^8 cpm/μg by nick-translation in the presence of [α-^{32}P]dCTP and [α-^{32}P]dTTP (11). Hybridizing fragments were observed by autoradiography.

RNA Blotting Analysis. Total RNA (15 μg) was denatured in 10 mM sodium phosphate, pH 6.5, containing 2.2 M formaldehyde and 50% (vol/vol) formamide for 15 min at 60°C; the samples were electrophoresed through 1.5% agarose gels at 100 V for 4 hr in 10 mM sodium phosphate, pH 6.5, containing 2.2 M formaldehyde. RNAs were blotted onto nitrocellulose (12), hybridized to ^{32}P-labeled pMK908 DNA, and observed by autoradiography. RNA sizes were estimated by using the mouse ribosomal RNAs as markers.

RESULTS

Polymorphism in the 908 Structural Gene. To examine the nature of the structural gene encoding the 908 RNAs, we used ^{32}P-labeled pMK908 DNA to probe Southern blots (10) of mouse liver DNA after treatment with restriction nuclease *Hin*dIII. As shown in Fig. 1, DNA from strain C57BL/6J exhibited a single major band of length 1.9 kilobases (kb). Prolonged exposure of the autoradiogram revealed weak hybridizing bands at 9 kb and 0.3 kb (data not shown); the nature of these minor species is not known and has not been investigated

Abbreviations: RI, recombinant inbred; kb, kilobase(s); MUP, major urinary protein.

Table 1. Segregation of chromosome 7 markers in the BXD RI strains

	BXD RI strains																							
Locus	1	2	5	6	8	9	11	12	13	14	15	16	18	19	20	21	22	23	24	25	27	28	29	30	31	32
Coh	B	D	B	D	B	B	B	D	B	D	B	B	B	B	B	B	B	B	B	B	B	B	B	B	—	—
											×										×	×				
Gpi-1	B	D	B	D	B	B	B	D	B	D	D	B	B	B	B	B	B	B	B	B	D	D	B	B	D	D
										×																×
908 DNA	B	D	B	D	B	B	B	D	B	B	D	B	B	B	—	B	B	B	B	B	D	D	B	B	D	B
			×							×	×						×			×	×					×
Tam-1	B	D	D	D	B	B	B	D	B	D	B	B	B	B	B	B	D	B	B	D	B	D	B	B	D	D
		×		×		×			×		×		×				×		×							
Hbb	B	B	D	B	B	D	B	D	D	D	D	B	D	B	B	B	B	B	D	D	B	D	B	B	D	D

The 908 DNA phenotype was determined by Southern blotting of *Hin*dIII-digested DNA. Distributions for coumarin hydroxylase (*Coh*), glucose-phosphate isomerase (*Gpi-1*), tosylarginine methylesterase-1 (*Tam-1*), and β-globin (*Hbb*) have been determined previously (4, 14, 15). Inheritance of a C57BL/6J progenitor phenotype is indicated by B, and inheritance of a DBA/2J phenotype is indicated by D. A crossover between markers within a strain is indicated by ×. A — indicates the phenotype has not been examined.

further. In DNA from strain DBA/2J the 1.9-kb band was replaced by one of length 2.1 kb (Fig. 1). F_1 hybrid animals, generated by mating C57BL/6J with DBA/2J, contained both bands (Fig. 1). Thus, variation in the 908 structural gene exists between the two strains.

Genetic analysis of interstrain variation is facilitated by the use of RI strains. These are sets of inbred strains that are generated by inbreeding the F_2 progeny from a cross of two progenitor strains (13). Recombinant chromosomes containing various combinations of progenitor genes become fixed in homozygous form in RI strains; thus, by comparing the segregation pattern of alleles of a variant gene with the distributions of other variant loci whose map locations are known, it is possible to assign linkage of the variant gene to a particular site in the mouse genome.

To map the 908 structural gene, we determined the 908 DNA phenotype in each RI strain of the BXD set, which consists of 26 strains derived from progenitor strains C57BL/6J and DBA/2J (4). The results for some of the strains are shown in Fig. 1. Each strain contained either the 1.9-kb band characteristic of C57BL/6J or the 2.1-kb band characteristic of DBA/2J. The distribution for the fragment size was compared to the distribution for a variety of other loci that segregate in the BXD strains and that have been mapped. Table 1 shows distributions for several markers on chromosome 7 (4, 14, 15), and indicates cosegregation of the 908 DNA phenotype with *Gpi-1* in 24 strains, with *Tam-1* in 19 strains, and with *Coh* in 19 strains. Thus, the 908 structural gene, which we term *RP2*, is closely linked to *Gpi-1*.* Preliminary analysis of segregating mouse–hamster cell lines shows correlation of the presence of chromosome 7 with 908 DNA (unpublished data), supporting this linkage assignment. The data in total do not allow defining the precise location of *RP2*; however, preliminary evidence (unpublished data) suggests that it is located on the distal side of *Gpi-1*.

Polymorphism in the 908 RNAs. Using the RNA blotting technique of Alwine *et al.* (17), we previously suggested that pMK908 hybridizes to two RNAs, with lengths of 2.5 and 1.5 kb (3). The analysis has been extended by using higher-resolution protocols and the emerging picture appears more complex. Total RNA was subjected to agarose gel electrophoresis in the presence of 2.2 M formaldehyde after denaturation in 50% formamide and 2.2 M formaldehyde. RNA in the gel was blotted onto nitrocellulose and hybridized to ^{32}P-labeled pMK908 DNA. Fig. 2 shows the results for kidney RNAs from both untreated and androgen-treated females of strains C57BL/6J and DBA/2J and an F_1 hybrid derived from the two. The shorter exposure shown in Fig. 2*A* allows visualizing the androgen-induced RNAs. In strain C57BL/6J (lane 4), there is a very broad band having an average length of 2.5 kb and a discrete band with a length of 1.5 kb; this is consistent with the data reported earlier (3). RNA from strain DBA/2J (lane 5) shows a strikingly different pattern. There is a relatively discrete band at a length of about 2.8 kb along with the 1.5 kb band. F_1 hybrids have a pattern that appears to be a combination of the C57BL/6J and DBA/2J patterns (lane 6). These results suggest that the broad band in C57BL/6J RNA represents a heterogeneous group of hybridizing RNAs. DBA/2J RNA is less heterogeneous; in fact, each band in the DBA/2J pattern may be a single RNA species. Thus, the 908 RNAs appear to be heterogeneous in size. A polymorphism exists between strains C57BL/6J and DBA/2J that alters the heterogeneity pattern, as reflected in RNA blots.

A longer exposure of the autoradiogram (Fig. 2*B*) reveals the RNA patterns in untreated females (lanes 1–3). The patterns are similar to those for the androgen-treated animals. However, untreated females do appear to be slightly different from testosterone-treated animals of the same strain, as indicated by the presence of an altered molecular weight distribution of the RNA species. This suggests that the pattern of expressed 908 RNAs may be regulated, in part, by testosterone; however, more experiments will be necessary to substantiate this point.

Examination of the 908 RNA pattern in several other strains

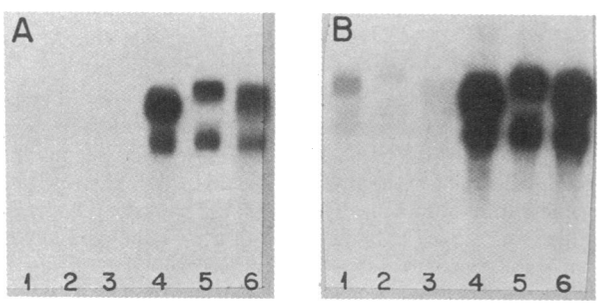

FIG. 2. RNA blot analysis of the 908 RNAs. Total kidney RNA was subjected to electrophoresis through an agarose gel, blotted onto nitrocellulose, hybridized to ^{32}P-labeled pMK908 DNA, and observed by autoradiography after exposure for 16 hr (*A*) or 3 days (*B*). RNA was from control (lanes 1–3) and testosterone-treated (lanes 4–6) female mice from strain C57BL/6J (lanes 1 and 4), strain DBA/2J (lanes 2 and 5), and an F_1 hybrid derived from C57BL/6J and DBA/2J (lanes 3 and 6).

* Nomenclature is according to recently proposed guidelines for naming genetic loci identified by DNA polymorphism (16).

Table 2. Segregation of the 908 DNA and RNA phenotypes in AKXL RI strains

	AKXL RI strains																	
Phenotype	5	6	7	8	9	12	13	14	16	17	21	24	25	28	29	37	38
908 DNA	A	L	A	L	A	L	L	L	A	A	L	A	A	A	L	A	A
908 RNA	A	L	A	L	A	–	L	L	A	A	–	A	–	–	–	–	A

Phenotypes were determined as described in the text. Inheritance of the AKR/J progenitor phenotype is indicated by A, and inheritance of the C57L/J phenotype is indicated by L. A — indicates the phenotype has not been examined.

extended these findings (data not shown). Each strain had the pattern of either C57BL/6J or DBA/2J. Thus, in both normal and hormone-treated mice, the 908 RNAs exhibited the characteristics of a regulated RNA family.

Correlation Between DNA and RNA Polymorphisms. For the strains initially surveyed, we observed that those containing the C57BL/6J type RNA pattern (i.e., C57BL/6J, BALB/cByJ, C3H/HeJ, and AKR/J) contain the 1.9-kb genomic fragment; conversely, strains exhibiting the DBA/2J type RNA pattern (i.e., DBA/2J, SWR/J, and C57L/J) contain the 2.1-kb fragment. This concordant distribution of the DNA and RNA polymorphisms suggested that the genetic determinants for the RNA pattern may be linked to, or within, the 908 structural gene on chromosome 7.

To further test the hypothesis that these traits are linked, we have examined the 908 DNA and RNA phenotypes in RI strains of the AKXL set, which are derived from progenitor strains AKR/J and C57L/J (5). Strain AKR/J is like C57BL/6J with respect to the 908 DNA and RNA variations; strain C57L/J is like DBA/2J. The results (Table 2) show that the two phenotypes segregate together in the 11 AKXL strains that have been analyzed and indicate that the locus or loci determining the 908 RNA pattern is genetically linked to the 908 structural gene.

We have examined the 908 DNA in 13 inbred and 42 RI strains, and we have determined the 908 RNA pattern in 32 inbred and 11 RI strains. For the 21 strains that have been analyzed for both phenotypes, there is complete concordance between the structural gene allele and the pattern of RNAs expressed. Thus far, no deviations from this correlation have been noted.

DISCUSSION

A powerful approach to the understanding of mammalian gene expression is the study of genetic variation in systems amenable to molecular analysis. In this report, we have used a cDNA probe to identify polymorphism for a hormone-sensitive gene and its corresponding transcripts in inbred mice.

A DNA polymorphism was identified by variation in the size of the major *Hin*dIII fragment that hybridizes to pMK908 DNA in a Southern blot. In the BXD RI lines, this variation in the 908 structural gene was found to cosegregate with the *Coh*, *Gpi-1*, and *Tam-1* loci on chromosome 7 (Table 1). A recently identified B-lymphocyte antigen, Lyb-8.2, is encoded by a gene, *Lyb-8*, that maps to this region also (18). We have called the 908 structural gene *RP2*, with the designations $RP2^b$ and $RP2^d$ for alleles carried by strains C57BL/6J and DBA/2J, respectively. Using equations that have been described (13), we can estimate that the *RP2* is 5.9 ± 3.6 centimorgans from *Coh*, 2.3 ± 1.8 centimorgans from *Gpi-1*, and 12.1 ± 6.7 centimorgans from *Tam-1*.

The number of 908 structural genes is not accurately known. However, present evidence is consistent with a single structural gene. First, only one major *Hin*dIII fragment hybridizes to pMK908 (Fig. 1); it is therefore unlikely that there are multiple, dispersed structural genes as is the case for the major urinary protein (MUP) structural genes (19). Second, the low intensity of the hybridization signal in the Southern blots makes it unlikely that there are a large number of gene copies. Finally, the fact that mouse strains contain one or the other DNA band in Southern blots, with no strain having both, suggests variation of a single gene. Confirmation of the structural gene number must await direct analysis of genomic DNA.

Examination of a variety of inbred strains led to the conclusion that the 908 RNAs are actually a sequence-related group of several species that exhibit extensive size heterogeneity; the lack of resolution on RNA blots (see Fig. 2) precludes determining the exact number of RNAs present. The pattern of expressed RNAs is determined by a genetic site that is associated with *RP2*. We have found no evidence for recombination between *RP2* and the genetic determinant for the RNA phenotype. Table 3 lists the alleles of *RP2* that are carried by the inbred strains examined to date. We do not know whether or not the RNA polymorphism is a consequence of the DNA polymorphism. The fact that strains carrying different *RP2* alleles exhibit size variation in both the 908 DNA and RNA suggests that the molecular bases for the two phenotypes may be the same. However, distinguishing this from the notion of coincident distribution of linked, yet separate, mutations must await more detailed analysis of the structure of the *RP2* region.

The existence of tightly linked structural and regulatory information has been well documented (20, 21) and is probably applicable to a number of systems. One that is regulated similarly to *RP2* is the MUP family in mouse liver. A gene cluster on chromosome 4 contains some 20–25 MUP structural genes; associated with the cluster is a locus, *Mup-a*, that controls the relative production of the different MUPs (19). *Mup-a* appears to function, as does *RP2*, by regulating the relative levels of different RNAs (22, 23). A major difference between the two systems lies in the fact that the heterogeneity displayed by the 908 RNAs is generated by what is most likely a single structural gene, whereas that within the MUP family derives from 20–25 genes. It will be interesting to determine if the mode of regulation characteristic of the MUP and 908 systems is a general feature of gene function within the mouse and other higher organisms.

The protein products of the 908 RNAs have not been completely characterized. A 43,000-dalton polypeptide has been detected by *in vitro* translation of purified total 908 RNA from $RP2^b$ as well as $RP2^d$ mice (ref. 3; unpublished data). In order to define the relationships between RNA heterogeneity and translational capacity, it is necessary to determine which mem-

Table 3. Alleles of *RP2* carried by inbred mouse strains

Allele		Strains
$RP2^b$	Inbred	A/J, AKR/J, BABT, BALB/cByJ, BUB/BnJ, C57BL/6J, C57BL/6ByJ, C57BL/10Sn, C3H/HeJ, C3H/HeHa, C3H/Str, CBA/J, Ha/ICR, JBT, LT/Sv, SEA/GnJ, SEC/1ReJ, YEBT, YBR
	RI	AKXL-5, -7, -9, -16, -17, -24, -25, -28, -37, -38 BXD-1, -5, -8, -9, -11, -13, -14, -16, -18, -19, -21, -22, -23, -24, -25, -29, -30, -32
$RP2^d$	Inbred	C57L/J, C57BR/cdJ, C58/J, CE/J, DBA/2J, DBA/2Ha, DE/Cv, I/LnJ, I/Str, MA/MyJ, 020, RF/J, SJL/J, SM/J, SWR/J, 129/J
	RI	AKXL-6, -8, -12, -13, -14, -21, -29 BXD-2, -6, -12, -15, -20, -27, -28, -31

Alleles were assigned by determining the DNA phenotype in 3 inbred and 31 RI strains, the RNA phenotype in 22 inbred strains, and both phenotypes in 10 inbred and 11 RI strains.

bers of the 908 RNA family encode this polypeptide. It is possible that the RNA polymorphism results in aberrant RNA processing, as has been noted for mutations in the human β-globin (24) and rat calcitonin (25) genes; such an alteration may or may not obliterate production of the corresponding proteins.

The molecular nature of the 908 RNA heterogeneity invites further exploration. Several genes encode heterogeneous RNAs, and several mechanisms for generating the size heterogeneity have been noted. Variation at the 3′ end of the RNA, reflecting the use of different poly(A) addition sites, appears within the mammalian immunoglobulin (26), chicken ovalbumin (27), and mammalian dihydrofolate reductase (28, 29) systems. Heterogeneity at the 5′ end may result from transcription of different 5′ exons (30) or from altered transcription initiation within a particular exon (31). Thus, the eukaryotic cell offers several strategies for creating heterogeneous RNAs. Further study of the 908 RNAs should be particularly informative due to the existence of the RNA polymorphism and its potential use in examining the regulation of transcription.

We thank Dr. Verne Chapman for making mice available from his colony. We also are grateful to Dr. Les Kozak of The Jackson Laboratory for providing facilities for isolating RNA from many of the strains listed in this report. Finally, we thank Dr. A. V. Le, Barbara Quarantillo, Colleen Hohman, and Leon Hall III for technical assistance. This work was supported by Grants GM-19521 and GM-28464 from the National Institutes of Health.

1. Swank, R. T., Paigen, K., Davey, R., Chapman, V. M., Labarca, C., Watson, G., Ganschow, R. E., Brandt, E. J. & Novak, E. (1978) *Rec. Prog. Horm. Res.* **34,** 401–436.
2. Bardin, C. W. & Catterall, J. F. (1981) *Science* **211,** 1285–1294.
3. Berger, F. G., Gross, K. W. & Watson, G. (1981) *J. Biol. Chem.* **256,** 7006–7013.
4. Traina, V. L., Taylor, B. A. & Cohen, J. C. (1981) *J. Virol.* **40,** 735–744.
5. Taylor, B. A. & Meier, H. (1976) *Genet. Res. Camb.* **26,** 307–312.
6. Swank, R. T. & Brandt, E. J. (1978) *Am. J. Pathol.* **92,** 755–772.
7. Jeffries, A. J. & Flavell, R. A. (1977) *Cell* **12,** 429–439.
8. Cox, R. A. (1968) *Methods Enzymol.* **12,** 120–129.
9. Labarca, C. & Paigen, K. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 4462–4465.
10. Southern, E. M. (1975) *J. Mol. Biol.* **98,** 503–517.
11. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113,** 237–251.
12. Derman, E., Kranter, K., Walling, L., Weinberger, C., Ray, M. & Darnell, J. E. (1981) *Cell* **23,** 731–739.
13. Taylor, B. A. (1978) in *Origins of Inbred Mice,* ed. Morse, H. C. (Academic, New York), pp. 423–438.
14. Skow, L. C. (1978) *Genetics* **90,** 713–724.
15. Wood, A. W. & Taylor, B. A. (1979) *J. Biol. Chem.* **254,** 5647–5651.
16. Elliott, R. W. (1981) *Mouse News Letter* **64,** 87.
17. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5350–5354.
18. Symington, F. W., Subbarao, B., Mosier, D. E. & Sprent, J. (1982) *Immunogenetics,* in press.
19. Bennett, K. L., Lalley, P. A., Barth, R. K. & Hastie, N. D. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1220–1224.
20. Berger, F. G., Breen, G. A. M. & Paigen, K. (1979) *Genetics* **92,** 1187–1203.
21. Paigen, K. (1979) *Annu. Rev. Genet.* **13,** 417–466.
22. Hastie, N. D., Held, W. A. & Toole, J. J. (1979) *Cell* **17,** 449–457.
23. Szoka, P. R., Gallagher, J. F. & Held, W. A. (1980) *J. Biol. Chem.* **255,** 1367–1373.
24. Orkin, S. H., Kazazian, H. H., Antonarakis, S. E., Goff, S. C., Boehm, C. D., Sexton, J. P., Waber, P. G. & Giardina, P. J. V. (1982) *Nature (London)* **296,** 627–631.
25. Rosenfeld, M. G., Lin, C. R., Amara, S. G., Stolarsky, L., Roos, B. A., Ong, E. S. & Evans, R. M. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1717–1721.
26. Early, P., Rogers, J., Davis, M., Calame, K., Bond, M., Wall, R. & Hood, L. (1980) *Cell* **20,** 313–319.
27. Heilig, R., Perrin, F., Gannon, F., Mandel, J. L. & Chambon, P. (1980) *Cell* **20,** 625–637.
28. Setzer, D. R., McGrogan, M., Nunberg, J. H. & Schimke, R. T. (1980) *Cell* **22,** 361–370.
29. Setzer, D. R., McGrogan, M. & Schimke, R. T. (1982) *J. Biol. Chem.* **257,** 5143–5147.
30. Young, R. A., Hagenbuchle, O. & Schibler, U. (1981) *Cell* **23,** 451–458.
31. Carlson, M. & Botstein, D. (1982) *Cell* **28,** 145–154.