Point mutation generates constitutive expression of an inducible eukaryotic gene

(arginase/Saccharomyces cerevisiae/transcriptional regulation)

ROBERTA A. SUMRADA* AND TERRANCE G. COOPER*

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, PA 15260

Communicated by Harland G. Wood, September 10, 1984

ABSTRACT We describe the analysis of two cis-dominant mutations that result in constitutive expression of the inducible CAR1 gene from yeast. One mutation results from insertion of a Ty element just upstream from the point where CAR1-specific transcription begins. The other mutation is a C-to-G transversion at position -153. Isolation of this point mutation, outside of the transcribed region of CAR1, suggests that expression of this gene is regulated at transcription. It also demonstrates the feasibility and usefulness of analyzing the regulatory sequences of eukaryotic genes on a nucleotide-by-nucleotide basis.

The control of gene expression in bacteria has been an area of intensive and productive study. In several cases, investigators have analyzed sequences regulating transcription one base at a time through the use of mutations located in the pertinent control region (1). These studies have provided an increasingly detailed picture of the molecular events involved with the expression of bacterial operons and control of this process. There is every reason to believe that a similar approach would yield equally detailed information about the control mechanisms of gene expression in eukaryotic cells. However, a paucity of cis-dominant control mutations in eukaryotic systems has raised questions about whether or not mutational analysis of eukaryotic regulatory sequences will be feasible. Two cis-dominant regulatory mutations have been reported to occur at the CAR1 locus in Saccharomyces cerevisiae (2). These mutations alter expression of the gene encoding arginase.

Arginase (CARI) and ornithine aminotransferase (CAR2) catalyze the first two steps of arginine degradation (3). The production of both enzymes has been shown to be inducible (3), and arginine has been identified as the native inducer (4). Enzyme production is also subject to nitrogen-catabolite repression when strains are grown in the presence of readily used nitrogen sources such as asparagine or glutamine. Two classes of recessive, regulatory mutations have been isolated. Strains carrying lesions at the arg80, arg81, or arg82 (argRI, argRII, and argRIII) loci are unable to induce production of arginase and ornithine aminotransferase (2, 5). On the other hand, strains with mutations at the car80 locus produced both enzymes constitutively (5). Mutations at car80 have been shown to be epistatic to those at the arg80-82 loci (2, 5).

We previously suggested that expression of the arginase gene was regulated at a point prior to the initiation of protein synthesis; transcriptional control was considered a likely candidate (6). To test this hypothesis, we cloned the *CAR1* gene and used it as a probe to measure the steady-state levels of *CAR1*-specific RNA (7). We found that *CAR1* mRNA increased dramatically when arginine was added to the culture medium of wild-type cells. This increase was not seen if as-

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.

paragine was added along with inducer (i.e., CAR1 mRNA production appeared to be sensitive to nitrogen catabolite repression). RNA derived from a car80 mutant contained a high constitutive level of CAR1 mRNA even when the mutant cells were grown in the absence of inducer. However, CAR1 mRNA dropped to nearly undetectable levels when the car80 culture was provided with a repressive nitrogen source such as asparagine. This result again demonstrated the presence of nitrogen repression on CAR1 gene expression and eliminated inducer exclusion as a basis for it. The qualitative differences in CAR1 mRNA levels, generated by alteration of the strain's genotype and growth conditions, led us to suggest that control of arginase production was predominantly exerted at the level of CAR1 transcription (6).

As one means of testing for the existence of *CAR1* transcriptional regulation, we have cloned the two *cis*-dominant mutant alleles mentioned above and determined their nucleotide sequences. The mutant lesions were each found to be located outside of the transcribed region.

Alteration of the *CAR1* induction and repression processes by mutations situated upstream from the beginning of the transcript is consistent with *CAR1* regulation being exerted at the transcriptional level. A preliminary report of this work has already appeared.[†]

METHODS

The strains of S. cerevisiae and Escherichia coli used in this work are listed in Table 1. Standard genetic methods were used throughout (10). Yeast carbon-base medium (Difco) was used for growth of transformants and enzyme assay. Nitrogen sources were provided at 0.1% final concentration, while double nitrogen sources were supplied at 0.05% each. Detailed procedures for cloning (7), DNA sequence analysis, and construction of the fusion plasmids will appear elsewhere. Yeast arginase activity was assayed by the methods of Whitney and Magasanik (4). β -Galactosidase activity was measured by the method of Rose and Botstein (11); units are expressed as nmol of O-nitrophenyl β -D-galactoside (ONPG) hydrolyzed per min/mg of protein.

RESULTS

Isolation and Nucleotide Sequence Analysis of the $CAR1-0^h$ and $CAR1-0^-$ Alleles. As a first step toward understanding how the $CAR1-0^h$ and $CAR1-0^-$ mutations generate constitutive expression of the CAR1 gene, we isolated them on chimeric plasmids. This was accomplished by the integration-excision method described by Stiles (12). The $CAR1-0^h$ excision plasmid derived from this procedure was isolated, and a

Abbreviation: bp, base pair(s).

^{*}Present address: Department of Microbiology and Immunology, University of Tennessee Center for the Health Sciences, Memphis, TN 38163.

[†]Sumrada, R. A. & Cooper, T. G., Eleventh International Conference on Yeast Genetics and Molecular Biology, Sept. 13–17, 1982, Montpellier, France, p. 92 (abstr.).

Table 1. Strains used in this work

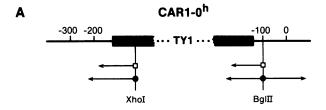
Strain	Genotype
S. cerevisiae	
RH218	MATa trp1 CUP1 gal2 SUC2 Mal-
7094a	MATα carl-0 ^h
7204b	MATα carl-0 ⁻
VT27	MATa trpl CUPl carl
M1360-21A and 7C	MATa trp1 his6 ura3-52 car1
M1394-17D	MATa trp1 ura3-52 car1-0
M1369-10D	MATα trp1 ura3-52 car1-0 ^h
E. coli	•
SF-8	hsdR hsdM recB recC lop-1 supE44 gal-96 Sm ^r leuB thi-1 thr ⁻
HB101	hsdR hsdM recA13 supE44 lacY1 proA2 Sm ^r leuB thi-1

The yeast strains used in this work have been described previously (7-9).

restriction map of it was constructed (data not shown). A portion of this map was found to be colinear with that of plasmid D15 (13), thereby indicating that the $CAR1-0^h$ phenotype was generated by insertion of a Ty sequence just upstream from the Bgl II site at the 5' end of the CAR1 gene.

The same procedure was used to clone the $CAR1-0^-$ allele (plasmid pRS20) of the CAR1 gene. The cloned $CAR1-0^-$ gene generated a typical mutant phenotype (constitutive arginase production that was partially resistant to nitrogen catabolite repression) when its expression was assayed following transformation of the plasmid into a $car1\ S.\ cerevisiae$ recipient. This observation indicated that the $CAR1-0^-$ allele was definitely contained on the fragment excised from the mutant genome. Subcloning experiments localized the mutation to a 2.45-kilobase region upstream from the beginning of the CAR1 gene coding region. A detailed restriction map of the excised fragment was prepared and was found to be precisely the same as its wild-type counterpart, indicating that the $CAR1-0^-$ phenotype did not arise from a gross, chromosomal rearrangement like the one seen with $CAR1-0^h$.

Structural analysis of the constitutive alleles was pursued to the nucleotide sequence level by using the strategy depicted in Fig. 1. A restriction map of plasmid pRS21 indicated



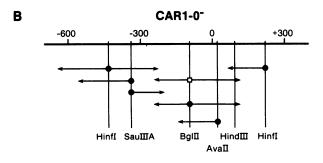


Fig. 1. Sequence strategy for the CARI mutant alleles. Vertical lines indicate the labeling sites used for sequencing the $CARI-0^h(A)$ and $CARI-0^-(B)$ mutant alleles. Arrows indicate the extent of sequence read from a single labeled site. \Box , Fragments 3'-end-labeled with Klenow fragment; \bullet , fragments 5'-end-labeled with polynucleotide kinase. Numbers are relative to the CARI ATG at position +1.

Since the $CARI-0^-$ lesion had not been localized, we determined the nucleotide sequence of the region adjacent to the 5' end of the mutant gene (positions +219 to -600) and found it to be absolutely identical to the wild type with only one exception. A guanidine nucleotide (cytidine in the sense strand) at position -153 of the wild type was replaced with a cytidine nucleotide in the mutant allele (Fig. 2).

Construction of CAR1-LACZ Fusion Plasmids. The finding of a single C-to-G transversion (position -153 of the sense strand; Fig. 3 and ref. 14) in the CAR1-0⁻ mutant prompted two questions: (i) could this base change be verified in an independent manner and (ii) was this alteration sufficient to generate the mutant phenotype? In other terms, was there a second lesion upstream of the 600-base-pair (bp) fragment whose sequence we determined? The first question was addressed directly. The wild-type CAR1 sequence in the vicinity of the CAR1-0⁻ mutation is 5' T-A-G-C-C-G-C-G-A 3'. The mutant sequence is 5' T-A-G-C-G-G-C-C-G-A 3'. This mutant sequence contains the recognition site for Xma III (5' C-G-G-C-C-G 3'). Therefore, the mutant allele was expected to possess a Xma III site that is missing from the wild-type allele. We tested this prediction by end-labeling a 526-bp Ava II-Ava II fragment (containing the region in question, positions -516 to +10 in Fig. 3) from wild-type and mutant plasmids. The two fragments then were digested with Xma III, and the digestion products were resolved on an acrylamide gel. Xma III did not cleave the wild-type fragment but did

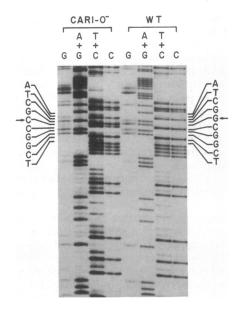


Fig. 2. Comparison of CAR1-0⁻ and wild-type sequences. Sequence reaction products from 5'-end-labeled CAR1-0⁻ DNA (Left) and wild-type DNA (Right) were resolved on a 10% acrylamide/urea gel. Arrows indicate the single base-pair transversion at position -153.

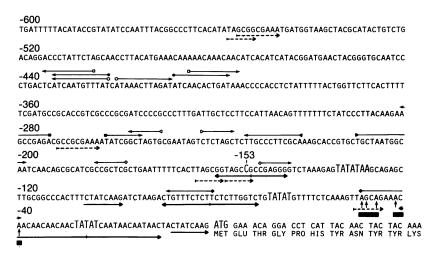


Fig. 3. The 5' regulatory region of the CARI gene. Paired arrows terminating in symbols above the sequence indicate the regions of inverted repeat homologies (20% mismatch). Arrows below the sequence not terminating in symbols indicate directly repeated sequences. Palindromic sequences are denoted by divergent arrows joined at the point of symmetry. Major 5' termini of CAR1 transcripts are indicated by vertical arrows (14). Thick bars denote the extent of 5' terminal heterogeneity on these mRNA species. All numbers are relative to the ATG at position +1. Symbols \Box , \Diamond , \Diamond , \triangleright , \diamond , \blacktriangleright , \blacktriangleleft , and \bullet identify paired sequences.

cleave the mutant fragment into pieces of approximately 370 and 150 bp (Fig. 4). This result indicates the presence of an *Xma* III site in the mutant fragment and, hence, a C-to-G transversion at position -153 in the *CARI-0*⁻ mutant DNA.

To determine whether the single base alteration we observed was, alone, sufficient to cause the mutant phenotype, we constructed CARI-LACZ fusions using wild-type and mutant DNA as the source of CAR1 sequences. The essential basis of this strategy is the fact that the wild-type and mutant sequences fused to LACZ were derived from fully sequenced fragments that differed by only the C-to-G transversion at position -153. The detailed strategy, which is depicted in Fig. 5, produced an in-frame fusion of the LACZ gene to a site 13 bp into the coding region of CAR1. The 5 end of the CAR1 fusion fragments was at position −516. After verifying the structures of fusion plasmids pRS45 and pRS46, they were transformed into yeast strain RH218. The respective transformants were isolated and grown under conditions in which the CAR1 gene should be expressed at an uninduced basal level (glutamate), an induced level (arginine), and an induced but nitrogen-catabolite-repressed level (arginine and asparagine) of arginase activity. The wild-type fusion (pRS46) exhibited a normal pattern of regulation for the fused CARI-LACZ gene (Fig. 6)—i.e., it was both inducible (compare Arg vs. Glu) and repressible (compare Arg vs. Arg + Asn). The mutant fusion (pRS45) exhibited a pattern of β -galactosidase regulation identical to that observed for arginase in the $CAR1-0^-$ mutant. β -Galactosidase was constitutively produced on glutamate, there was modest dere-

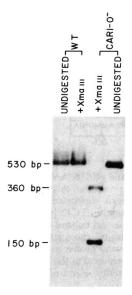


FIG. 4. Xma III digestion products from wild-type (WT) and CAR1-0⁻ DNA fragments. Ava II DNA fragments of 0.53 kilobase (positions -517 to +10 relative to ATG) were isolated from wild-type and CAR1-0⁻ DNA and 5'-end-labeled. They were then digested with restriction enzyme Xma III under supplier's reaction conditions. The digests were then resolved on a 7% acrylamide gel adjacent to undigested fragments and unlabeled standards.

pression when arginine replaced glutamate as the nitrogen source, and asparagine repressed β -galactosidase production by about half. This result indicated that the C-to-G transversion alone accounts for the phenotype of the $CAR1-0^-$ mutation.

A third fusion was constructed in a manner similar to that described in Fig. 5. Here, the objective was to assess whether sequences in the vicinity of the $CARI-0^-$ mutation (position -153) were necessary for CARI gene expression. Therefore, all of the nucleotide sequences 5' from the Bgl II

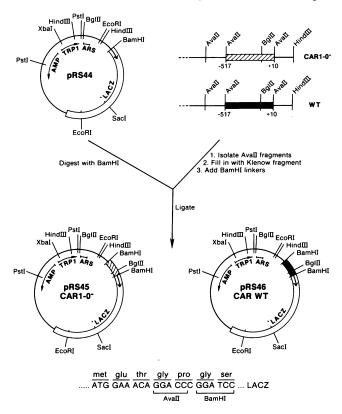


FIG. 5. Construction of $CARI-0^-$ and wild-type (WT) CARI-LACZ fusions. Ava II fragments (positions -517 to +10 relative to ATG) were prepared from wild-type and $CARI-0^-$ DNA and were made blunt-ended by the Klenow fragment of DNA polymerase I. BamHI molecular linkers (10-mers) were then added and trimmed by subsequent addition of DNA ligase and BamHI. The resulting BamHI fragments were ligated into plasmid pBR322 and then subcloned into the LACZ vector, pRS44, thereby producing the fusion plasmids (pRS45 and pRS46), which were identified by minipreparations. The resulting reading frame of the CARI-LACZ fusions is shown at the bottom of the figure.

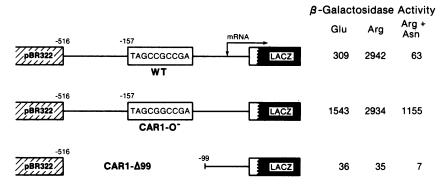


FIG. 6. β -Galactosidase activity of the CAR1-LACZ fusion plasmids. Plasmids pRS45 (CAR1-0⁻-LACZ fusion), pRS46 (WT CAR1-LACZ fusion) were transformed into yeast strain RH218 as described. A transformant from each plasmid was grown in the presence of the nitrogen source indicated and assayed for β -galactosidase activity. The plasmid designated CAR1- $\Delta 99$ contains a deletion extending from position -516 to position -100 relative to the ATG of the fusion plasmid.

site (position -100) were deleted. Position -99 is 50 bp upstream of the first *CAR1* mRNA start site and 31 bp 5' to the "TATA" box that probably functions in the expression of this gene. As shown in Fig. 6, deletion of the sequences between positions -516 and -99 results in nearly complete loss of β -galactosidase production, thus pointing to their necessity for *CAR1* gene expression.

While attempting to localize the CAR1 gene by subcloning, we had subcloned a CAR1-containing fragment that terminated at this Bgl II site (plasmid pRS18 in Fig. 7 and in fig. 2 of ref. 7) and had found this construction to complement a carl mutation. Arginase production supported by this subclone is shown in Fig. 7, where it is compared to a larger fragment (plasmid pTC42) containing several kilobases of DNA upstream of CAR1. The CAR1 gene, borne on plasmid pRS18, was inducible but to a lesser extent than was observed for the CAR1 gene borne on plasmid pTC42. Although induction was somewhat decreased, it was definitely present. The most straightforward explanation of this result is that expression of the CARI gene on plasmid pRS18 is dependent on vector sequences that are not present in the pRS47 ($carl-\Delta 99$) construction. We assessed this possibility by searching the tetracycline-resistance gene sequence on plasmid pBR322 (600 bp from the BamHI site toward the Ava I site) for homology to sequences upstream of the CAR1 ATG (positions -600 to +1 of *CAR1*). The sequence 5' T-A-G-G-C-C-A-G 3' was found at position -161 from the CARI ATG in plasmid pRS18, thus placing it in the TET gene of the plasmid. This sequence is very similar to the decanucleotide 5' T-A-G-C-C-G-C-C-G-A 3' found at position -157 in the wild-type CAR1 gene.

All of the above observations point to the importance of the sequence 5' T-A-G-C-C-G-A 3' in the regulation of CARI gene induction. If this conclusion is correct, one might expect to find the same sequence in the 5' regulatory region of the CAR2 gene because both arginase and ornithine aminotransferase production are similarly inducible (2, 3, 9). Therefore, we isolated the CAR2 gene and determined its complete nucleotide sequence (unpublished data). When we compared the 5' flanking sequences of the CARI and CAR2 genes, we observed two areas of homology: the sequence T-A-T-A-T at position -114 and the sequence 5' T-A-G-C-C-G-C-C-G-A 3' at position -177. The latter sequence is exactly the same one we observed in the wild-type CAR1 gene, and it is the sequence that contains the C-to-G transversion in the CAR1-0⁻ mutant. The position of the decanucleotide with respect to ATG is also approximately the same in both CAR1 and CAR2.

DISCUSSION

Several observations identify the sequence 5' T-A-G-C-C-G-C-C-G-A 3' as potentially important to the regulation of *CAR1* gene expression in *S. cerevisiae*. First, the sequence is found adjacent to the 5' end of both the *CAR1* and *CAR2* genes, which are known to be similarly inducible (2, 3, 9).

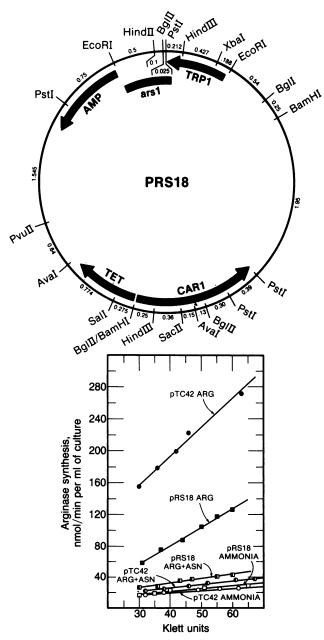


FIG. 7. (Upper) Restriction map of plasmid pRS18. Arrows indicate the direction of transcription of genes present on this plasmid. (Lower) Differential rate of arginase synthesis supported by plasmid-borne copies of the CARI gene. Plasmids pTC42 [wild-type CARI (7)] and pRS18 [CARI gene with sequences deleted upstream of the Bgl II site at position -100] were transformed into a carl yeast strain (VT27), and the transformants were grown on the indicated nitrogen sources. Cell samples were removed at the indicated cell densities and assayed for arginase activity. Klett units were measured with a green filter.

Deletion of the sequence resulted in a dramatic decrease in expression of a CAR1-LACZ fusion. More or less normal induction was then restored when the deleted sequences were replaced with those derived from the coding portion of the tetracycline-resistance gene of plasmid pBR322. A search of this gene revealed two sequences, 5' T-A-G-G-C-G-C-A-G 3' and 5' G-C-C-G-C-C-G 3', that are closely related to the decanucleotide mentioned above. Finally, mutation of a single nucleotide in this sequence resulted in constitutive expression of the CARI gene that was partially resistant to nitrogen catabolite repression. Although these observations point to a significant regulatory role for the sequence 5' T-A-G-C-C-G-C-G-A 3', it is important to recognize that the termini of this sequence were designated by the homology observed between CAR1 and CAR2 sequences. Whether this sequence alone is sufficient for the regulation of CAR1 gene expression remains to be directly tested.

The single base change shown to account for the constitutive CAR1-0⁻ phenotype generates an important insight into control of CAR1 gene expression. This observation along with the existence of unlinked recessive mutations, which also result in constitutive production of the CAR1 and CAR2 gene products, are characteristics expected for genes that are negatively controlled. While this suggestion does not exclude other more complicated models, it does serve as a useful premise on which to base future biochemical experiments. The possibility of negative control for the CAR genes contrasts with that observed for several other yeast genes that appear to be regulated in a positive manner (15, 16). cis-Dominant, constitutive mutants have been sought in many of these systems, but in all cases the searches were fruitless. In the case of several amino acid biosynthetic genes, Fink and his colleagues have shown that the pertinent regulatory sequence is represented in multiple, closely spaced copies adjacent to the 5' end of the genes (17, 18). Demonstration that only one copy of these sequences was required for expression of the gene explained the failure to isolate cis-dominant point mutations in these systems. In converse manner, the isolation of cis-dominant CAR1-0 and CAR2-0 mutations (2) indicates that the regulatory sequence of these genes is present only once in their flanking sequences. In agreement with this expectation, the sequence 5' T-A-G-C-C-G-C-C-G-A 3' is present only once in each gene.

The postulation of negative regulation for CAR gene expression is consistent with most but not all of the data available. An inconsistency arises in the explanation of data derived from the $\Delta 99$ fusion (plasmid pRS47) described in Fig. 6. If the mutated control sequence in the CAR1-0 strain is associated with a negative control process, then deletion of the sequence should produce the same phenotype as a point mutation. In contrast to this expectation, deletion of the sequences between positions -516 and -99 resulted in total loss of expression. This observation raises the distinct possibility of positive control and predicts the existence of a new class of recessive mutations that are epistatic to mutations at car80 and result in a loss of ability to induce CAR1 and CAR2 gene expression. According to this model, these mutations would reside in a gene that encodes a positive control element whose existence has heretofore escaped identification. Moreover, if such an element exists, it is pertinent to ask whether the CAR80 gene product regulates in a negative manner by inhibiting the function of the positively acting element. Put in another way, do the sequences that presumably interact with the CAR80 gene product overlap with those that interact with the putative positively acting regulatory element or, alternatively, do the regulatory elements interact with one another?

The fact that a single point mutation results in both constitutive expression of CARI and partial resistance to nitrogen-

catabolite repression argues that the target sites for these two regulatory processes either overlap or otherwise interact with one another. The observation also raises the alternative possibility of interactions occurring between the regulatory elements associated with induction and nitrogen-catabolite repression.

The position and phenotype of the CAR1-0⁻ mutation also address the level at which CAR1 gene expression is regulated. We have demonstrated that the CAR1-0 mutation dramatically alters arginase induction and its sensitivity to nitrogen-catabolite repression (7). The fact that the CAR1-0⁻ lesion has been shown to be a single point mutation situated 104 bp upstream from the start of CARI mRNA synthesis strongly suggests that both types of arginase regulation are exerted at transcription. We also have shown that the 5' terminus of CAR1-specific mRNA is the same regardless of the physiological or genetic state of the cells—i.e., the state of induction or repression in wild-type or control mutant cells (14). This observation eliminates the possibility of the mutated nucleotide being included in the transcript regardless of cell genotype or growth conditions used. Although these data do not prove that predominant regulation of the CAR1 gene occurs at the level of expression, they do provide substantial support for that hypothesis.

In a more general way, the observations presented here demonstrate the value of analyzing the sequences of putative eukaryotic control regions one base at a time. This approach has been highly instrumental in elucidating the molecular events associated with the control of prokaryotic gene expression. The identification of a C-to-G transversion as the molecular basis for the $CAR1-0^-$ mutation argues that the approaches applied so successfully to prokaryotic systems are likely to be equally productive when applied to eukaryotic systems.

The authors thank Dr. Leonard Guarente for providing the plasmid that contained the *LACZ* gene used in our fusion experiments, Dr. Robert Lawther for his help with the linker technologies, and Diane Platoniotis for verifying the recloned *CAR1-0*⁻ phenotype. We also appreciate the comments from other members of the laboratory who critically read and improved the manuscript. This work was supported by U.S. Public Health Service Grant GM-24383 from the National Institute of General Medical Sciences.

- Youderian, P., Bouvier, S. & Susskind, M. M. (1982) Cell 30, 843– 853.
- Dubois, E., Hiernaux, D., Grenson, M. & Wiame, J. M. (1978) J. Mol. Biol. 122, 383-406.
- 3. Middlehoven, W. J. (1970) Antonie van Leeuwenhoek 36, 1-19.
- Whitney, P. A. & Magasanik, B. (1973) J. Biol. Chem. 248, 6197–6202.
- 5. Messenguy, F. & Dubois, E. (1983) Mol. Gen. Genet. 189, 148-156.
- 6. Bossinger, J. & Cooper, T. G. (1977) J. Bacteriol. 131, 163-173.
- Sumrada, R. A. & Cooper, T. G. (1982) Mol. Cell. Biol. 2, 1514– 1523.
- Farabaugh, P. J. & Fink, G. R. (1980) Nature (London) 286, 352– 356.
- Deschamps, J., Dubois, E. & Wiame, J. M. (1979) Mol. Gen. Genet. 174, 225-243.
- 10. Fink, G. R. (1970) Methods Enzymol. 17, 59-78.
- 11. Rose, M. & Botstein, D. (1983) Methods Enzymol. 101, 167-180.
- 12. Stiles, J. (1983) Methods Enzymol. 101, 290-300.
- 13. Cameron, J. R., Loh, E. Y. & Davis, R. W. (1979) Cell 16, 739-751.
- Sumrada, R. A. & Cooper, T. G. (1984) J. Bacteriol. 160, 1078– 1087.
- Hinnebusch, A. G. & Fink, G. R. (1983) Proc. Natl. Acad. Sci. USA 80, 5374-5378.
- Guarente, L., Lalonde, B., Gifford, P. & Alani, E. (1984) Cell 36, 503-511.
- Donahue, T. F., Davies, R. S., Lucchini, G. & Fink, G. R. (1983) Cell 32, 89-98.
- Hinnebusch, A. G. & Fink, G. R. (1983) J. Biol. Chem. 258, 5238–5247.