Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 743–746, February 1988
Biochemistry

Specific interaction between a *Saccharomyces cerevisiae* protein and a DNA element associated with certain autonomously replicating sequences

(DNA-binding protein/repetitive and single-copy autonomously replicating sequences)

SHLOMO EISENBERG*, CHRIS CIVALIER*, AND BIK-KWOON TYE†

*Department of Microbiology, School of Medicine, The University of Connecticut Health Center, Farmington, CT 06032; and †Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853

Communicated by Mary J. Osborn, October 9, 1987

ABSTRACT **We have isolated a protein from *Saccharomyces cerevisiae* that binds specifically to a nucleotide sequence associated with the autonomously replicating sequence (ARS) ARS*120*, located in the telomeric region of a yeast chromosome. "Footprinting" analysis revealed that a 26-base-pair DNA sequence, 5′-CAAGTGCCGTGCATAATGATGTGGGT-3′, was protected by this protein from DNase I digestion. A plasmid containing 48 direct tandem repeats of this oligonucleotide was constructed and used to affinity-purify the binding activity. The purified protein, OBF1 (origin binding factor), showed specific binding to ARS*120*. The 26-base-pair OBF1-protected sequence was sufficient for the recognition and binding of the protein, since the mobility of a DNA fragment containing the synthetic binding site was retarded in agarose gels when incubated with OBF1. By performing competition experiments with a number of different ARSs, we showed that OBF1 binds tightly to some but not all ARSs. Interestingly, OBF1 does not appear to have a discernible affinity for ARS*1* or the ARSs associated with mating type loci, *HMLα* and *HMR*a, which are substrates for a DNA-binding activity reported by others. Since OBF1 appears to bind to DNA associated with a number of ARSs, we suggest that this protein may have a function related to ARS activity, perhaps in the initiation of DNA replication at selected ARSs.**

DNA replication, a tightly regulated process in the eukaryotic cell, occurs in the well-defined S phase of the cell cycle. The replication of DNA proceeds by multiple initiations along a chromosome, which results in the formation of multiple replicons. In *Saccharomyces cerevisiae*, the initiation site at each replicon is believed to be specific and to correspond to the autonomously replicating sequences (ARSs). An ARS is defined as a DNA fragment that when cloned into a plasmid supports the autonomous replication of the plasmid in yeast (1). Ample evidence suggests that the ARSs serve as origins for the initiation of plasmid replication. *In vitro* studies on the replication of ARS-bearing plasmids correlated the initiation sites for plasmid replication with the location of the ARSs (2–4). Similarly, electron microscopy of *in vivo* replicating intermediates of rRNA-encoding DNA (rDNA) and the 2-μm plasmid placed the initiation sites at locations of the ARS elements (5, 6).

Most ARSs are present as single-copy DNA in the genome. However, exceptions are found in two families of repetitive ARSs. They are the ARSs that are located on the tandemly repeating rDNA units on one arm of chromosome XII (7) and those that are associated with the telomeres of yeast chromosomes (8). The telomeric ARSs can be further divided into two classes, X and Y′, which are associated with moderately conserved and highly conserved subtelomeric sequences, respectively. All ARSs share a common 11-base-pair (bp) consensus sequence, which is essential but not sufficient for ARS function (9). Nucleotide sequences on either side of the core consensus are required. However, the amount of flanking sequences required varies depending on the ARS (10–12). It is not known what role these sequences may play in DNA replication. If these sequences function in the initiation of DNA replication, then, by analogy to prokaryotic and viral replication systems, we anticipate the existence of proteins that will interact specifically with DNA elements in the ARSs. This belief has prompted us to search for ARS-binding proteins in *S. cerevisiae*. Here we report the isolation of a protein (OBF1) identified by its binding activity to a DNA element associated with the X-telomeric ARS*120*.

MATERIALS AND METHODS

Yeast Strains. The DNA-binding protein OBF1 was partially purified from the *S. cerevisiae* strain 20B-12 (*MATα, trp1, gal2, pep4-3*), which was obtained from E. Phiziky (University of Rochester, NY).

Plasmids. Plasmids used, their ARS inserts, and plasmid sizes are as follows. L2-773 (ARS*120*), 8.7 kilobases (kb); YCp131C (ARS*131C*), 14.6 kb; YRp121 (ARS*121*), 13.2 kb; and YRp131S (ARS*131S*), 11.6 kb, were as described (13). YIp5-Δ637 (*HMLE* ARS), 5.9 kb, was provided by J. Broach (Princeton University, Princeton, NJ), and p120.23 (*HMRE* ARS, cloned in vector YIp-5), 5.8 kb, was provided by J. Abraham (California Biotechnology, Palo Alto, CA), p19ARS1 (ARS*1*), 3.4 kb, was constructed by inserting the *Hin*dIII/*Eco*RI fragment of YRp7 (containing ARS*1*) into the *Hin*dIII and *Eco*RI sites of the pUC19 polylinker; pSN120, 2.9 kb, was constructed by blunt-end ligation of the *Sac* II/*Nco* I fragment of ARS*120* (see Fig. 1) into the *Sal* I site of pUC19 using T4 DNA polymerase and ligase; and pBS120-3, 2.8 kb, contains three direct repeats of the following sequence: GTCGACGAGGCAAGTGCCGTGCATAATGATGTGGGTGCCTCGAG. The above oligonucleotide was prepared in a DNA synthesizer and contains the 26-bp sequence (boldface type) that is protected from DNase I digestion by the OBF1 protein, flanked on the 5′ and 3′ ends by the *Sal* I and *Xho* I restriction sites, respectively. This oligonucleotide was hybridized with its complement and was self blunt-end ligated to form a multimeric unit. The multimeric chain was subsequently treated with the restriction enzymes *Xho* I and *Sal* I, which generated single units of the synthetic binding site terminated by *Xho* I and *Sal* I protruding ends. These were self-ligated in the presence of *Xho* I and *Sal* I and cloned into the *Sal* I and *Xho* I sites of

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.

Abbreviation: ARS, autonomously replicating sequence.

plasmid pUC13X (14). We have selected a clone, pBS120-3, that contained three tandem repeats of the synthetic oligonucleotide in a direct head-to-tail orientation.

Preparation of the DNA-Binding Protein OBF1. The yeast cells were grown, harvested, and then lysed with zymolyase as described (15). A protein fraction was prepared by ammonium sulfate precipitation followed by chromatography through a DEAE-cellulose column as previously described (15). Fractions that were active in binding to ARS*120* DNA were further fractionated on a DNA-cellulose affinity column prepared as follows: 48 copies of the above synthetic oligonucleotide were inserted into plasmid pUC19 (provided by D. Lazinski, Health Science Center, University of Connecticut) by a procedure similar to that described (16). The DNA enriched for the synthetic binding site was coupled to cellulose by the method of Alberts and Herricks (17). DEAE-cellulose fractions containing the binding activity (15) were loaded on the DNA-cellulose affinity column at a salt concentration of 0.4 M NaCl, and the protein containing the binding activity was eluted with 0.6 M NaCl. This resulted in at least a 1000-fold-purified ARS*120* DNA-binding activity. A more detailed description of the purification will be published elsewhere.

DNA-Binding Reaction and Analysis of Protein–DNA Complexes. The DNA-binding reaction mixture (25 μl) contained 50 mM Hepes (pH 7.5), 10 mM $MgCl_2$, bovine serum albumin at 0.1 mg/ml, 10 mM dithiothreitol, 1–2 fmol of a ^{32}P-labeled DNA fragment, and an aliquot of the affinity-purified protein (OBF1). The reaction mixture was incubated at room temperature for 15 min, placed on ice, and then electrophoresed in the cold room in 1.4% agarose gels for 4 hr at a constant current of 75 mA. The gels were dried under a fan at room temperature and autoradiographed using intensifying screens and Kodak XAR-5 film.

Preparation of ^{32}P-Labeled DNA Fragments. Plasmid DNA was digested with a restriction enzyme (as indicated in the legends) followed by incubation with [α-^{32}P]dATP and reverse transcriptase. Labeled DNA was separated from unincorporated dATP by chromatography in Sephadex G-50. The DNA was precipitated with ethanol and incubated with a second restriction enzyme to generate the fragment of interest. DNA fragments were occasionally isolated preparatively by electroelution, following electrophoresis in agarose.

"Footprinting." DNA sequences protected by OBF1 from DNase I digestion were analyzed by a footprinting technique (18). The nucleotide sequence in the DNase I-resistant region was determined by the Maxam and Gilbert procedure (19).

RESULTS

DNA-Cellulose Affinity-Purified OBF1 Protein Binds to a Specific Nucleotide Sequence in ARS*120* DNA. We have previously identified, using an agarose gel mobility retardation assay, a protein in crude extracts that binds specifically to a nucleotide sequence contained within ARS*120* DNA (15). By footprinting analysis, we found that a 26-bp nucleotide sequence, CAAGTGCCGTGCATAATGATGTGGGT, was protected from DNase I digestion by a protein in crude extracts. This nucleotide sequence is located 200 bp 3′ to the ARS*120* core consensus sequence (Fig. 1). We constructed a plasmid containing 48 direct tandem repeats of the 26-bp sequence. The plasmid was coupled to cellulose, and the DNA-cellulose was used for the purification of the binding activity. Protein–DNA complex formation was detected by the gel mobility retardation assay as described in *Materials and Methods*. Unlike crude extracts, the purified OBF1 fraction did not require addition of carrier DNA to the reaction mixture to demonstrate a specific protein–DNA interaction. The purified protein bound specifically to the

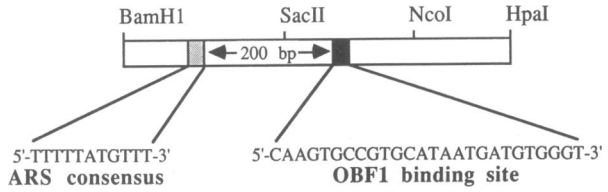

FIG. 1. Schematic presentation of ARS*120*. The diagram represents a restriction map of a DNA fragment isolated from YRp120 (7). This fragment (0.57 kb) promotes autonomous replication of plasmids in yeast. The arrows indicate the relative positions of the ARS core consensus sequence (5′-TTTTTATGTTT-3′) and the binding site (5′-CAAGTGCCGTGCATAATGATGTGGGT-3′) for OBF1.

ARS*120* *Sac* II/*Nco* I fragment (Fig. 2*A*) in the absence of carrier DNA. Binding to the ^{32}P-labeled ARS*120* *Sac* II/*Nco* I fragment could be completely blocked by competition by the addition of a 20-fold molar excess of plasmids bearing either the intact ARS*120* or the *Sac* II/*Nco* I fragment (Fig. 2 *B* and *C*). In contrast, plasmids pUC19 (Fig. 2*A*) and plc5 (data not shown), which do not contain ARS DNA, were ineffective in competition, indicating that the protein–DNA interaction observed in Fig. 2*A* is specific. This was further demonstrated by the footprint analysis shown in Fig. 3. Titration of the protein in the binding reaction revealed that the 26-bp DNA sequence in the *Sac* II/*Nco* I fragment was protected from DNase I digestion, confirming the specificity of the protein–DNA interaction.

The Information Needed for the Specific Protein–DNA Interaction Is Entirely Included in the Synthetic 26-bp Nucleotide Sequence. The fact that OBF1 could be isolated by its affinity to a plasmid enriched for the 26-bp binding site suggests that the DNA sequence necessary for this interaction is included within the 26-bp nucleotide sequence. We have proven that this indeed is the case by the experiment described in Fig. 4. A plasmid (pBS120-3) was constructed containing three direct tandem repeats of the 26-bp nucleotide sequence. The ^{32}P-labeled DNA fragment containing the three repeats was incubated with increasing amounts of OBF1 followed by electrophoresis in agarose gels. The DNA

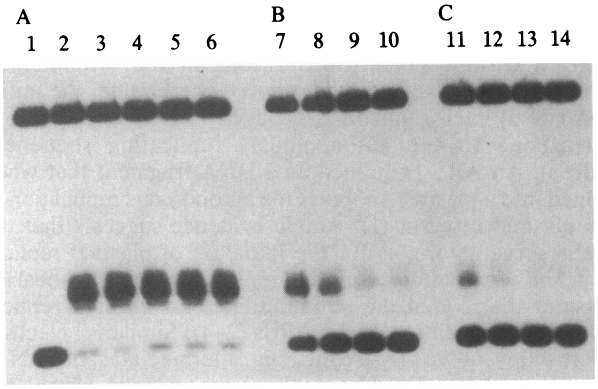

FIG. 2. Specific interaction of OBF1 with ARS*120* DNA. Plasmid pSN120 was digested with the *Hin*dIII and *Xba* I followed by end-labeling with [α-^{32}P]dATP. The labeled DNA (1.3 fmol) was incubated with OBF1 (0.5 μl) and competing DNA and was analyzed for the formation of protein–DNA complexes as described in *Materials and Methods*. (*A*) Lane 1, no OBF1. The bottom band contains the *Sac* II/*Nco* I fragment of ARS*120* DNA (see Fig. 1). Lane 2, 5 μl of OBF1 without competing DNA. Lanes 3–6, competition with plasmid pUC19 DNA at a 4-, 8-, 20-, and 40-fold molar excess, respectively. (*B*) Lanes 7–10, same as lanes 3–6 in *A* except that the competing DNA was the plasmid pSN120. (*C*) Lanes 11–14, same as lanes 3–6 in *A* except that the competing DNA was the plasmid L2-773.

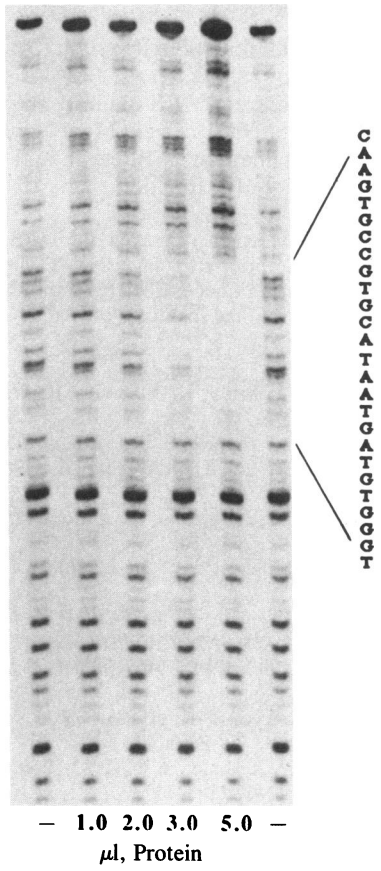

FIG. 3. Footprinting analysis of the OBF1-binding site. Plasmid pSN120 DNA was digested with *Hin*dIII followed by labeling with reverse transcriptase and [α-^{32}P]dATP. The labeled DNA was purified and was further digested with *Eco*RI. The DNA fragment, which included the *Sac* II/*Nco* I region of ARS*120* was isolated preparatively using an IBI electroeluter. DNA-binding reactions included 5 fmol of labeled DNA and increasing amounts of protein as indicated. DNase I digestion of the individual reactions and DNA sequence analysis were carried out as described in *Materials and Methods*. The DNA sequence analysis of the region that was protected from DNase I digestion revealed the sequence CAAGTGCCGTGCATAATGATGTGGGT.

fragment was retarded to assume three new positions, which is consistent with each position being represented by a fragment in which either one, two, or all three binding sites were occupied by OBF1. At high concentrations of protein, all input DNA was shifted to the most retarded position, suggesting that the three binding sites were saturated with OBF1.

OBF1 Binds to a Number of ARSs. Although the initial identification of the DNA-binding protein OBF1 was carried out with the X-telomeric ARS*120*, it was of interest to determine if this protein could recognize and bind to other ARSs as well. We have carried out competition experiments using plasmids containing different ARSs as described in Fig. 5. The ARSs tested were ARS*121* and ARS*1* (two single-copy DNA ARSs), ARS*120* and ARS*131C* (two class X-telomeric ARSs that show extensive sequence similarity to each other), ARS*131S* (a class Y′-telomeric ARS that contains no detectable sequence similarity to the class X-telomeric ARSs), and the *HMRE* and *HMLE* ARSs (two ARSs located in the silent mating type loci *HMR***a** and *HML*α, respectively). In each experiment, a ^{32}P-labeled ARS*120* *Sac* II/*Nco* I fragment was incubated with OBF1 in the presence of a 4–40 molar excess of competing plasmid DNA. When added plasmid was competing for binding, a gradual decrease in the amount of labeled protein–DNA

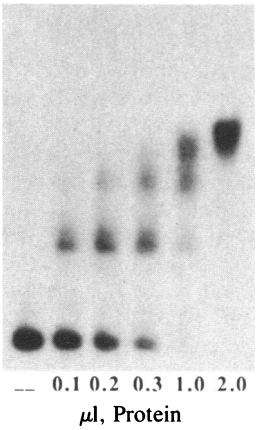

FIG. 4. A synthetic oligonucleotide contains all the information required for protein binding. Plasmid pBS120-3 was digested with the *Hin*dIII followed by labeling with reverse transcriptase and [α-^{32}P]dATP. The labeled DNA was purified and was digested with *Eco*RI. The labeled DNA fragment containing three direct repeats of the synthetic binding site was isolated preparatively and was used in the binding reactions. The binding reactions included in addition to labeled DNA (1.5 fmol) the amount of OBF1 protein indicated. The analysis of protein–DNA complex formation was carried out as described in *Materials and Methods*.

complex along with a concomitant increase in the intensity of the original labeled DNA fragment was observed. Thus, ARS*121* (Fig. 5*A*) and ARS*131C* (Fig. 5*B*) effectively competed with labeled ARS*120* DNA for the OBF1 protein. Almost complete competition was achieved by a 20- to 40-fold molar excess of competing DNA. On the other hand, ARS*131S* (Fig. 5*C*) appeared to be a weaker competitor. A 20- to 40-fold molar excess of the plasmid YRp131S resulted in about 50% competition. In contrast, plasmids bearing ARS*1*, *HMRE* ARS, and *HMLE* ARS competed poorly, if at all, for OBF1 (Fig. 5 *D–F*). These results are consistent with the interpretation that OBF1 has differential affinity for different ARSs. ARSs that have a high affinity for OBF1 included single-copy (ARS*121*) as well as repetitive (ARS*120* and ARS*131C*) ARSs.

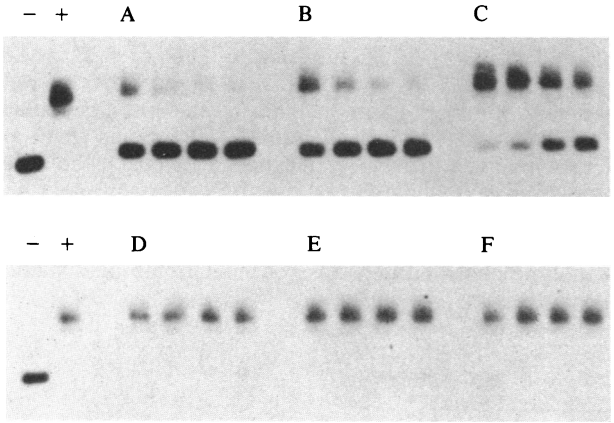

FIG. 5. DNA binding competition reactions with plasmids bearing various ARS sequences. ^{32}P-labeled DNA fragment containing the *Sac* II/*Nco* I ARS*120* DNA fragment was prepared as described in Fig. 3. The condition for DNA-binding reactions, competition, and DNA–protein complex analysis were as described in Fig. 2. – and + lanes represent reactions without and with OBF1, respectively. (*A*) DNA binding reactions in the presence of increasing (left to right) amounts (4-, 8-, 20-, and 40-fold molar excess) of plasmid YRp121. (*B–F*) The same as *A* except that the competing DNA was plasmid YCp131C, YRp131S, p19ARS1, p120.23, or YIp5-Δ637, respectively.

DISCUSSION

We have identified and purified by DNA-cellulose affinity chromatography a protein, OBF1, that binds to a specific sequence in the X-telomeric ARS*120*. This nucleotide sequence, protected from DNase I digestion by OBF1, was found to be present in all other X-telomeric ARSs that have been characterized (ref. 20 and V. Chang, C. Chan, and B.-K.T., unpublished observations). Indeed, the X-telomeric ARSs that we have tested in the binding reaction, ARS*131C* (Fig. 5) and ARS*131A* (data not shown), bound tightly to the protein. Thus, it is tempting to speculate that the protein may have a role in the structural organization of telomeric sequences. However, we found that OBF1 did not bind to X-telomeric ARSs exclusively. It bound tightly to the single-copy ARS, ARS*121*, and more weakly to the class Y-telomeric ARS, ARS*131S*. Therefore, we believe that OBF1 may have a function related to ARS activity. Such a protein may play a role in the assembly or be part of a protein complex responsible for the initiation of replication at the ARS. Alternatively, OBF1 might have a regulatory role in determining the efficiency of initiation of a particular ARS sequence. This in turn may determine the temporal initiation of different sets of replicons, which has been suggested to occur during the S phase of the *S. cerevisiae* cell cycle (21). In ARS*120*, the binding site is located 200 bp 3′ to the core consensus sequence. It is of interest to determine the relative location of the OBF1-binding site to the ARS core consensus sequence in ARS*121* and ARS*131S*. Recently, we found that the ARS function and the OBF1-binding site in ARS*121* both reside within a 0.48-kb DNA fragment (S. Francesconi, S. Walker, and S.E., unpublished work). A detailed DNA analysis of the above ARS may shed light on the functional relationship of the OBF1-binding site and ARS activity.

We also found that plasmids bearing ARS*1*, *HMRE* ARS, and *HMLE* ARS DNA did not appear to compete significantly with ARS*120* in the binding reaction. Therefore, we presume that these ARS sequences do not contain a binding site for OBF1. Recently, Shore *et al.* (22) reported the identification of two DNA-binding proteins. SBF-B binds to a site 3′ to the core consensus in ARS*1* and *HMRE* ARS, similar to the relative position of the OBF1-binding site in ARS*120*. The SBF-E protein interacts with the E box in the "silencer" of the *HMLα* and *HMRa* loci. In light of the fact that the plasmids p19ARS1, YIp5-Δ637, and p120.23 bearing ARS*1*, *HMRE* ARS, and *HMLE* ARS, respectively, were ineffective competitors in the binding reaction, it is unlikely that SBF-B and OBF1 are the same. If SBF-B and OBF1 are factors involved in DNA replication, it is conceivable that the two proteins may be functionally equivalent parts of a protein complex involved in the initiation of replication at an ARS. Similar suggestions have been made to explain the phenotype observed in a class of minichromosome maintenance-defective mutants (Mcm^-) that seem to affect the function of some but not all ARSs (23, 24). The results reported here and by Shore *et al.* (22) on the substrate specificity of the different ARS-binding proteins and the ARS specificity exhibited by certain Mcm^- mutants suggest that ARSs may be inherently different and may require different protein factors for function.

We thank Clarence Chan for making the ARS*120* nucleotide sequence data available prior to publication. This work was supported in part by American Cancer Society Grant MV-368 awarded to S.E. and in part by National Institutes of Health Grant GM34190 grant awarded to B.-K.T.

1. Stinchcomb, D. T., Struhl, K. & Davies, R. W. (1984) *Nature (London)* **289**, 39–43.
2. Celniker, S. E. & Campbell, J. L. (1982) *Cell* **31**, 201–213.
3. Kojo, H., Greenberg, B. D. & Sugino, A. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 7261–7265.
4. Jazwinski, M. S., Niedzwiecka, A. & Edelman, G. M. (1983) *J. Biol. Chem.* **258**, 2754–2757.
5. Saffer, L. D. & Miller, O. L., Jr. (1986) *Mol. Cell. Biol.* **6**, 1148–1157.
6. Newlon, C. S., Devenish, R. J., Suci, P. A. & Roffis, C. J. (1981) *ICN-UCLA Symp. Mol. Cell. Biol.* **22**, 501–516.
7. Saffer, L. D. & Miller, O. L., Jr. (1983) *J. Cell Biol.* **97**, 111a (abstr.).
8. Chan, C. S. M. & Tye, B.-K. (1983) *Cell* **33**, 563–573.
9. Broach, J. R., Li, Y. Y., Feldman, J., Jayram, M., Abraham, J., Nasmyth, K. A. & Hicks, J. B. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 1165–1173.
10. Tschumper, G. & Carbon, J. (1980) *Gene* **10**, 157–166.
11. Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E. & Campbell, J. L. (1984) *Mol. Cell. Biol.* **4**, 2455–2466.
12. Kearsy, S. (1984) *Cell* **37**, 299–307.
13. Chan, C. S. M. (1985) Dissertation (Cornell Univ., Ithaca, NY).
14. Xiao, H. & Lis, J. T. (1986) *Mol. Cell. Biol.* **6**, 3200–3206.
15. Eisenberg, S. & Tye, B.-K. (1987) *ICN-UCLA Symp. Mol. Cell. Biol.* **47**, 391–401.
16. Rosenfeld, P. J. & Kelly, T. J. (1986) *J. Biol. Chem.* **261**, 1398–1408.
17. Alberts, B. & Herrick, G. (1971) *Methods Enzymol.* **21**, 192–217.
18. Johnson, A. D., Meyer, B. J. & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5061–5065.
19. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
20. Button, L. L. & Astell, C. R. (1986) *Mol. Cell. Biol.* **6**, 1352–1356.
21. Fangman, W. L., Hice, R. H. & Chlebowicz-Sledziewska, E. (1983) *Cell* **32**, 831–838.
22. Shore, D., Stillman, D. J., Brand, A. H. & Nasmyth, K. A. (1987) *Eur. J. Biochem.* **6**, 461–467.
23. Maine, G. T., Sinha, P. & Tye, B.-K. (1984) *Genetics* **106**, 365–385.
24. Gibson, S., Surosky, R., Sinha, P., Maine, G. & Tye, B.-K. (1987) *ICN-UCLA Symp. Mol. Cell. Biol.* **47**, 341-354.