Table 2 Amino-acid sequences at the α-lactalbumin elbow

Protein	79	80	81	82	83	84	85	86	87	88
α-lac, human	Lys	Phe	Leu	Asp	Asp	Asp	Ile	Thr	Asp	Asp
α-lac, bovine	Lys	Phe	Leu	Asp	Asp	Asp	Leu	Thr	Asp	Asp
Lys, hen	Ala	Leu	Leu	Ser	Ser	Asp	Ile	Thr	Ala	Ser
Lys, human	Ala	Leu	Leu	Gln	Asp	Asn	Ile	Ala	Asp	Ala
Lys, horse	Lys	Leu	Leu	Asp	Glu	Asn	Ile	Asp	Asp	Asp
Calcium ligands in α-lac	+			*		+			*	*

The residue numbering is based on human α-lactalbumin. Add three to these numbers to get the classical hen lysozyme numbering.
* Residues that contribute a side-chain calcium ligand; + residues that contribute a carbonyl oxygen ligand.

elbow are relevant to protein evolution. The divergence of α-lactalbumin from lysozyme involves the formation of a novel enzyme system, and seems to be related to the formation of an extremely well developed Ca^{2+} binding ability, a surprising innovation. We would argue against an alternative evolutionary interpretation, that an ancestral lysozyme possessed Ca^{2+} binding activity and the current lysozyme structure retains a 'memory' of this, because the proposed serial homologues (phage, goose and hen type lysozymes[27]) do not possess this loop as a common feature. It is an interesting coincidence, however, that the loop of phage lysozyme once proposed as a calcium binding loop is probably homologous to a region fairly close to the elbow in α-lactalbumin[27,28]. Clearly the evolution of a novel functional property of the α-lactalbumin molecule has not come about through exon shuffling or any sophisticated structural modifications. It seems instead to be a structurally localized property that arose principally through simple amino-acid substitutions at three positions to produce a calcium binding site of exquisite specificity. While other cases of divergence of structures to fulfil radically different functions are known (for example, the various functionally diverse immunoglobulin domains), here the process is so recent that we feel that we have almost caught evolution *flagrante delicto*. Naturally α-lactalbumin may lend support to the rather simplistic approach to protein engineering that considers function to be a local property. Such a hypothesis can be tested easily in this case by investigating the effects of mutagenic modification of the lysozyme loop in an attempt to produce a calcium binding site (as may have already been accomplished naturally in the horse milk lysozyme). We would be pleased if our observations stimulate such studies and studies on the biological significance of calcium binding to the lactose synthetase. Although the lactose synthetase complex is not cytosolic, it is interesting to see essentially intracellular calcium binding at a site with no homology to the EF-hand. At any event it is encouraging to see nature innovating rather than restricting itself to spare-functional-part protein design.

This work was supported by the MRC and SERC.

Received 14 April; accepted 18 August 1986.

1. Levine, B. A. & Williams, R. J. P. in *Calcium and Cell Function*, Vol. II (ed. Cheung, W. Y.) 1-38 (Academic, New York, 1982).
2. Kretsinger, R. H. & Nockolds, C. E. *J. biol. Chem.* **248,** 3313-3326 (1973).
3. Brew, K. & Campbell, P. N. *Biochem. J.* **102,** 258-269 (1967).
4. Aschaffenburg, R. & Drewry, J. *Biochem. J.* **65,** 273-277 (1957).
5. Qasba, P. K. & Safaya, S. K. *Nature* **308,** 377-380 (1984).
6. Hiraoka, Y. *et al. Biochem. biophys. Res. Commun.* **95,** 1098-1104 (1980).
7. Mitani, M. *et al. J. biol. Chem.* **261,** 8824-8829 (1986).
8. Musci, G. & Berliner, L. J. *Biochemistry* **24,** 6945-6948 (1985).
9. Hiraoka, Y. & Sugai, S. *Int. J. Pept. Prot. Res.* **23,** 535-542 (1984).
10. Herzberg, O., Moult, J. & James, M. N. G. *J. biol. Chem.* **261,** 2638-2644 (1986).
11. Aschaffenburg, R. *et al. J. molec. Biol.* **67,** 525-528 (1972).
12. Fenna, R. E. *J. molec. Biol.* **161,** 203-210 (1982).
13. Aschaffenburg, R. *et al. J. molec. Biol.* **127,** 135-137 (1979).
14. Smith, S. G. *et al. Biochem. J.* (submitted).
15. Hall, L. *et al. Nucleic Acid Res.* **10,** 3503-3515 (1982).
16. Shewale, J. G., Sinha, S. K. & Bew, K. *J. biol. Chem.* **259,** 4947-4956 (1984).
17. Kuwajima, K., Hiroshima, Y. & Sugai, S. *Int. J. Pept. Prot. Res.* **27,** 18-27 (1986).
18. Herzberg, O. & James, M. N. G. *Biochemistry* **24,** 5298-5302 (1985).
19. Rees, D. C., Lewis, M. & Lipscomb, W. N. *J. molec. Biol.* **168,** 367-387 (1983).
20. Jones, T. A. & Thirup, S. *EMBO J.* **5,** 819-822 (1986).
21. Handoll, H. D. *thesis*, Univ. Oxford (1985).
22. Artymiuk, P. J. & Blake, C. C. F. *J. molec. Biol.* **152,** 737-762 (1981).
23. McKenzie, H. A. & Shaw, D. C. *Biochem. Int.* **10,** 23-31 (1985).
24. Schaer, J-J, Milos, M. & Cox, J. A. *FEBS Lett.* **190,** 77-80 (1985).
25. Koga, K. & Berliner, L. J. *Biochemistry* **24,** 7257-7262 (1985).
26. Chothia, C. & Lesk, A. M. *Trends biochem. Sci.* **111,** 116-118 (1985).
27. Weaver, L. H. *et al. J. molec. Evol.* **21,** 97-111 (1985).
28. Tufty, R. M. & Kretsinger, R. H. *Science* **187,** 167-169 (1975).
29. Bernstein, F. C. *et al. J. molec. Biol.* **112,** 535-542 (1977).
30. Stuart, D. I. thesis, Bristol Univ. (1979).
31. Rossmann, M. G. & Argos, P. *J. molec. Biol.* **109,** 99-129 (1977).
32. Blundell, T. L. & Johnson, L. N. in *Protein Crystallography*, 147 (Academic, New York, 1976).

Bent DNA at a yeast autonomously replicating sequence

Michael Snyder*†, Andrew R. Buchman‡ & Ronald W. Davis†

* Department of Biology, Yale University, New Haven, Connecticut 06511, USA
† Department of Biochemistry, Stanford University School of Medicine, Stanford, California 94305, USA
‡ Department of Cell Biology, Stanford University, Stanford, California 94305, USA

DNA fragments that show retarded electrophoretic mobility through polyacrylamide gels have been found in both prokaryotes and eukaryotes[1–7]. In the case of kinetoplast DNA, evidence has been presented that the DNA is curved or 'bent'[8,9]. Bent DNA has previously been found at the λ and simian virus 40 (SV40) DNA replication origins[6,7]. Here we show the existence of bent DNA at a yeast autonomously replicating sequence (ARS1), a putative replication origin. The bent DNA has been localized to a 40–55 base pair (bp) segment and contains six $(A)_{3-5}$ stretches (that is, six poly(A) stretches, three to five nucleotides in length) phased approximately every 10.5 bp. This region contains a DNA binding site for a yeast protein factor. This site lies at the 3′ end of the *TRP1* gene, in a region devoid of nucleosomes, and is positioned 80 bp away from the ARS consensus sequence; removal of this region impairs ARS function *in vivo*. The bent DNA may be involved in transcription termination or the prevention of nucleosome assembly in this region.

Yeast ARS sequences have properties expected of origins of replication: they allow the autonomous replication of sequences to which they are attached, occur at a frequency expected of chromosomal origins, replicate once per cell cycle using genes responsible for chromosomal replication and act as origins of replication *in vitro* (see ref. 10 for review). Electron microscopic data for ribosomal DNA and 2μ DNA has mapped nascent replication bubbles near ARS elements suggesting that ARSs act as replication origins *in vivo*[11,12].

The structure of ARS1, which is near the 3′ end of the yeast *TRP1* gene, is shown in Fig. 1. There are three elements that comprise ARS1, domains A, B and C. Domain A is an 11-bp consensus sequence found at most other ARS sequences.[13,14]. Mutations in this sequence at ARS1 or the ARS at the HO locus obliterate ARS function *in vivo*[13,15,16]. Domains B and C flank ARS1 and are important for its function[16–18]; mutations in either

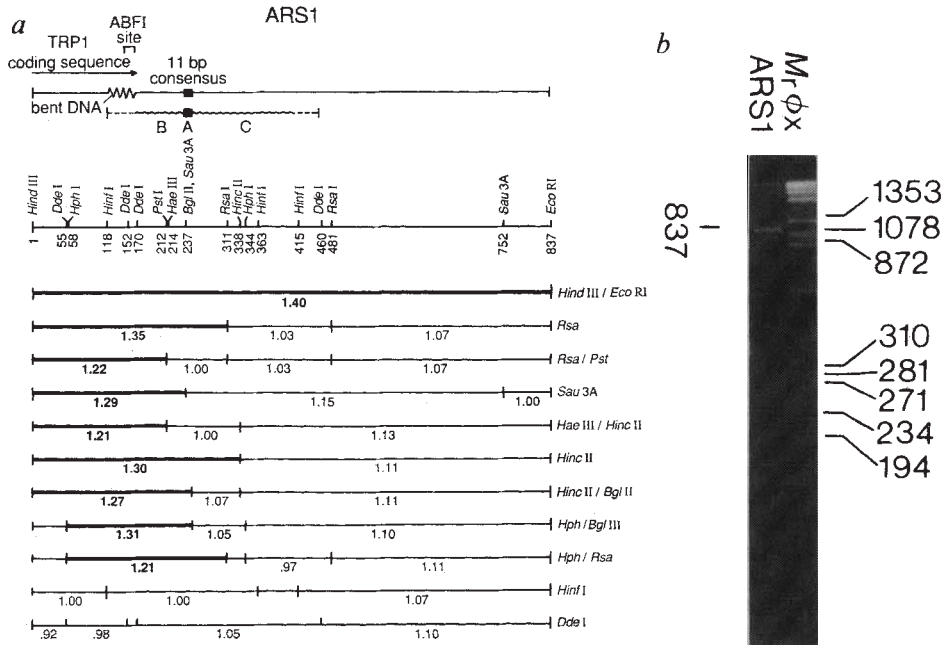

Fig. 1 *a*, Features of yeast ARS1; *b*, mobility of ARS1 restriction fragments on a 10% polyacrylamide gel. A (box), B and C (~ lines) are indicated. The broken lines indicate the degree of uncertainty of the extent of domains B and C. The *Hin*dIII/*Eco*RI ARS1 fragment was digested with the various restriction enzymes indicated, ethanol-precipitated and the electrophoretic mobilities of the fragments measured in a 10% polyacrylamide gel using 50 mM Tris base, 50 mM boric acid, 1 mM EDTA buffer[20]. Electrophoresis was at 2.9 V cm^{-1} for 16 h at 24.5 °C. Correct assignment of the DNA bands was determined by digestion with a second enzyme. Beneath the fragments is indicated the k value = expected distance migrated/observed migration distance. Fragments where $k > 1.20$ are indicated in bold type (not determined for fragments less than 60 bp). *b*, Migration of the 837 bp (determined from DNA sequence[23]) *Hin*dIII/*Eco*RI ARS1 restriction fragment in a 10% polyacrylamide gel. 837 indicates the position of the 837-bp ARS1 fragment. Sizes of ΦX174 *Hae*III marker fragments are given. The 281-bp fragment of ΦX174 migrates more slowly than expected. Other standards used are λ *Bst*EII markers and pBR322 *Hin*fI markers (not shown).

region reduce, but do not abolish, ARS function *in vivo*. Domain B begins very close to the 3′ end of the *TRP1* gene and may overlap it. A protein factor called ABFI (ARS binding factor I) that has been partially purified from yeast extracts,[20] binds in this region. The roles of the different ARS domains A, B and C, and of ABFI are unknown.

Because bent DNA has been found at bacteriophage and viral DNA replication origins[6,7], we tested for unusual DNA conformational features in the ARS1 region by subjecting restriction fragments of ARS1 DNA to electrophoresis in 10% polyacrylamide gels and identifying DNA fragments that exhibited unusually retarded mobility[19,21,22]. The mobilities were compared to 'normal' DNA fragments derived from pBR322, ΦX174 and λ DNAs (see Fig. 1 legend). Several DNA fragments were found that migrate 20–40% more slowly than predicted from the known DNA sequence[23] at 24.5 °C (Fig. 1). For example, an 837-bp fragment migrates to a position expected of a 1,170-bp fragment, 40% more slowly than anticipated. Using the method of Diekmann and Wang[19], several restriction fragments spanning the ARS1 region were tested for abnormal electrophoretic mobility and their k values determined (k = expected distance migrated/observed distance migrated for the DNA fragments). All DNA fragments that contain the 150-bp region between the *Hph*I and *Pst*I sites are found to have k values > 1.20. DNA fragments generated by cleaving the DNA with *Hin*fI and *Dde*I that cut within this 150-bp segment do not show abnormal mobility. Data presented below demonstrate that the *Hin*fI (118) and *Dde*I (152, 170) sites bracket a region which has properties characteristic of bent DNA. It is clear that there is a strong correlation between the position of the 118–170 *Hin*fI/*Dde*I region within a fragment and the observed k value. It has been noted that when the bent sequence is near the end of a fragment, the effect on retarded mobility is diminished[21]. Similar results were obtained after the digested ARS1 DNA was extracted with phenol or treated with proteinase K (data not shown). Thus reduced mobility does not appear to be due to association with proteins.

The abnormal DNA region was further localized by deletion analysis. A series of five deletions beginning at the *Hin*dIII site and extending increasing distances towards ARS1 were constructed (Fig. 2). These are called Δ108–Δ197; the numbers refer to the number of base pairs deleted. Restriction fragments of plasmids containing these deletions were generated by cleaving at sites located 150–200 bp outside the deletion regions in such a way that the putative bend is approximately in the centre of the fragment. The mobility and k values of these DNA fragments are shown in Fig. 2. Both the wild-type (wt) and Δ108 DNA fragments migrate more slowly than expected (31 and 32% more slowly, respectively). Δ149 (which differs from Δ108 by 41 bp) migrates normally. Thus at least some of the DNA that confers abnormal mobility lies between 108–149 bp on the map in Fig. 1.

Bent DNA tends to migrate more 'normally' at higher temperatures[19,22]. ARS1 DNA fragments from the deletion plasmids were analysed as above except that the electrophoresis was performed at 38 °C instead of 25 °C. As shown in Fig. 2, the wild-type and Δ108 fragments migrated faster at the higher temperature (lower k values). Relative mobilities of fragments lacking the 108–149 region were not affected by temperature. Thus, the effect of temperature on abnormal migration is consistent with the presence of bent DNA. Abnormal migration cannot be due to simple denaturation of DNA sequences, which would have the opposite effect with temperature. Inspection of the DNA sequence in the 108–149 region reveals the presence of six A_{3-5} stretches phased approximately every 10.5 bp (Fig. 3). This feature is characteristic of the bent DNA observed in kinetoplast DNA and λ replication origin[6,8,21]. It has been shown[22] that bending occurs at the 3′ junctions of A tracts, and is maximized when the 3′ A junctions are phased approximately every 10 bp.

Interestingly, DNA footprint analysis with partially purified yeast extracts demonstrates that a protein factor (ABFI) specifically binds to this region at residues 139–155 (Fig. 3). So far, five different binding sites for this factor have been identified in yeast DNA near ARS elements[24]. The binding site at ARS1 contains an $(A)_5$ stretch directly in its centre and other binding sites contain $(A)_{3-4}$ stretches. The bent conformation of the DNA at ARS1 may affect the affinity with which this factor binds DNA, similarly to SV40 T antigen where the nucleotides contacted by T antigen flank a DNA bend that is critical for binding[7].

To determine whether the bent DNA region is important for ARS1 function *in vivo*, the effect of the five deletions shown in Fig. 2 on ARS activity was assayed. Each ARS fragment was

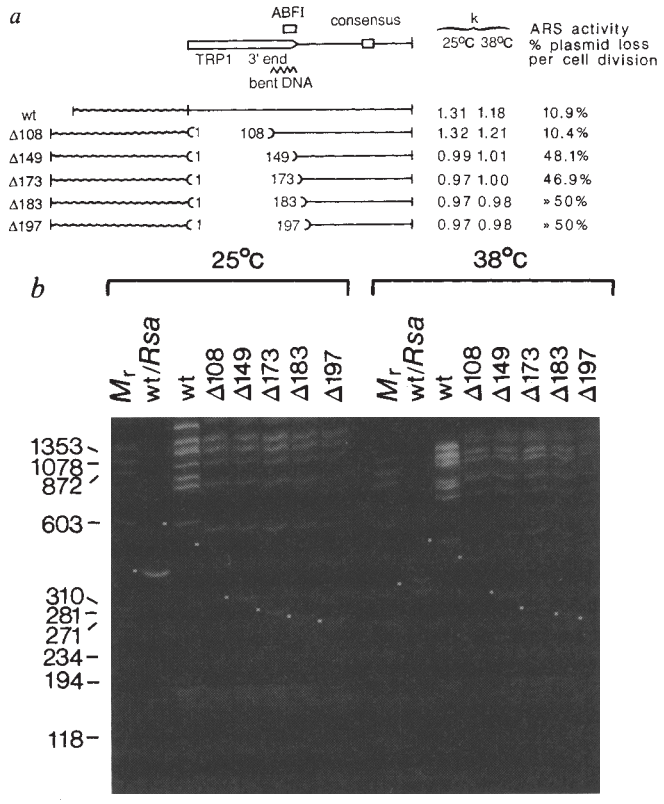

Fig. 2 Deletion analysis through the bent DNA region. *a*, The features of ARS1 and deletion fragments; *b*, mobility of digested deletion plasmids. *a*, All fragments were cloned into YRP14/CEN4[25]. The wild-type (wt) fragment was inserted between the *Hin*dIII and *Eco*RI sites. Deletions begin at the *Hin*dIII site and extend to the nucleotides indicated. *Bam*HI octanucleotide linkers were added and the *Bam*HI/*Eco*RI fragments cloned into YRP14/CEN4. For bent DNA analysis wt DNA was cleaved with *Rsa*I and *Eco*RV which leaves 156 bp of pBR322 sequences (⁓) adjacent to the *Hin*dIII/*Rsa*I ARS1 fragment. For the deletion plasmids, DNA was digested with *Rsa*I and *Sph*I which leave 187 bp of pBR322 sequences adjacent to ARS1 fragments. The lengths of the resulting fragments of interest are: wt 467 bp; Δ108, 396 bp; Δ149, 355 bp; Δ173, 331 bp; Δ183, 321 bp; Δ197, 307 bp. Electrophoresis was as in Fig. 1, at 25 °C. DNAs were also digested with enzymes that remove the pBR322 sequences and analysed on the same gel. The wt fragment still migrates abnormally, but all deletion fragments Δ108–Δ197 now migrate with *k* values <1.05 (not shown). ARS activity for each plasmid was assayed by transforming the plasmids into yeast strain YNN318 (a/α ura3-52/ura3-52 lys2-801/lys2-801 ade2-101/ade2-101 trp1-Δ901/trp1-Δ901 his3-Δ200/his3-Δ200). Cells harbouring plasmids were transferred from selective yeast minimal plates containing glucose to yeast minimal media supplemented with galactose (Calbiochem) and grown for 16 h at 30 °C (selecting for the plasmid) and an additional 10 h without selection, then plated on nonselective galactose plates containing limiting adenine; cells containing the plasmids appear pink while those lacking the plasmid are red[25]. The frequency of $\frac{1}{2}$ sectored colonies among colonies containing a single plasmid ($n = 296$ to 2,000) was measured to determine the rate of plasmid loss per cell division. Loss for wt and Δ108 plasmids was 4/1 1:0/2:0 segregation events while for Δ149 and Δ173 loss was >45/1 1:0/2:0 segregation events. Loss rates are very high for Δ183 and Δ197 constructs; all colonies appear homogeneously red. For Δ183 and Δ197 the number of plasmid-containing colonies plated was determined using selective glucose-containing plates. When similar ARS activity experiments were performed using glucose media, the relative stabilities of the wt and Δ108–Δ197 plamids are unchanged; however, the overall plasmid stability is increased. This medium effect is consistent with our observations that ARS plasmids are more stable on richer carbon sources[25,32]. *b*, Polyacrylamide gel showing mobility of the digested deletion plasmids. Left, DNAs analysed at 25 °C for 16 h; right, identical gel electrophoresis at 38 °C for 13 h. Lanes are as indicated and described above. Additional lanes are: M_r, ΦX174 *Hae*III digested DNA (sizes indicated in bp); wt/*Rsa*, *Rsa*I the *Hin*dIII/*Eco*RI ARS1 DNA digested with *Rsa*I. The 311-bp DNA fragment migrates coincident with a 356-bp DNA fragment at 25 °C (the 356-bp DNA fragment migrated 11% more slowly than expected on this particular gel).

cloned into YRP14/CEN4, a centromere-containing plasmid lacking an efficient replicator, and the rate of plasmid loss was determined by a colony colour assay (Fig. 2)[25]. Plasmids containing bent DNA (wild type and Δ108) are stable and lost at the

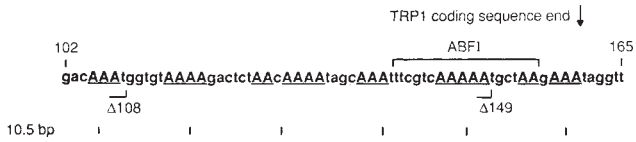

Fig. 3 DNA sequence in the bent DNA region of ARS1. All A residues that have neighbouring A residues are shown in bold type. 10.5-nucleotide intervals and the ABFI binding site are also indicated. A 25-bp region centred around this site is protected in DNase I protection experiments[20]. The endpoints of deletions Δ108 and Δ149 are shown. The ARS1 consensus sequence is located at position 241–251.

rate of 10% per cell division. The Δ149 plasmid, which lacks most of the bent DNA and half of the ABFI binding site, is fivefold less stable (Fig. 2). Thus, the bent DNA region is important for ARS function *in vivo*. It remains to be determined whether this effect is due to the bent DNA sequence, to the binding site for ABFI, or to both.

Bent DNA has been previously discovered at the bacteriophage λ and SV40 replication origins[6,7,26]. The discovery of such DNA at yeast ARS1 suggests that it may be important at chromosomal origins as well. Although the precise role of the bent DNA at ARS1 is presently unknown, the position of this region suggests several interesting possibilities. Since it is located at the 3′ end of the *TRP1* gene, it may be important for transcription termination. Transcription through ARS1 has been found to reduce its activity (M.S. and R.W.D., unpublished observations). Thus one role of domains B and C might be to act as transcription termination segments to protect domain A. The bent DNA with its associated protein could serve an important role in this process. In addition, the region immediately adjacent to the ARS1 consensus element on a *TRP1* ARS1 plasmid (nucleotides 60–242) has been shown to be free of nucleosomes[27,28]. Another function of the DNA bend might be to keep this region free of nucleosomes and readily accessible to the DNA replication apparatus and/or other proteins. The origin region of SV40 DNA has also been found to be free of nucleosomes[29–31].

We thank Ron Sapolsky for his help with sequence analysis and Roger Kornberg in whose laboratory some of these experiments were performed. This research was supported by NIH grant GM 21891.

Received 6 May; accepted 2 September 1986.

1. Bossi, L. & Smith, D. M. *Cell* **39,** 643–652 (1984).
2. Simpson, L. *Proc. natn. Acad. Sci. U.S.A.* **76,** 1585–1588 (1979).
3. Ross, W., Schulman, M. & Landy, A. *J. molec. Biol.* **156,** 505–529 (1982).
4. Stellwagen, N. C. *Biochemistry* **22,** 6186–6193 (1983).
5. Chalberg, S. S. & Englund, P. T. *J. molec. Biol.* **138,** 447–472 (1980).
6. Zahn, K. & Blattner, F. R. *Nature* **317,** 451–453 (1985).
7. Ryder, K., Silver, S., DeLucia, A. L., Fanning, E. & Tegtmeyer, P. *Cell* **44,** 719–725 (1986).
8. Marini, J. C., Levene, S. D., Crothers, D. M. & Englund, P. T. *Proc. natn. Acad. Sci. U.S.A.* **79,** 7664–7668 (1982).
9. Hagerman, P. J. *Proc. natn. Acad. Sci. U.S.A.* **81,** 4632–4636 (1984).
10. Campbell, J. L. *Genet. Engng* **5,** 109–155 (1983).
11. Newlon, C. S., Devenish, R. J., Suci, P. A. & Roffis, C. J. *ICN-UCLA Symp. molec. cell. Biol.* **22,** 501–516 (1981).
12. Saffer, L. D. & Miller, O. L. *Molec. cell. Biol.* **6,** 1148–1157 (1986).
13. Stinchcomb, D. C., Mann, C., Selker, E. & Davis, R. W. in *The Initiation of DNA Replication and Segregation of Yeast Chromosomes.* Vol. 3 (ed. Ray, D. S.) 473–488 (Academic, New York, 1981).
14. Broach, J. R. *et al. Cold Spring Harb. Symp. Quant. Biol.* **47,** 1165–1173 (1982).
15. Celniker, S. E., Sweder, K., Srienc, F., Bailey, J. E. & Campbell, J. L. *Molec. cell. Biol.* **4,** 2455–2466 (1984).
16. Kearsey, S. *Cell* **37,** 299–307 (1984).
17. Koshland, D., Kent, J. C. & Hartwell, L. H. *Cell* **40,** 393–403 (1985).
18. Srienc, F., Bailey, J. E. & Campbell, J. L. *Molec. Cell. Biol.* **5,** 1676–1684 (1985).
19. Diekmann, S. & Wang, J. C. *J. Molec. Biol.* **186,** 1–11 (1985).
20. Buchman, A., Kimmerly, W., Rine, J. & Kornberg, R., (manuscript in preparation).
21. Wu, H. & Crothers, D. M. *Nature* **308,** 509–513 (1984).
22. Koo, H. S., Wu, H.-M. & Crothers, D. M. *Nature* **320,** 501–506 (1986).
23. Tschumper, G. & Carbon, J. *Gene* **10,** 157–166 (1980).
24. Buchman, A. R. & Kornberg, R., (manuscript in preparation).
25. Hieter, P., Mann, C., Snyder, M. & Davis, R. W. *Cell* **40,** 381–392 (1985).
26. Zahn, K. & Blattner, F. R. *EMBO J.* **4,** 3605–3616 (1985).
27. Thoma, F., Bergman, L. W. & Simpson, R. T. *J. molec. Biol.* **177,** 715–733 (1984).
28. Thoma, F. & Simpson, R. T. *Nature* **315,** 250–252 (1985).
29. Varshavsky, A. J., Sundon, O. & Bohm, M. *Cell* **16,** 453–466 (1979).
30. Jakobovits, E. B., Brutosom, S. & Aloni, Y. *Nature* **285,** 263–265 (1980).
31. Saragosti, S., Moyne, G. & Yaniv, M. *Cell* **20,** 65–73 (1980).
32. Snyder, M. & Davis, R. W., (manuscript in preparation).