The pyrimidine/purine-biased region of the epidermal growth factor receptor gene is sensitive to S1 nuclease and may form an intramolecular triplex

Mikio KATO,*† Jun KUDOH* and Nobuyoshi SHIMIZU*‡

*Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160, and

†Laboratory of Molecular Biology, College of Integrated Arts and Sciences, University of Osaka Prefecture, 4-804 Mozuume-cho, Sakai 591, Japan

The pyrimidine/purine-biased region located upstream of the EGF (epidermal growth factor) receptor gene transcription initiation sites was sensitive to S1 nuclease when under superhelical tension. The structural basis of this specific reactivity to S1 nuclease was probed by the use of diethyl pyrocarbonate. The patterns of modification suggested that the H-form proposed by Mirkin, Lyamichev, Drushlyak, Dobrynin, Filippov & Frank-Kamenetskii [Nature (London) (1987) 330, 495–497], which includes an intramolecular triplex and a single-stranded region, was the most plausible model for the sequence tested. The results of dimethyl sulphate modification also supported this model.

INTRODUCTION

Many genes have S1 nuclease-hypersensitive sites in their regulatory regions. These S1 nuclease-hypersensitive segments of DNA often contain a purine/pyrimidine strand bias. Unusual DNA structures characterized by nuclease hypersensitivity or specific chemical reactivity occur around or in homopurine/homopyrimidine sequences and are stabilized when the segment is negatively supercoiled or in acidic pH (for review see [1]).

At least three models for the structural basis of this hypersensitivity have been proposed from the results obtained by use of chemical and enzymic probes. Evans & Efstratiadis [2] proposed that each strand of the purine/pyrimidine region has a different backbone conformation in the case of $(dG-dA)_n \cdot (dT-dC)_n$. This can be classified as heteronomous DNA [3] by analogy to the proposed structure for $(dA)_n \cdot (dT)_n$. Mirkin *et al.* [4,5] demonstrated the possibility of forming an intramolecular triplex structure termed the H-form when the nucleotide sequence has mirror symmetry. Pulleyblank *et al.* [6] proposed a structure with alternating Hoogsteen (G·C⁺) and Watson–Crick (A·T) base pairs in $(dG-dA)_n \cdot (dT-dC)_n$ based on the result that the methylation efficiency at N-7 of guanine by DMS (dimethyl sulphate) was reduced under acidic pH. The promoter region of the epidermal growth factor (EGF) receptor gene contains nearly identical pyrimidine/purine sequences [7] that can be regarded as possessing either direct repeats or mirror symmetry. The direct repeat sequences could conform to a DNA slippage model which would explain its S1 nuclease sensitivity [8]. The sequence with mirror symmetry is thought to be able to form an intramolecular triplex under mildly acidic or neutral conditions because the pK_a value of protonation of a Hoogsteen G·C⁺ base pair is about 8 [5]. Recently, Johnson *et al.* [9] have reported that the pyrimidine/purine region of the EGF receptor gene was sensitive to S1 nuclease under superhelical tension, and proposed that the structural basis for this sensitivity was DNA slippage. However, there seems to be no direct evidence to prove the slippage model. In order to determine the conformation of the promoter region of the EGF receptor gene in the superhelix, we have carried out chemical probing at various pH values using diethyl pyro-

carbonate (DEPC) which mainly carbethoxylates N-7 of unstacked adenine [10,11]. The positions attacked by DEPC were determined by the pausing of primer extension with the Klenow fragment of *Escherichia coli* DNA polymerase I. The primer extension method was initially developed by HuQu *et al.* [12] for probing the secondary structure of large RNA molecules with retroviral reverse transcriptase. The Klenow fragment was also expected to pause at a nucleotide whose base was modified so as to be unable to form Watson–Crick-type base pairing. We have also used DMS, which can methylate N-7 of guanine [13], as a chemical probe to confirm the conformation of the EGF receptor gene promoter. We hypothesized that an unusual DNA structure in the EGF receptor gene promoter region was generated with Hoogsteen base pairing, because DMS treatment made the region resistant to S1 nuclease by preventing the methylated guanine from participating in Hoogsteen base pairing.

MATERIALS AND METHODS

DNA, chemicals and enzymes

The plasmid pERCAT-1, which contains a 1.1 kb fragment including the promoter region of the EGF receptor gene, was kindly provided by Dr. Ira Pastan [14]. Three *Sac*II cleavage sites were mapped within the 1.1 kb fragment [15]. A 0.25 kb *Sac*II fragment containing the S1 nuclease-sensitive region was cut out from pERCAT-1. T4 DNA polymerase (3'→5' exonuclease activity) was used to generate blunt ends. The fragment was inserted into the *Sma*I site of pUC19 and the resulting plasmid, p7ERS, was cloned into *E. coli* strain JM105 [16]. The supercoiled plasmid DNA was prepared by ethidium bromide/caesium chloride equilibrium density gradient centrifugation [17]. The M13/pUC reverse and universal sequencing primers were purchased from Toyobo. [α -³²P]dCTP (3000 Ci/mmol) and [γ -³²P]ATP (3000 Ci/mmol) were obtained from Amersham International. DEPC and DMS were from Aldrich and Wako Pure Chemicals, respectively. The restriction endonucleases, Klenow fragment of *E. coli* DNA polymerase I and T4 DNA polymerase were from Toyobo. S1 nuclease was from Pharmacia.

Abbreviations used: EGF, epidermal growth factor; DEPC, diethyl pyrocarbonate; DMS, dimethyl sulphate.

‡ To whom correspondence and reprint requests should be addressed.

Fig. 1. Scheme of primer extension assay

Supercoiled p7ERS was treated with S1 nuclease or DEPC. After denaturation, the DNA was hybridized to 5'-end-labelled primer, and primer extension was performed with the Klenow fragment of DNA polymerase I. The positions of S1-cleavage and carbethoxylated nucleotides were detected by the length of the run-off DNA chains and the prematurely terminated DNA chains, respectively. The purine-rich strand was used as a template for DEPC-probing and both strands were used for S1-nicking. The solid line represents the insert in question.

S1 nuclease digestion

S1 nuclease treatment was performed in 320 μ l containing 16 μ g of plasmid DNA, 100 units of S1 nuclease, 30 mM-sodium acetate (pH 4.6), 220 mM-NaCl and 1 mM-ZnSO₄. The reaction was incubated for 5 min at 37 °C and stopped by addition of 2 μ l of 0.5 M-EDTA followed by phenol extraction. DNA was recovered by ethanol precipitation and digested using restriction endonucleases. The S1-cutting reaction was performed with high salt to inactivate any double-strand nicking activity which might be present in the enzyme preparation [18].

Fine mapping of S1 nuclease and DEPC target sites

In order to map the sites of S1 nuclease digestion and DEPC modification at nucleotide resolution, the primer extension assay was carried out using supercoiled p7ERS DNA. The procedures used here were essentially the same as described by Gralla [19], and represented schematically in Fig. 1. Carbethoxylation of p7ERS was performed in acidic and neutral conditions. Buffer 4E contained 60 mM-sodium acetate, 50 mM-NaCl and 1 mM-EDTA (pH 4.5). Buffer 7E contained 50 mM-sodium cacodylate, 50 mM-NaCl and 1 mM-EDTA (pH 7.0). The reactions containing 3 μ g of p7ERS in 200 μ l of each buffer were initiated by adding

3 μ l of DEPC and incubated for 10 min at 23 °C. The reactions were stopped and DNA was precipitated by addition of ethanol. DEPC-treated DNA was resuspended in 0.3 M-sodium acetate, precipitated with ethanol, rinsed with 80 % ethanol, and dried. Denaturation of the DNA was performed with 0.2 M-NaOH according to the standard pUC sequencing protocol [20]. S1 nuclease digestion of p7ERS was performed in a similar way to that described above but at 20 °C, and 3 μ g of digested p7ERS was subjected to mapping. This condition caused partial digestion, and the S1-nicking pattern was analysed on each strand. The M13/pUC reverse and universal sequencing primers (5'-CAGGAAACAGCTATGAC-3' and 5'-GTTTCCAGTCACGAC-3'; Toyobo) were labelled with ³²P at their 5'-end using T4 polynucleotide kinase and [γ -³²P]ATP. The kinase reactions were carried out in 20 μ l containing 8 pmol of primer, 50 mM-Tris/HCl (pH 7.6), 10 mM-MgCl₂, 10 mM-MgCl₂, 10 mM-2-mercaptoethanol and 4 units of kinase. The reactions were incubated for 60 min at 37 °C, and were stopped by heating at 90 °C for 3 min. The reaction mixtures were cooled and 5 μ l aliquots were added to each of the denatured DNA templates. The annealing mixtures were incubated at 90 °C for 3 min, and slowly cooled to allow primer hybridization. After hybridization, the extension mixture (5 μ l), containing 1 mM of each dNTP, and 2 units of Klenow fragment in 50 mM-Tris/HCl (pH 7.6), 10 mM-MgCl₂ and 10 mM-2-mercaptoethanol was added to each annealing mixture. Primer extension was performed in 10 μ l at 25 °C for 15 min, and the reaction was stopped by the addition of 5 μ l of loading solution containing 95 % formamide, 20 mM-EDTA and 0.25 % Bromophenol Blue. Aliquots were applied to denaturing 8 % polyacrylamide-gel electrophoresis. Dideoxy sequencing was also carried out as reference. The polymerase pauses or stops at modified or nicked positions, giving rise to prematurely terminated DNA chains whose ends can be located precisely by gel electrophoresis in parallel with the dideoxy sequencing lanes generated from the unmodified DNA templates.

Methylation of DNA by DMS

Methylation of p7ERS was performed in 200 μ l of reaction mixture containing 12 μ g of DNA, 100 mM-sodium cacodylate, 100 mM-sodium acetate and 5 mM-EDTA (pH 7.3). The reaction was started by adding 5 μ l of a 20 % (v/v) solution of DMS in ethanol. After incubation at 25 °C for 60 min, the reaction was terminated by adding 6 μ l of 2-mercaptoethanol, 50 μ l of 3 M-sodium acetate and 750 μ l of cold ethanol. DNA was recovered by precipitation, resuspended in 0.3 M-sodium acetate, precipitated with ethanol, rinsed with 80 % ethanol and dried. S1 nuclease treatment was essentially as described above and was followed by digestion with the restriction endonuclease *Dra*I whose recognition sequence is TTATAA. The digests were fractionated by 1 % agarose gel electrophoresis, and visualized by ethidium bromide staining.

RESULTS

S1 nuclease-sensitive sites in the EGF receptor gene promoter region

A characteristic pyrimidine/purine-biased sequence appears repeatedly in the EGF receptor gene upstream of the transcription initiation sites (Fig. 2a) [7]. Similar sequences have been found which are sensitive to S1 nuclease when under superhelical tension in the upstream promoter region of both the chicken and mouse α 2(I) collagen genes [8]. In order to examine whether such sequences in the EGF receptor gene can form a S1-sensitive structure, the plasmid pERCAT-1 with native bacterial super-

Fig. 2. Nucleotide sequence of the pyrimidine/purine-biased region of the EGF receptor gene promoter

(a) Sequence pattern and positions attacked by probes. The brackets with arrowheads indicate sequences with mirror symmetry. The horizontal lines with arrowheads indicate the direct repeat sequences. The asterisks indicate DEPC-modified purines. The bases specifically modified under acidic conditions are underlined. S1-nicking regions are indicated by open squares (□) for the pyrimidine strand and closed squares (■) for the purine strand. The nucleotide numbering is according to Johnson *et al.* [9]. The major *in vivo* transcription start site is -258 . (b) Possible models of secondary structure for the mirror symmetry and the direct repeat pyrimidine/purine sequences.

helicity was treated with S1 nuclease. S1 nuclease cleaved the superhelical circular DNA and the cleavage sites were mapped approximately to the pyrimidine/purine-biased region by restriction enzyme digestion and fractionation on a polyacrylamide gel (Fig. 3). S1 nuclease did not cleave linearized DNA (results not shown). The structural basis for the S1 sensitivity could be due to either H-form DNA or DNA slippage (Fig. 2b). Johnson *et al.* [9] have also shown that there is a region sensitive to S1 nuclease at the pyrimidine/purine stretch of the EGF receptor gene promoter. In order to map the sites that were sensitive to S1 nuclease at the nucleotide level, we subcloned the 0.25 kb *Sac*II fragment containing the pyrimidine/purine-biased region. This plasmid, termed p7ERS, maintained S1-sensitivity (see Fig. 5b, lane 4). The plasmid p7ERS was treated with S1 nuclease or DEPC and the sites of cleavage were mapped by primer extension.

Determination of S1- and DEPC-target sites

The primer extension assay showed that S1 nuclease cleaved at three sites in the pyrimidine/purine-biased region on the purine strand (Fig. 4a, lane 3), and at two sites on the pyrimidine strand (Fig. 4b, lane 1). S1 nuclease attacked at the joint of each pyrimidine/purine cluster (Fig. 2a). Three cutting sites were observed by the S1 nuclease treatment of pERCAT-1 (Fig. 3), but S1-nicking occurred with different efficiencies at each site in p7ERS. We cannot yet explain what caused this difference between pERCAT-1 and p7ERS. DEPC also attacked at the joint of each cluster. The purines of the second cluster (-352 to -349) were specifically modified by DEPC at an acidic pH but not under neutral conditions (Fig. 4a, lanes 1 and 2). This observation is consistent with the H-form model [4,5]. The pyrimidines of the second cluster could form Hoogsteen base pairing with purines of the preceding cluster by a sharp bending of the DNA at the joint between the first and second clusters. Purines of the second cluster would be left single-stranded (forming an intramolecular triplex and a single-stranded region) because the first and second clusters have mirror symmetry (Figs.

Fig. 3. S1 nuclease-sensitive sites in supercoiled pERCAT-1

Ethidium bromide-stained polyacrylamide gel and the restriction map of the *Hind*III fragment of pERCAT-1. H, S and B indicate the restriction sites of *Hind*III, *Sac*II and *Bgl*II, respectively. The solid line corresponds to the pyrimidine/purine-biased region. S1-sensitive sites were approximately mapped to this region by gel mobility assignment. Lanes 1 and 2, pERCAT-1 was digested with *Hind*III and *Bgl*II; lanes 3 and 4, pERCAT-1 digested with *Hind*III; lanes 5 and 6, pERCAT-1 digested with *Bgl*II; lanes 1, 3 and 5, DNAs were treated with S1 nuclease prior to the restriction endonuclease digestion. Positions of two marker fragments (*Sau*3AI-digested pUC 19) are indicated.

Fig. 4. DEPC-modification and S1-nicking patterns of supercoiled p7ERS

(a) Probing the purine strand with DEPC (lanes 1 and 2) and S1 nuclease (lane 3). DEPC treatment was performed in buffer 7E (lane 1) or buffer 4E (lane 2). (b) Probing the pyrimidine strand with S1 nuclease (lane 1). Lanes T, C, G, A refer to the dideoxy sequencing reactions. Note that the 5'-phosphorylated DNA (lanes 1–3 of *a* and lane 1 of *b*) migrated slightly faster (less than one nucleotide) than the control DNA (lanes T, C, G, A) having a 5'-hydroxy group.

2a and 2b). DEPC would attack these purines. Efficient forming of Hoogsteen C⁺·G pairing requires acidic conditions, since the pK_a value of the protonation is about 8 [5]. This explains the pH-dependency of DEPC modification. The different specificity of S1-nicking on each strand also supported the H-form rather than

Fig. 5. Effect of DMS treatment on S1-specific cutting of p7ERS

DNA fragments generated by enzyme digestion were visualized by ethidium bromide staining. (a) Lane M, *Hind*III-digested λ DNA marker; lane 1, DMS-treated p7ERS; lane 5, intact p7ERS; lanes 2 and 6, p7ERS heated at 75 °C for 30 min; lanes 3 and 7, p7ERS digested with S1 nuclease; lanes 4 and 8, p7ERS digested with *Dra*I; DNAs in lanes 2–4 were methylated by DMS prior to heating or enzyme digestion. T, trimer or open circular dimer; LD, linear dimer; SCD, supercoiled dimer; OC, open circular p7ERS; L, linearized p7ERS; SC, supercoiled p7ERS. (b) Lane 1, DMS-treated/*Dra*I-digested p7ERS; lane 2, DMS-treated/S1 nuclease-treated/*Dra*I-digested p7ERS; lane 3, *Dra*I-digested p7ERS; lane 4, S1 nuclease-treated/*Dra*I-digested p7ERS; lane M, *Hind*III-digested λ DNA marker. The bands marked with arrows were generated by the S1 nuclease treatment prior to *Dra*I digestion.

the slippage model. On the pyrimidine strand, the hypothetical sharp bending regions were nicked by S1 nuclease (Fig. 4b, lane 1; Fig. 2a, □). In the slippage model, S1-nicking should have equally occurred at both edges of the repeat sequences.

Effect of DMS treatment on S1-sensitivity

DMS treatment did not cause breakage of supercoiled p7ERS (Fig. 5a, compare lane 1 with lane 5). However, S1-specific

bands were markedly decreased by DMS treatment (Fig. 5b, compare lane 2 with lane 4). Since DMS treatment was performed at neutral pH, the supercoiled p7ERS was not in an unusual structure. The methylation by DMS which mainly occurs at N-7 of guanine prevented the structural transition, namely generation of S1-hypersensitive sites of supercoiled p7ERS. Quantitative analysis of the methylated bases was not done, but the presence of modified bases was evaluated. N-7-methylated guanine is fragile and breakage of the glycosidic bond is easily induced by heating at neutral pH [21]. Heating at 75 °C for 30 min fragmented DMS-treated p7ERS but had no effect on untreated p7ERS (Fig. 5a, lanes 2 and 6). The restriction endonuclease *DraI* recognizes the sequence TTAAA [22]. *DraI* digestion of p7ERS generated two visible bands (Fig. 5a, lane 8); *DraI* digestion of DMS-treated p7ERS resulted in an additional band (linear) caused by partial digestion, even when an excess amount of enzyme was used (Fig. 5a, lane 4). This was thought to be due to the presence of methylated adenines which cannot be recognized by *DraI*. DMS methylates N-3 of adenine less efficiently than it does N-7 of guanine [13]; the methylated guanines are probably also present. *EcoRI* treatment of DMS-treated p7ERS showed that about 20% remained supercoiled and 50% was nicked (results not shown). *EcoRI* recognizes GAATTC [23] and DMS treatment makes DNA resistant to digestion by *EcoRI* [24]. Thus methylated bases are present in DMS-treated p7ERS and superhelicity was maintained during DMS treatment.

Assignment of a plausible model for the secondary structure of the EGF receptor gene promoter

Since the guanine methylated at N-7 can form Watson-Crick base pairs but cannot form Hoogsteen base pairs, the loss of S1-sensitivity by DMS treatment of the supercoiled p7ERS indicated that the structure of the S1-hypersensitive sites was based on Hoogsteen base pairing. Although there was no direct evidence that methylation of purine bases did not affect formation of the heteronomous DNA structure proposed by Evans & Efstratiadis [2], which is S1-sensitive, a pH dependency of the carbethoxylation patterns (Fig. 4a) strongly suggested the involvement of an H-form structure with Hoogsteen base pairing. Although Johnson *et al.* [9] proposed previously that the S1-sensitivity of this region was caused by DNA slippage, we believe that the H-form model is more plausible. Unequal sensitivity of the S1-nicking sites between two strands may rule out the slippage model. Htun & Dahlberg [25] showed that the H-form was not present in (T-C)₁₈·(G-A)₁₈ at neutral pH with native bacterial superhelicity, but was clearly formed with increased superhelicity, even at pH 8, when probed by DEPC [25]. So, this H-form may be significant under physiological conditions.

DISCUSSION

The physiological role of unusual DNA structures *in vivo* has not been completely defined. Regions required for EGF receptor gene expression have been identified by the use of a series of deletion mutants. The potential H-form region of the EGF receptor gene promoter was found to be dispensable for maximal transcription activity *in vivo* and *in vitro* [14,26]. No proteins have been identified which directly interact with this region, but some protection patterns were observed around this region when analysed by DNAase I footprinting and exonuclease III protection [14,26]. A recent report by Johnson *et al.* [9] indicated that a mutation introduced into this region decreased the activity *in vivo* of the EGF receptor gene promoter 3–5-fold in transient expression. They showed that the S1-sensitive region of the EGF receptor gene promoter contained the target sites of two specific proteins, one of which was Sp1 [9]. In the case of EGF receptor

gene expression, the role of the S1-sensitive region is differently interpreted by different assays. There are reports that this region is not essential [14,26] and that it affects expression [9].

The human U1 and U2 RNA genes also have S1-sensitive sequences which are comprised of (dG-dA)_n·(dT-dC)_n. The accumulation of human U1 RNA was observed after a fragment of the human U1 RNA gene with the S1-sensitive region deleted was injected into frog oocytes [27]. But the (dG-dA)_n·(dT-dC)_n stretches of the human U1 and U2 RNA gene in chromatin maintained their S1-sensitivity [28,29]. The promoter region of the chicken adult β^A -globin gene showed an unusual reactivity to a single-strand-specific chemical reagent in a supercoiled plasmid and in the chromatin of erythrocytes; however, this reactivity was not found in chromatin of embryonic erythrocytes which express embryonic β -globin in contrast to adult β^A -globin [30]. In such cases, actively transcribed genes may have region of altered secondary structure. These observations indicate that unusual DNA structures related to S1 nuclease-sensitivity or specific chemical reactivity have a role in gene regulation, even though these DNA conformations have not been shown to be the target of specific regulatory proteins. Weintraub [31] demonstrated that DNA sequences capable of forming S1-sensitive sites in supercoiled plasmids bind nucleosomes poorly *in vitro*. Dissociation of a nucleosome *in vivo* might cause localized superhelical tension which could allow formation of an altered local DNA secondary structure [32]. Decreased affinity for nucleosomes may be necessary for recognition of regions which are the target sites of specific binding proteins.

Recent works have proposed how such DNA structures occur [5,34] and in addition, evidence *in vivo* for triplex DNA has been presented [35]. How unusual DNA structures play a role in gene regulation is a critical question. Here, we propose how such structures act to regulate transcription. Weintraub [31] made the observations that DNA sequences capable of forming S1-sensitive sites show low affinities for nucleosomes and that S1 nuclease requires a fully supercoiled molecule for cutting. In addition, it is known that transcription drives the supercoiling of DNA [36,37]. Generation of the unusual DNA structures occurs to relax negative supercoils [6,34]. These facts taken together form the basis of our model in which as follows: when one nucleosome dissociates from the nucleosome bead, two helical turns are released and negative superhelical tension is introduced into the

Fig. 6. Model for the role of unusual DNA structure in regulating transcription

Structural transition cancels negative superhelical tension generated by nucleosomes release. The solid bent line represents the unusual DNA region and the closed square represents a *cis*-acting element.

DNA molecule. Subsequently, a supercoiling-induced structural transition occurs and the negative superhelical tension is absorbed. After the transition, the naked DNA region has a low affinity for the nucleosome, since the negative supercoiling has been already compensated for at this time and the reconstitution of nucleosome introduces positive supercoiling [38]. If the transition between the unusual DNA structure and usual B-DNA is much faster than that between nucleosome-release and nucleosome-reconstitution, the naked DNA regions remain open ready to be recognized by specific DNA binding proteins. These steps are schematically illustrated in Fig. 6. In the case of EGF receptor gene, the potential sharp-bending centre (position -353) of the triplex locates about 110 bp upstream of EGF receptor gene-specific transcription factor (ETF) binding site [33] and close to a number of Sp1 binding sites [9]. While the RNA polymerase advances on the DNA template, negative superhelicity increases behind the polymerase and positive superhelicity accumulates in front [36]. Topoisomerase may act to remove these tensions, both positive and negative, and the generation of unusual DNA structure may be necessary to absorb the negative superhelicity created behind the polymerase. The introduced positive superhelicity in front of the polymerase may also be removed by release of nucleosomes.

The expression of specific transcription factors is a step in gene regulation. However the regulation of chromatin structure is also important for gene regulation. We do not feel that the unusual DNA structures are positively recognized by specific DNA-binding proteins. They may be the result of modulation of the superhelicity of DNA necessary to allow for transcription.

We are grateful to Dr. Ira Pastan (National Cancer Inst.) for providing the plasmid pERCAT-1. We also would like to thank Professor Masahiro Ishigami (University of Osaka Prefecture) for encouragement. This work was supported in part by Grants-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

REFERENCES

- Wells, R. D. (1988) *J. Biol. Chem.* **263**, 1095–1098
- Evans, T. & Efstratiadis, A. (1986) *J. Biol. Chem.* **261**, 14771–14780
- Arnott, S., Chandrasekaran, R., Hall, I. H. & Puigjaner, L. C. (1983) *Nucleic Acids Res.* **11**, 4141–4155
- Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A. & Frank-Kamenetskii, M. D. (1987) *Nature (London)* **330**, 495–497
- Voloshin, O. N., Mirkin, S. M., Lyamichev, V. I., Belotserkovskii, B. P. & Frank-Kamenetskii, M. D. (1988) *Nature (London)* **333**, 475–476
- Pulleyblank, D. E., Haniford, D. B. & Morgan, R. (1985) *Cell* **42**, 271–280
- Ishii, S., Xu, Y.-Hh., Stratton, R. H., Roe, B. A., Merlino, G. T. & Pastan, I. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 4920–4924
- McKeon, C., Schmidt, A. & Crombrugge, B. (1984) *J. Biol. Chem.* **259**, 6636–6640
- Johnson, A. C., Jinno, Y. & Merlino, G. T. (1988) *Mol. Cell. Biol.* **8**, 4174–4184
- Ehresmann, C., Baudin, F., Mougél, M., Romby, P., Ebel, J.-P. & Ehresmann, B. (1987) *Nucleic Acids Res.* **15**, 9109–9128
- Herr, W. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8009–8013
- HuQu, L., Michot, B. & Bachelier, J. P. (1983) *Nucleic Acids Res.* **11**, 5903–5920
- Lawley, P. D. & Brookes, P. (1963) *Biochem. J.* **89**, 127–138
- Johnson, A. C., Ishii, S., Jinno, Y., Pastan, I. & Merlino, G. T. (1988) *J. Biol. Chem.* **263**, 5693–5699
- Kawasaki, K., Kudoh, J., Omoto, K. & Shimizu, N. (1988) *Jpn. J. Cancer Res. (Gann)* **79**, 1174–1183
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33**, 103–109
- Radloff, R., Bauer, W. & Vinograd, J. (1967) *Proc. Natl. Acad. Sci. U.S.A.* **57**, 1514–1521
- Schon, E., Evans, T., Welsh, J. & Efstratiadis, A. (1983) *Cell* **35**, 837–848
- Gralla, J. D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3078–3081
- Hattori, M., Hidaka, S. & Sakaki, Y. (1985) *Nucleic Acids Res.* **13**, 7813–7827
- Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560–564
- Purvis, I. J. & Moseley, B. E. B. (1983) *Nucleic Acids Res.* **11**, 5467–5474
- Hedgepeth, J., Goodman, H. M. & Boyer, H. W. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 3448–3452
- Cantor, C. & Efstratiadis, A. (1984) *Nucleic Acids Res.* **12**, 8059–8072
- Htun, H. & Dahlberg, J. E. (1988) *Science* **241**, 1791–1796
- Kageyama, R., Merlino, G. T. & Pastan, I. (1988) *J. Biol. Chem.* **263**, 6329–6336
- Skuzeski, J. M., Lund, E., Murphy, J. T., Steinberg, T. H., Burgess, R. R. & Dahlberg, J. E. (1984) *J. Biol. Chem.* **259**, 8345–8352
- Htun, H., Lund, E. & Dahlberg, J. E. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 7288–7292
- Htun, H., Lund, E., Westin, G., Pettersson, U. & Dahlberg, J. E. (1985) *EMBO J.* **4**, 1839–1845
- Kohwi-Shigematsu, T., Gelinas, R. & Weinraub, H. (1983) *Proc. Natl. Acad. Sci.* **80**, 4389–4393
- Weintraub, H. (1983) *Cell* **32**, 1191–1203
- Lasen, A. & Weintraub, H. (1982) *Cell* **29**, 609–622
- Kageyama, R., Merlino, G. T. & Pastan, I. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 5016–5020
- Htun, H. & Dahlberg, J. E. (1989) *Science* **243**, 1571–1576
- Burkholder, G. D., Latimer, L. J. P. & Lee, J. S. (1988) *Chromosoma* **97**, 185–192
- Futcher, B. (1988) *Trends Genet.* **4**, 271–272
- Tsao, Y.-P., Wu, H.-Y. & Liu, L. F. (1989) *Cell* **56**, 111–118
- Laskey, R. A., Honda, B. M., Mills, A. D. & Finch, J. T. (1978) *Nature (London)* **275**, 416–420

Received 19 July 1989/8 January 1990; accepted 15 January 1990