BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **234,** 73–78 (1997)
ARTICLE NO. RC976581

Stalling of Human Methyltransferase at Single-Strand Conformers from the Huntington's Locus

Steven S. Smith and David J. Baker
Department of Cell and Tumor Biology, City of Hope National Medical Center, Duarte, California 91010-3000

Received March 25, 1997

We describe evidence for a sequence of events in which the Human DNA(cytosine-5)methyl-transferase first methylates spontaneous single-stranded conformers (SSCs) and then stalls at the methylated site to produce a complex with the conformationally unusual DNA. This property of the enzyme is a result of its ability to respond to a general loss of symmetry at its CG recognition site. The data suggest that DNA methyltransferase, itself, may physically participate in biological processes that distinguish between DNA that is in the normal Watson–Crick paired conformation and DNA that is conformationally unusual (*e.g.* a hairpin loop or misassembled replication intermediate). The *in vitro* methylation of spontaneous SSCs from the Huntington's locus illustrates the phenomenon.

As a consequence of its mechanism of action, human DNA methyltransferase necessarily recognizes conformationally unusual DNA (1, 2). This specificity suggests that the evolution of DNA methyltransferase and DNA methylation have been driven by the occurrence of unusual structures in DNA (1-3). In addition to a possible role in the stabilization of the differentiated state, the enzyme may have roles in repair, recombination and general chromosome stability (3). This is further supported by the association of methylation with sites of DNA damage (4-7) and by the demonstration of the same capacity for unusual conformer recognition in bacterial systems (8-11).

The capacity of single-stranded DNA to self-associate is well known. This tendency makes it necessary to use highly chaotropic solutions to resolve single-stranded DNAs by molecular weight for sequencing purposes (12). Even under those conditions, genomic DNA with a high G+C content can retain secondary structure that impedes separation by molecular weight (12). More recently, the recognition that this secondary structure is very sensitive to base sequence has been used to detect single-base changes in DNA sequence as changes in the electrophoretic mobility of single-stranded conformers (SSCs) under non-denaturing conditions (13-14). While the tendency of these sequences to adopt unusual single-strand conformations provides a useful tool in detecting mutations, it may also provide a significant impediment to the normal replication and transcription processes and may promote clastic events like those associated with insertion, deletion, expansion, recombination and transposition (3, 15-19). DNA methyltransferase is currently the only enzyme that has been shown to interact with the SSCs (15-17). In those studies, the human enzyme was shown to interact with SSCs from the human *ras* gene and the CCG triplet repeats associated with the dynamic mutations observed in the human *FMR*-1 gene of fragile X.

The function of the interaction between DNA methyltransferase and SSCs from *FMR*-1 is currently unclear. One possible function is implicit in the twofold symmetry of the CG dinucleotide pair. Since this symmetry is present in a normally-paired duplex and not in a single-strand conformer of the same sequence, it can be used by the methyltransferase to distinguish between them. Here, we report evidence for a two-step process in which the human DNA methyltransferase first methylates SSCs and then stalls at the methylated site to produce a non-productive complex with the conformationally unusual DNA.

MATERIALS AND METHODS

Oligodeoxynucleotides. These molecules were synthesized by β-cyanoethyl phosphoramidite chemistry on a 1 μmol scale with a Cyclone Plus MilliGen/Biosearch instrument and were purified by OligoPak columns from Millipore. To form hairpin loops, single-strands were annealed by heating at 95° C for 5 minutes in: 10 mM Tris pH 7.4, 1 mM EDTA, 100mM NaCl. Incubation at 50°C for 1 hour was followed by gradual cooling to room temperature and then to 0°C prior to use. Oligodeoxynucleotide concentrations were estimated by measuring the optical density (OD) at 260 nm. To form duplexes, equimolar quantities of complementary strands were annealed in the same way.

0006-291X/97 $25.00

Enzymatic methylation. DNA methyltransferase from human placenta (20-90 units) (2) was incubated with DNA at the concentrations indicated and S-Adenosyl-L-[methyl-^{3}H]methionine (6μM; 15 Ci/mmol; Amersham) in a mixture containing 50mM HEPES (4-[2-hydroxyethyl]-1-piperazinethane-sulfonic acid) (pH 7.0), 50mM NaCl, 2mM dithiothreitol, 75 μM spermine and 10% v/v glycerol in a total volume of 100 μl. Purification procedures and the methods used for the determination of velocities with oligodeoxynucleotide substrates were as described in ref. 2.

Gel electrophoretic analysis of SSCs. For electrophoretic analysis, oligodeoxynucleotide SSCs were formed by annealing as described above. Samples were separated by electrophoresis through a 12% polyacrylamide non-denaturing gel or 20% polyacrylamide denaturing gels as described (15). DNAs were visualized by staining the gels in a solution of 0.02% (w/vol.) methylene blue (Mallinckrodt) followed by destaining in water. Images were recorded with a Speedlight Platinium gel documentation system (Lightools Research, Encinitas CA).

RESULTS

Like the CCG triplet from the *FMR*-1 gene of fragile X (15-17), the CAG triplet from the Huntington's locus can self-associate to form an SSC containing a repeating element composed of two G·C base pairs and a mispair (18). Although the oligodeoxynucleotides composed of repeating CAG triplets can form SSCs under physiological conditions, they are not good substrates for the human methyltransferase (Table 1) because they can not form the three-nucleotide recognition motif required for productive DNA methylation by the human enzyme (2,19). On the other hand, an oligodeoxynucleotide 48mer spanning a segment of the CAG repeat and its 3′ flank in exon1 of the Huntington's locus also forms an SSC. Relative to oligodeoxythymine standards, this oligodeoxynucleotide has the mobility expected for a 48mer when separated by electrophoreses under denaturing conditions. Under nondenaturing conditions, it forms a compact band with a mobility corresponding to that of a 26bp linear duplex, suggesting that the SSC is a single species (Fig. 1). Taken together, these properties suggest that it is folded by internal base-pairing into a unimolecular

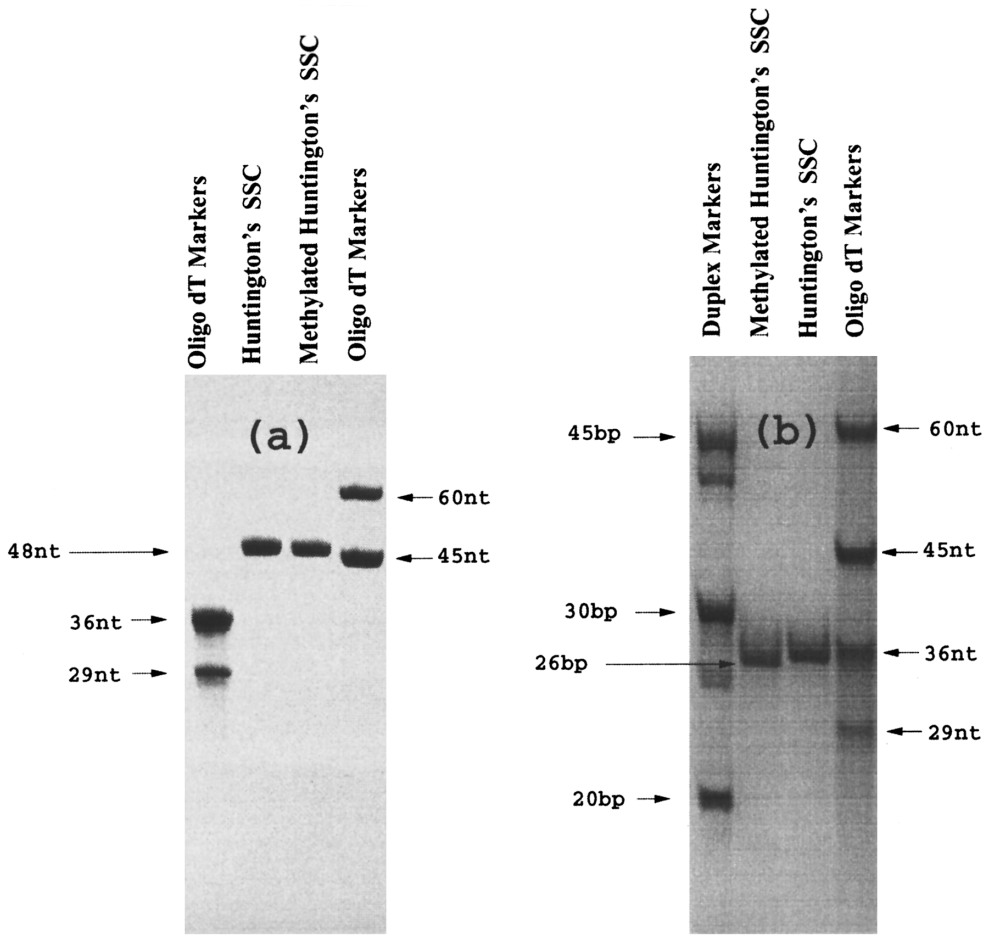

FIG. 1. Detection of secondary structure in oligodeoxynucleotides corresponding to Exon 1 nt 403-450 of the Huntington's gene. Markers are oligo dT molecules of the indicated lengths. (a) Electrophoretic separation under denaturing conditions. (b) Electrophoretic separation under nondenaturing conditions.

TABLE 1

Sequences[a]	3-Nucleotide[b] recognition sites: Number	3-Nucleotide[b] recognition sites: Type	Reaction[c] rate (fmol/min)
5′CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG3′	0	—	2 ± 0.8
5′CAGCAGCAGCAGCAGCAGCAG 3′GACGACGACGACGACGACGAC	0	—	
5′CAGCAGCAGCAGCAGCAGCAGCAGCAACAGCCGCCACCGCCGCCG3′	0	—	812 ± 76
5′CAGCAGCAGCAGCAGCAGCAG 3′GCCGCCGCCACCGCCGACAACGAC	4	C·N G·C	

[a] Sequences were single-stranded, but each adopts a single-stranded conformation (SSC) involving unimolecular folding. The purine-rich CAG triplet repeat is known to fold as shown under the conditions of the methyltransferase reaction (18).

[b] Human DNA (Cytosine-5)methyltransferase requires a Three-Nucleotide Recognition motif for active methylation of DNA. The mispaired cytosine (*) is targeted for methylation. Only the folded form of the Huntington's SSC offers this motif to the enzyme. One possible conformation for this folded SSC is shown in the table. This conformation presents four activated recognition sites to the enzyme (*i.e.*, sites in which the target cytosines are mispaired and are therefore transition state analogues). Other conformations are possible, but each can present no more than four activated motifs to the enzyme.

[c] Reaction rates are the mean of five independent determinations. The substrate oligodeoxynucleotide concentration was 1 μM in each case.

hairpin as reported for DNA from c-Ha-*ras,* and other DNA containing triplet repeats (15-18).

This SSC is recognized and rapidly methylated by the human methyltransferase (Table 1 and Fig. 2, left) with a Km of about 0.60μM. The rapid reaction rate observed under these conditions implies that the SSC contains one or more recognition sites for the human methyltransferase that are activated by an unusual conformation at the cytosines targeted by the methyltransferase (1,2,8,19). In control experiments (not shown) we found that M·*Sss*I did not methylate the SSC at a significant rate indicating that CG dinucleotide pairs held together by normal Watson-Crick base pairing are not present in the folded SSC (8). As noted above, a control CAG triplet repeat (Table 1, top) was not a good substrate for the human enzyme, suggesting that the CAG triplets occurring in the Huntington's SSC from exon1 (Table 1, bottom) are not the sites of methylation for the human enzyme. Three-nucleotide recognition motifs for the human methyltransferase can occur at any of the four CG dinucleotides in the Huntington's SSC from exon1. With this oligodeoxynucleotide the only two mispairs that can form at the targeted cytosine are the A·C mispair and the C·C mispair. Both are known to be structural analogs of the transition state for the enzymatic catalysis (1,2,8,19). Taken together these considerations suggest a tentative structure for the SSC in which a simple hairpin loop forms four methyltransferase recognition sites in which the A·C mispair presents a transition state analog (*i.e.* an unstacked cytosine) at each of the four recognition sites for the human enzyme (Table 1, and Fig. 2).

Although the reaction rate is rapid, it is important to note that the saturation curve is not hyperbolic at high substrate concentrations, indicating a pronounced substrate or product inhibition. When cytosines in the CG sites of the oligodeoxynucleotide are substituted with 5-methyldeoxy-cytosine during the chemical synthesis of the DNA, the methylated oligodeoxynucleotide is converted to a poor substrate for the enzyme (Fig. 2) indicating that the CG sites are the sites of enzymatic methylation in the product, as has been previously determined for this enzyme (2). Thus, a sequence in which these sites were methylated during chemical synthesis could be used to test for product inhibition.

The data (Fig. 2, right) indicate that the CG-methylated oligodeoxynucleotide is a potent inhibitor of the

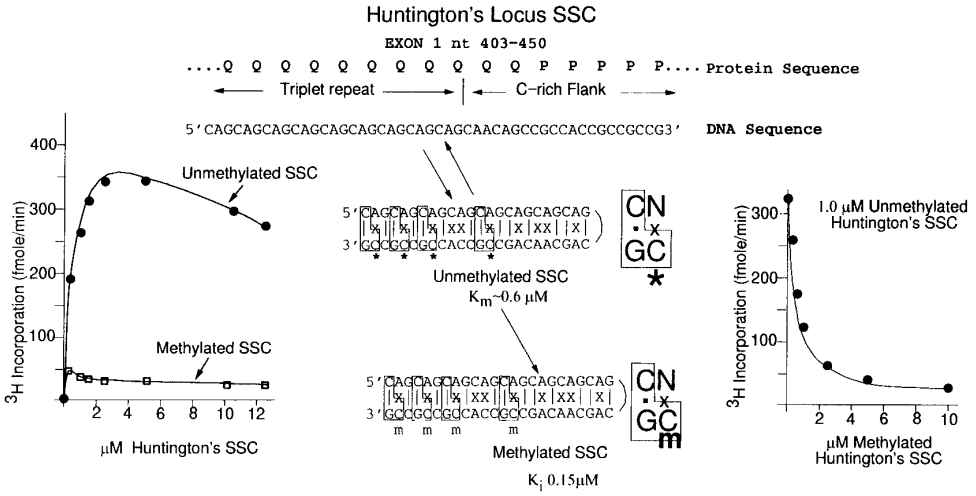

FIG. 2. Stalling at the SSC from the Huntington's locus. Rapid methylation of a folded structure (center) is shown in the saturation plot on the left. While the symmetrical folding depicted here is consistent with the electrophoretic data, asymmetrically folded (uneven) hairpins with altered pairing schemes may also form in a dynamic equilibrium. Each of the possible hair-pin structures is expected to produce mispairing at the CG dinucleotide. Recall that enzymes have a higher affinity for their catalytic transition states than they have for either their substrates or their products (18,19). As a consequence of their mechanism of action, which requires that the targeted cytosine unstack from the helix (1), DNA (cytosine-5) methyltransferases from both prokaryotes and eukaryotes necessarily recognize mispairs involving the cytosine in the CG site (*) as analogs of their catalytic transition state (1,2,7,8,10,11). This results in rapid methylation of the folded structure, as the saturation plot shows. Inhibition of the reaction by the asymmetrically methylated hairpin indicates the formation of an inactive complex between the enzyme and its methylated cytosine product in the hairpin.

methyltransferase with a K_i of about 0.15 μM. This result clearly demonstrates product inhibition through the formation of a tight complex between the methylated SSC and the enzyme.

DISCUSSION

The electrophoretic mobility of the Huntington's SSC requires that it contain internal base-pairs (Fig. 1). The active methylation of this SSC *in vitro* requires that at least one of the CG sites in the molecule participate in a three-nucleotide motif (Table 1) containing a target cytosine activated by mispairing so that it presents a transition state analog to the methyltransferase (1, 2, 8, 19). These results suggest a folded molecule of the type shown in Fig. 2 (center). Moreover, the active methylation of the SSC from the Huntington's locus suggests that the CG sites in the region flanking the CAG repeat will become methylated if they participate in an SSC during repeat expansion at the Huntington's locus. To the best of our knowledge, studies of the methylation state of the Huntington's locus have not been reported. Thus, the validity of this prediction can not be ascertained.

The central finding of the kinetic studies reported here is the formation of a tightly bound complex between an SSC and the human DNA methyltransferase. In addition to the SSC from the Huntington's locus, complexes between SSC's that form spontaneously in $(CCG)_{16}$, $(CAG)_{16}$, and $(CTG)_{16}$ 48mers have been detected in gel-retardation studies (in preparation). Each of these complexes was stable for at least 2.5 hrs at 37°C. Moreover, the complexes could also be stored frozen and subsequently isolated by electrophoresis. In addition, the complex formed between the enzyme and the Huntington's locus SSC studied here, has been observed in high molecular weight aggregates. The molecular basis for this aggregation is currently under investigation.

Neither the kinetic data presented here nor the electrophoretic data (in preparation) address the exact nature of the chemical linkages in the methyltransferase-SSC complex. A complex in which the product 5-methycytosine is bound at the active site is an attractive one. A reversible covalent link, through the enzyme nucleophile, to C6 of 5-methylcytosine is possible, as is the formation of an irreversible complex with a dimethyl substitution at C5. Our experiments do not distinguish between these possibilities.

Previous work with the human enzyme has clearly shown that the *de novo* reaction and the methyl-directed reactions are carried out by the same enzymes (7). This has now been confirmed for other eukaryotic enzymes like that from the mouse (20-21). Since the same enzyme carries out both reactions, we conclude that the human enzyme displays a capacity for differentiating between an SSC and a normally paired duplex. This capacity will cause it to probe nascent or

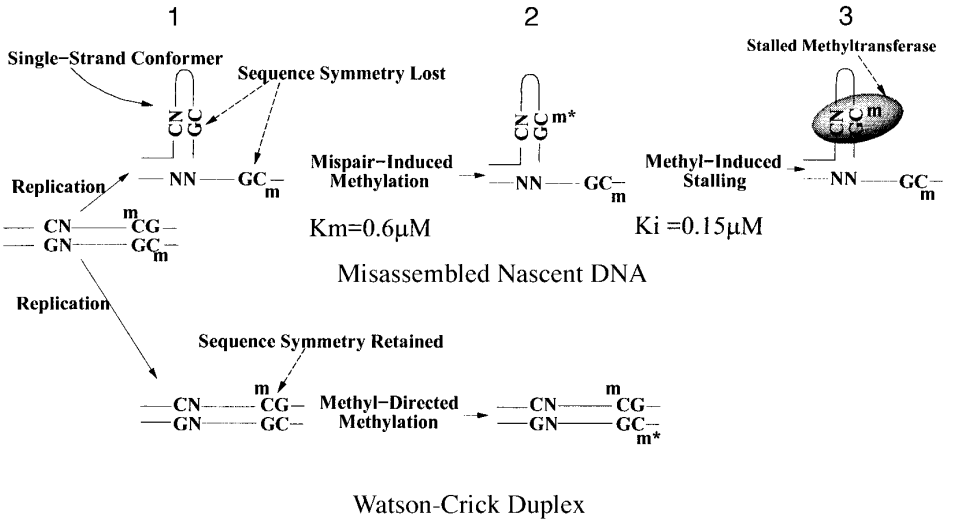

FIG. 3. A Model for Loop-Tagging by the Human Methyltransferase. Recognition of the asymmetrically methylated CG dinucleotide in the normally paired nascent duplex proceeds rapidly with quick release of the enzyme from the enzyme-substrate complex (lower track). In contrast, recognition of an SSC involves the recognition of a heteroduplex region in the stem-loop in which the CG dinucleotide has lost its normal two-fold symmetry because it is mispaired. The mispair targets the cytosine for active methylation (upper track). 1. As a consequence of the relaxed sequence specificity of these enzymes (2, 12), mispairs produced in the heteroduplex stem are methylated first at the cytosine residue on the strand segment containing the recognition motif (CG for the human enzyme). 2. Methyltransferase then probes the heteroduplex and stalls at all sites lacking a cytosine at a previously methylated site. 3. Methyltransferase stalled at the methylated mispair in the stem of the loop.

transcriptionally active DNA until it encounters an SSC where it will methylate mispaired cytosines at CG sites until it eventually stalls (Fig. 3). The enzymology does not resolve the question of whether or not the recognition is active or passive. Passive recognition would simply require that a biological mechanism exist for removing stalled methyltransferase from SSCs. Active recognition would imply that the methyltransferase, itself, is equipped to play an active role in biological processes that must distinguish between DNA that is in the normal Watson-Crick paired conformation and DNA that is conformationally unusual (*e.g.* a hairpin loop or misassembled replication intermediate). Processes of this type could include DNA repair, DNA recombination, and the maintenance of the differentiated state.

We favor active recognition and direct participation of the enzyme in such processes for two reasons. First, enzymes generally have higher affinities for their transition states than they do for their substrates or products (22-23). Thus, at the target site, the mispaired cytosine is a transition state analog for the reaction (2,3,7) which dominates the binding of the enzyme to the substrate. The stalling process that we have observed can be thought of as resulting from the inability of the enzyme to release a product that is fixed in a conformation similar to the transition state by an SSC. If the process were passive, there would be a strong selective pressure for methyltransferases to adopt a transition state structure that does not resemble the mispaired sites present in SSCs.

Second, unlike the comparison between asymmetrically methylated and unmethylated DNA, where the catalytic rate (as measured by V_{max}) is enhanced but the affinity (as measured by the K_m or $S_{0.5}$) is not (24), the affinity for the SSC is enhanced by methylation at CG sites (as measured by K_i) while the reaction rate (V_{max}) is dramatically reduced (Figure 2, center) until binding to the product is achieved. Each of these considerations suggests that the recognition is both active and highly evolved.

ACKNOWLEDGMENTS

This work was supported by Grant 0595 from the Smokeless Tobacco Research Council, Inc. S. Smith is a member of the Clinical Cancer Center of the City of Hope (CA33572).

REFERENCES

1. Baker, D. J., Kan, J. L. C., and Smith, S. S. (1988) *Gene* **74,** 207–210.
2. Smith, S. S., Kan, J. L. C., Baker, D. J., Kaplan, B. E., and Dembek, P. (1991) *J. Mol. Biol.* **217,** 39–51.
3. Smith, S. S. (1991) *Mol. Carcinogenesis* **4,** 91–92.
4. Lieb, M. (1987) *J. Bact.* **169,** 5241–5246.
5. Oberlé, T., Rousseau, F., Heitz, D., Kretz, C., Devys, D., Hanauer, A., Boué, J., Bertheas, M. F., and Mandel, J. L. (1991) *Science* **252,** 1097–1102.

6. Jones, C., Penny, L., Mattina, T., Yu, S., Baker, E., Voullaire, L., Langdon, W. Y., Sutherland, G. R., Richards, R. I., and Tunnacliffe, A. (1995) *Nature* **376,** 145–149.
7. De Bustros, A., Nelkin, B. D., Silverman, A., Ehrlich, G., Poiesz, B., and Baylin, S. B. (1988) *Proc. Natl. Acad. U.S.A.* **85,** 5693–5697.
8. Baker, D. J., Laayoun, A., and Smith, S. S. (1993) *Biochem. Biophys. Res. Commun.* **196,** 864–871.
9. Laayoun, A., Baker, D. J., Riley, J., and Smith, S. S. (1994) *Gene* **150,** 195–196.
10. Yang, A. S., Shen, J.-C., Zingg, J.-M., Mi, S., and Jones, P. A. (1995) *Nucleic Acids Res.* **23,** 1380–1387.
11. Klimasauskas, S., and Roberts, R. J. (1995) *Nucleic Acids Res.* **23,** 1388–1395.
12. Maxam, A. M., and Gilbert, W. (1980) *Methods in Enzymology* **65,** 499–560.
13. Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K., and Sekiya, T. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86,** 2766–2770.
14. Suzuki, Y., Orita, M., Shiraishi, M., Hayashi, K., and Sekiya, T. (1990) *Oncogene* **5,** 1037–1043.
15. Smith, S. S., Laayoun, A., Lingeman, R. A., Baker, D. J., and Riley, J. (1994) *J. Mol. Biol.* **243,** 143–151.
16. Chen, X., Mariappan, S. V. S., Catasti, P., Ratliff, R., Moyzis, R. K., Laayoun, A., Smith, S. S., Bradbury, E. M., and Gupta, G. (1995) *Proc. Natl. Acad. Sci. U.S.A.* **92,** 5199–5203.
17. Laayoun, A., and Smith, S. S. (1995) *Nucleic Acids Res.* **23,** 1584–1589.
18. Gacy, A. M., Goellner, G., Juranic, N., Macura, S., and McMurry, C. T. (1995) *Cell* **81,** 533–540.
19. Smith, S. S. (1994) In Cohn, W. E., and Moldave, K. (eds.), *Prog. Nucl. Acid Res. and Mol. Biol.*, Academic Press, San Diego, **49,** 65–111.
20. Tollefsbol, T. O., and Hutchison III, C. A. (1995) *J. Biol. Chem.* **270,** 18543–18550.
21. Flynn, J., Glickman, J. F., and Reich, N. O. (1996) *Biochemistry* **35,** 7308–7315.
22. Haldane, J. B. S. (1930) Enzymes, Longmans, Green, New York. [reprinted by MIT Press, Cambridge, MA, 1965]
23. Pauling, L. (1948) *American Scientist* **36,** 50–58.
24. Smith, S. S., Hardy, T. A., and Baker, D. J. (1987) *Nucleic Acids Res.* **15,** 6899–6916.