Targeted Chemical Nucleases

DAVID S. SIGMAN,* THOMAS W. BRUICE, ABHIJIT MAZUMDER, AND CHRISTOPHER L. SUTTON

Department of Biological Chemistry, School of Medicine, Department of Chemistry and Biochemistry, and Molecular Biology Institute, University of California, Los Angeles, California 90024-1570

Received June 18, 1992

The design of functional molecules has been an active area of research in chemistry and molecular biology during the past decade, encompassing topics such as host–guest chemistry,[1] ribozymes,[2] catalytic antibodies,[3] and most recently, fullerenes.[4] Nucleolytic agents that cleave the phosphodiester backbone of DNA and RNA have been another important focus of activity.[5] Interest in this subject has paralleled, and relied on, rapid advances in molecular biology, genome mapping and sequencing, and also the chemical synthesis of deoxyoligonucleotides and peptides. The goals of this research, the methodological approaches, and the insights into chemical biology provided by these molecules are the topic of this Account, which will emphasize research with the chemical nuclease activity of 1,10-phenanthroline–copper by our laboratory and others but will also summarize the innovative studies of Dervan and colleagues using ferrous–EDTA linked to targeting molecules.[6-8]

Chemical nucleases are defined here as redox active coordination complexes that cleave DNA by an oxidative pathway. The term "nuclease" is used as a shorthand for "nucleolytic activity". Since these reagents are generally used in excess of their substrates and also self-destruct because they generate reactive intermediates, turnover has not been rigorously demonstrated. In addition to 1,10-phenanthroline–copper[9] and ferrous–EDTA,[10] other chelates capable of cleaving the phosphodiester backbone of DNA under physiological conditions include metalloporphyrins,[11-13] uranyl acetate,[14] and octahedral rhodium complexes.[15-17] The generation of piperidine sensitive cleavage sites resulting from guanosine oxidation has been achieved with ruthenium[18] and nickel complexes.[19,20] Chemical methods for cleaving the phosphodiester backbone by hydrolysis or nucleophilic catalysis are not yet efficient enough to be useful.[21-23]

Nuclease Activity of 1,10-Phenanthroline–Copper. The nuclease activity of 1,10-phenanthroline–copper (OP–Cu) was reviewed seven years ago in these *Accounts*.[24] At that time, work had focused on the untargeted reactivity of the coordination complex. Briefly, the tetrahedral 2:1 1,10-phenanthroline cuprous complex binds reversibly to the minor groove of DNA and then is subjected to the one-electron oxidation by hydrogen peroxide.[25] The oxidative species formed then attacks the C-1 hydrogen of the deoxyribose, leading to a series of elimination reactions resulting in the production of 3′ and 5′ phosphorylated ends, free bases, and 5-methylene-2(5*H*)-furanone as stable products.[26,27]

$$(OP)_2Cu^+ + DNA \underset{}{\overset{K_d}{<\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!>}} (OP)_2Cu^+\text{--}DNA \xrightarrow[k_{cat}]{H_2O_2} \text{scission products} \quad (1)$$

(5′ and 3′ phosphomonoester ends, free bases and 5-methylene furanone)

The secondary structure specificity of OP–Cu in its scission of DNA,[28] its sequence-dependent reactivity in the scission of B-DNA,[29-33] and its exceptional

David S. Sigman was born in New York, NY, in 1939. After receiving his Ph.D. at Harvard in 1965 with F. H. Westheimer, he was a postdoctoral fellow with E. R. Blout at Harvard Medical School. Since joining the UCLA faculty in 1968, his research interests have encompassed enzymatic reaction mechanisms, nucleic acid structures, DNA–protein interactions, and the design of site-specific nucleases.

Thomas W. Bruice was born in Los Angeles, CA, in 1952. After receiving his Ph.D. at UCSF in 1980 with D. V. Santi and G. L. Kenyon, he was a postdoctoral fellow with J. R. Knowles at Harvard. He became a resident in clinical pathology at UCLA after receiving an M.D. from the University of Miami in 1984. After postdoctoral work with D. S. Sigman, he joined Isis Pharmaceuticals.

Abhijit Mazumder was born in Bombay, India, in 1965. He received his B.A. from Johns Hopkins University in 1986 and his Ph.D. at the University of Maryland in 1990 with J. A. Gerlt. He is currently a postdoctoral fellow with D. S. Sigman at UCLA.

Christopher L. Sutton was born in Omaha, NE, in 1961. He received his B.S. from the University of Wisconsin in 1984 and is currently engaged in graduate work at UCLA with D. S. Sigman.

* Address correspondence to this author at the Molecular Biology Institute. Phone 310-825-8903.

(1) Cram, D. J. *Science* **1988**, *240*, 760–767.
(2) Cech, T. R. *Science* **1987**, *236*, 1532–1539.
(3) Lerner, R. A.; Benkovic, S. J.; Schultz, P. G. *Science* **1991**, *252*, 659–667.
(4) Smalley, R. E. *Acc. Chem. Res.* **1992**, *25*, 98–105.
(5) Sigman, D. S. *Biochemistry* **1990**, *29*, 9097–9105.
(6) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861–6863.
(7) Dervan, P. B. *Science* **1986**, *232*, 464–471.
(8) Dervan, P. B. *Methods Enzymol.* **1991**, *208*, 497–515.
(9) Sigman, D. S.; Graham, D. R.; D'Aurora, V.; Stern, A. M. *J. Biol. Chem.* **1979**, *254*, 12269–12272.
(10) Hertzberg, R. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 313–315.
(11) Le Doan, T.; Perrouault, L.; Helene, C.; Chassignol, M.; Thuong, N. T. *Biochemistry* **1986**, *25*, 6736–6739.
(12) Ward, B.; Skorobogaty, A.; Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 6875–6883.
(13) Groves, J. T.; Farrell, T. P. *J. Am. Chem. Soc.* **1989**, *111*, 4998–5000.
(14) Nielsen, P. E.; Jeppesen, C.; Buchardt, O. *FEBS Lett.* **1988**, *235*, 122–124.
(15) Barton, J. K. *Science* **1986**, *233*, 727–734.
(16) Pyle, A. M.; Long, E. C.; Barton, J. K. *J. Am. Chem. Soc.* **1989**, *111*, 4520–4522.
(17) Pyle, A. M.; Morii, T.; Barton, J. K. *J. Am. Chem. Soc.* **1990**, *112*, 9432–9434.
(18) Mei, H. Y.; Barton, J. K. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1339–1343.
(19) Chen, X.; Burrows, C. J.; Rokita, S. E. *J. Am. Chem. Soc.* **1992**, *114*, 322–325.
(20) Chen, X.; Rokita, S. E.; Burrows, C. J. *J. Am. Chem. Soc.* **1991**, *113*, 5884–5886.
(21) Chin, J. *Acc. Chem. Res.* **1991**, *24*, 145–152.
(22) Modak, A. S.; Gard, J. K.; Merriman, M. C.; Winkeler, K. A.; Bashkin, J. K.; Stern, M. K. *J. Am. Chem. Soc.* **1991**, *113*, 283–291.
(23) Morrow, J. R.; Buttrey, L. A.; Shelton, V. M.; Berback, K. A. *J. Am. Chem. Soc.* **1992**, *114*, 1903–1905.
(24) Sigman, D. S. *Acc. Chem. Res.* **1986**, *19*, 180–186.

0001-4842/93/0126-0098$04.00/0

reactivity with single-stranded DNA formed at transcription start sites[34,35] are manifestations of the stability of the essential noncovalent intermediate. The copper complex of 5-phenyl-1,10-phenanthroline shows particular specificity at transcription start sites[35] and in regions where the DNA is bent by protein binding.[36] Enhanced binding of this complex resulting from the intercalation of the pendant phenyl group into the underwound DNA most likely underlies this increased reactivity.

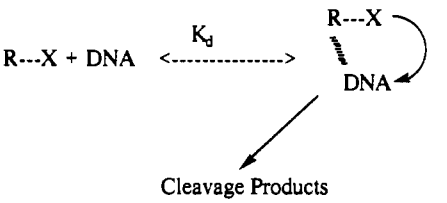
In its reaction with RNA, OP–Cu exhibits a strong kinetic preference for single-stranded loop regions relative to double-stranded stem structures.[37] The molecular structure of the intermediate between the tetrahedral complex and loop regions remains unclear. However, recent crystallographic studies have indicated that unexpected hydrogen-bond patterns appear in RNAs.[38-40] Therefore, all loops may have some residual structure which might generate a more favorable binding site for the tetrahedral 2:1 1,10-phenanthroline–cuprous complex than the A-helix of the stem structure. The high affinity for the dyes cibracon blue and reactive blue 4 that can be demonstrated in RNAs probably arises from the variety of unusual structures that may be generated in loop regions.[41] This finding was anticipated by earlier studies which noted the potent inhibition of enzymes by aromatic dyes unrelated in structure to normal substrates.[42]

Nucleolytic Activity of Ferrous–EDTA. In contrast to OP–Cu, ferrous–EDTA does not form a complex with nucleic acids in its scission reaction with RNA and DNA and therefore is not selective unless it is linked to a ligand with specificity for some aspect of nucleic acid structure.[43-46] It generates hydroxyl radicals and has been widely used for "hydroxyl radical footprinting" because the *free* coordination complex cuts evenly at all sequence positions as long as the phosphodiester backbone is solvent accessible. In its reaction with RNA, it has been used to distinguish nucleotides on the exterior from those on the interior.[45]

(25) Marshall, L. E.; Graham, D. R.; Reich, K. A.; Sigman, D. S. *Biochemistry* **1981**, *20*, 244–250.
(26) Kuwabara, M.; Yoon, C.; Goyne, T. E.; Thederahn, T.; Sigman, D. S. *Biochemistry* **1986**, *25*, 7401–7408.
(27) Goyne, T. E.; Sigman, D. S. *J. Am. Chem. Soc.* **1987**, *109*, 2846–2848.
(28) Pope, L. E.; Sigman, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3–7.
(29) Yoon, C.; Kuwabara, M. D.; Law, R.; Wall, R.; Sigman, D. S. *J. Biol. Chem.* **1988**, *263*, 8458–8463.
(30) Yoon, C.; Kuwabara, M. D.; Spassky, A.; Sigman, D. S. *Biochemistry* **1990**, *29*, 2116–2121.
(31) Veal, J. M.; Merchant, K.; Rill, R. L. *Nucleic Acids Res.* **1991**, *19*, 3383–3388.
(32) Veal, J. M.; Rill, R. L. *Biochemistry* **1991**, *30*, 1132–1140.
(33) Veal, J. M.; Rill, R. L. *Biochemistry* **1989**, *28*, 3243–3250.
(34) Spassky, A.; Sigman, D. S. *Biochemistry* **1985**, *24*, 8050–8056.
(35) Thederahn, T.; Spassky, A.; Kuwabara, M. D.; Sigman, D. S. *Biochem. Biophys. Res. Commun.* **1990**, *168*, 756–762.
(36) Frantz, B.; O'Halloran, T. V. *Biochemistry* **1990**, *29*, 4747–4751.
(37) Murakawa, G. J.; Chen, C.-H. B.; Kuwabara, M. D.; Nierlich, D.; Sigman, D. S. *Nucleic Acids Res.* **1989**, *17*, 5361–5369.
(38) Varani, G.; Cheong, C.; Tinoco, I., Jr. *Biochemistry* **1991**, *30*, 3280–3289.
(39) Bartel, D. P.; Zapp, M. L.; Green, M. R.; Szostak, J. W. *Cell* **1991**, *67*, 529–536.
(40) Chastain, M.; Tinoco, I., Jr. *Prog. Nucleic Acid Res. Mol. Biol.* **1991**, *41*, 131–177.
(41) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818–822.
(42) Glazer, A. N. *Proc. Natl. Acad. Sci. U.S.A.* **1970**, *65*, 1057–1063.
(43) Tullius, T. D.; Dombroski, B. A. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 5469–5473.
(44) Dixon, W. J.; Hayes, J. J.; Levin, J. R.; Weidner, M. F.; Dombroski, B. A.; Tullius, T. D. *Methods Enzymol.* **1991**, *208*, 380–413.
(45) Celander, D. W.; Cech, T. R. *Biochemistry* **1990**, *29*, 1355–1361.
(46) Hertzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934–3945.

Goals in the Design of Chemical Nucleases

The design of chemical nucleases relies on the kinetic scheme postulated above for the chemical nuclease activity of OP–Cu. Like affinity labeling reagents used in the site-specific modification of an enzyme, one aspect of their structure is responsible for recognition (R in eq 2) and the second is devoted to chemical attack on the DNA target (X in eq 2).[47] R can be an oligoribonucleotide, a protein/peptide, or a DNA-binding drug; X can be 1,10-phenanthroline–copper, ferrous–EDTA, or a metalloporphyrin. The formation of a productive noncovalent intermediate provides specificity in the reaction.



(2)

The obvious challenge in the design of chemical nucleases is choosing a recognition element (R) specific for the desired target that is consistent with the reactive properties of X. In the case of OP–Cu, the coordination complex must be directed or accessible to the C-1 hydrogen of the deoxyribose. In B-DNA, the C-1 hydrogen is located at the floor of the minor groove. For ferrous–EDTA, the scission chemistry apparently can be initiated by oxidation at either the C-1 or C-4 hydrogens in the minor groove or the C-3 hydrogen in the major groove.

Chemical nucleases have been useful in two contexts. In the first, the pattern of scission observed with the chemical nuclease has been used to provide information on the binding of the affinity element.[7] This approach for studying the binding of ligands to nucleic acids has been called affinity cleaving and bears a strong resemblance to photoaffinity labels for enzymes.[48] It is not necessary for these nucleases to be efficient to provide valuable insights into the structure of biologically important complexes under physiological conditions.

The second motivation for designing chemical nucleases is to transform the carrier into an efficient cleaver. This application not only requires the specificity inherent in affinity cleaving but also demands that the covalent chemistry be robust. For example, if the goal is to design alternatives to restriction enzymes, the yield of the scission reaction must be high and the product distribution limited. Alternatively, if an oligonucleotide is synthesized as an antisense reagent, linking it to a cutting reagent should enhance its ability to inactivate a messenger RNA. Presently, chemical nucleases have been more useful in defining ligand binding than in generating a family of new reagents or drugs.

(47) Glazer, A. N.; DeLange, R.; Sigman, D. S. *Chemical Modification of Proteins. Selected Methods and Analytical Procedures*; North-Holland American Elsevier Publishing Company: New York, 1975.
(48) Chowdhry, V.; Westheimer, F. H. *Annu. Rev. Biochem.* **1979**, *48*, 293–325.

Chemical Nucleases Linked to Oligonucleotides and Proteins

Seven years ago, our research program began focusing on the design of site-specific nucleases using OP–Cu.[49] Among the reasons that we concentrated on this chemical nuclease activity were our familiarity with its reactive properties, its known chemical mechanism of scission, and the ready availability of substituted phenanthrolines which could allow its convenient coupling to both proteins and nucleic acids. In addition, since the chemistry did not depend on diffusible hydroxyl radicals, background scission might be minimal. This is important for the design of scission reagents of high molecular weight DNAs. Although our work has been weighted toward the protein targeting of OP–Cu, results obtained using deoxyoligonucleotides as targeting vehicles will be briefly reviewed.

Oligonucleotide-Linked Reagents. The first example of a targeted scission reagent using OP–Cu was synthesized by linking a 21-nucleotide-long deoxyoligonucleotide at its 5-phosphorylated end to 5-glycylamido-1,10-phenanthroline by condensation with a water soluble carbodiimide.[49] The scission pattern observed in the complementary strand was consistent with the formation of a double-helical structure with antiparallel strands and established that one phenanthroline was sufficient to ensure the chemical reactivity of the copper–oxo species. Two phenanthrolines per copper ion are essential without a targeting ligand.[25]

Oligonucleotide-targeted scission of DNA using the nucleolytic activity of ferrous–EDTA had previously been demonstrated using a carrier linked at the 5′-end[50] and at an interior nucleotide.[51] Strict comparison between the efficiency of the scission using ferrous–EDTA and 1,10-phenanthroline as the nucleolytic agents is not possible although the scission by 1,10-phenanthroline–copper is less diffuse and concentrated in fewer bands.

Scission of RNA. OP-linked deoxyoligonucleotides also cleave complementary RNA sequences with rates comparable to that observed with DNA yielding patterns anticipated for antiparallel hybridization.[52] If the target sequence is within a stable stem structure, the RNA must be denatured to permit the hybridization of the OP-linked deoxyoligonucleotide. Disruption of the RNA structure is not necessary if the loop region is the target. Although RNase H, which is specific for RNA–DNA heteroduplexes, can be used in conjunction with deoxyoligonucleotides to detect single-stranded regions, OP-linked deoxyoligonucleotides have the advantage that they can nick nucleotides which are adjacent in space but not continuous in sequence. In Figure 1, the scission sites in the stem loop are presented. The nicks at 13–15 are anticipated, but those at +3 and +4 must arise because these sequence positions are adjacent to the 1,10-phenanthroline of the hybridized heptamer. These initial experiments have clearly indicated that linking the nucleolytic activity to an oligonucleotide will not improve its efficiency as an antisense reagent. The complexes are simply too flexible for the cleavage reaction to be very robust. The principle utility of these targeted nucleases is to explore the conformation of RNAs.

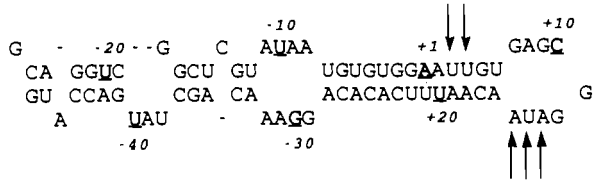


Figure 1. Sites of scission in an *E. coli* lac mRNA fragment by 5′-OP–ATCCGCT-3′, a deoxyoligonucleotide complementary to the loop.

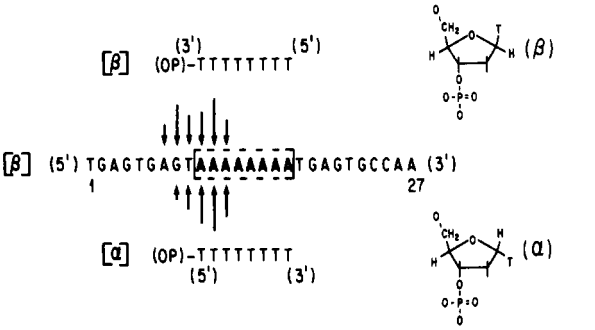


Figure 2. Scission patterns indicate that abiological α-deoxyoligonucleotides bind parallel to their complementary sequence while natural β-deoxyoligonucleotides bind antiparallel.

Novel Mechanisms of Hybridization. Targeted chemical nucleases have been useful for probing the hybridization of biological nucleotides and the formation of triple-stranded structures. For example, as part of their innovative program of developing modified oligonucleotides as specific inhibitors of gene expression, Helene and colleagues have explored the hybridization properties of nuclease resistant α-deoxyoligonucleotides by affinity cleaving.[53,54] As demonstrated in Figure 2, these abiological deoxyoligonucleotides hybridize with an orientation parallel to the phosphodiester backbone.[54]

Triple-stranded structures were initially proposed 25 years ago to account for a complex, detected spectroscopically, composed of two poly U strands and one poly A strand. In the model, a poly U strand (the guest) binds in the major groove of a poly A:poly U duplex (the host) using Hoogsteen hydrogen bonds. Analogous structures have been proposed for guanosine and cytosine containing polynucleotides although protonation of cytosine is required to form stable Hoogsteen base pairs. Intramolecular triple-stranded structures appear responsible for the sensitivity of DNA within chromatin to S-1 and single-strand specific chemical reagents.[55,56]

Since triplex formation does not require the disruption of a preexisting duplex structure, it is a potentially powerful method for targeting a homopurine–homopyrimidine stretch for scission either to modulate the biological activity of a gene or to target a given sequence

(49) Chen, C.-H. B.; Sigman, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 7147–7151.
(50) Chu, B. C. F.; Wahl, G. M.; Orgel, L. E. *Nucleic Acids Res.* **1983**, *11*, 6513–6529.
(51) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528–5529.
(52) Chen, C.-H. B.; Sigman, D. S. *J. Am. Chem. Soc.* **1988**, *110*, 6570–6572.
(53) Helene, C.; Toulme, J.-J. *Biochim. Biophys. Acta* **1990**, *1049*, 99–125.
(54) Sun, J. S.; Francois, J. C.; Lavery, R.; Saison-Behmoaras, T.; Montenay-Garestier, T.; Thuong, N. T.; Helene, C. *Biochemistry* **1988**, *27*, 6039–6045.
(55) Kohwi, Y.; Kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3781–3785.
(56) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D.; Cantor, C. R. *Nucleic Acids Res.* **1988**, *16*, 2165–2178.

domain for scission. While high-resolution NMR studies are providing detailed structures of triplexes,[57-59] affinity cleaving using both ferrous–EDTA and OP-linked deoxyoligonucleotides is a rapid and reliable approach for investigating strand polarity, stringency, and stability of these molecular complexes.[60-65] These results will influence the potential application of triplex formation in the modulation of gene expression as well as in the development of scission reagents suitable for chromosomal mapping. Initial results in the cleavage of double-stranded DNA by these reagents have been reported for SV-40 DNA (about 5 kB),[64] λ phage DNA (about 50 kB),[66] and yeast chromosome III (about 340 kB)[67] into which a target sequence was inserted by homologous recombination. Yields in the latter case were about 20–30% with deoxyoligonucleotides approximately 20 bases long.

Targeting the Chemical Nucleases with DNA-Binding Proteins. DNA-binding proteins (e.g., repressors, transcription factors) derivatized with scission reagents can be used to probe their binding properties. In addition, chimeric proteins efficient in DNA scission might be useful in analyzing high molecular weight DNA because they would recognize DNA sequences 17 bp long. The first example of the transformation of a DNA-binding protein into a DNA cutter was carried out with the *Escherichia coli* Trp repressor which represses its own synthesis as well as three operons involved in tryptophan metabolism in the presence of L-tryptophan.[68-72] Since the protein lacks cysteine residues, it was derivatized with 1,10-phenanthroline by modification of the lysine residues with iminothiolane[73] and then alkylation of the resulting sulfhydryl groups with 5-(iodoacetamido)-1,10-phenanthroline. The modified protein retained high-affinity DNA binding as measured by DNase I footprinting and cleaved its target operators in a reaction that required its corepressor, L-tryptophan. The scission patterns with the *aro H* and *trp EDCBA* operators indicated that Trp repressor binds to these two operators in different orientations.[68] However, for this conjugated protein, double-stranded scission was relatively inefficient: only about 17%.

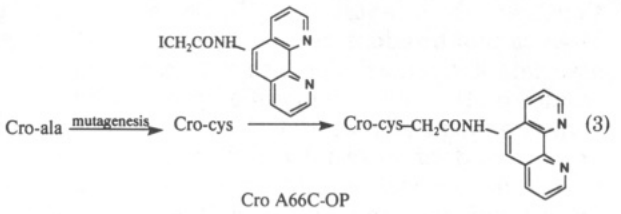
The likelihood of converting any given DNA-binding protein into a nucleolytic agent by the two-step modification protocol is not predictable. Even if a productive scission reagent is generated, subtle questions as to the orientation of the protein to the target DNA cannot be addressed because it is not possible to associate a given scission site with a particular lysyl residue on the multiply modified protein. In order to generate chemically discrete chimeric proteins capable of efficient double-stranded DNA scission, we decided to introduce cysteine residues at unique positions in DNA-binding proteins genetically and then alkylate them with 5-(iodoacetamido)-1,10-phenanthroline. The proteins chosen were the λ phage cro protein and the Trp repressor.

While this work was in progress, Ebright and colleagues chemically transformed the wild-type *E. coli* catabolite activating protein into a site-specific nuclease by alkylating the unique cysteine residue (cysteine 178) at amino acid position 10 of the helix–turn–helix domain.[74] Although alkylation decreases the DNA-binding affinity to the 22-base-pair recognition sequence by a factor of 400, scission is observed at exactly the positions predicted from the X-ray structure.[75] It therefore provides strong confirmation of the congruence of the crystal results with those in solution.

Targeting Scission with the Cro Protein. In pioneering studies on the cro protein, Matthews and colleagues first recognized the helix–turn–helix motif that is frequently found in prokaryotic DNA-binding proteins.[76-78] Model-building studies further suggested that the five C-terminal residues of cro were accessible to the minor groove, possibly contributing additional stability to binding.[79] NMR and crystallographic studies have indicated that these residues are flexible.[80-82] Site-directed mutagenesis of this region has shown that these residues are not the principal source of stability for the protein–DNA interaction.[83]

Nevertheless, the conversion of this protein into a site-specific scission reagent was appealing because it was possible to direct the nucleolytic activity of 1,10-phenanthroline–copper toward its chemically reactive site in the minor groove without diminishing its high-affinity binding in the major groove. The design of our experiment is indicated schematically in eq 3.[84,85] To achieve this goal, the C-terminal alanine residue was changed into a cysteine residue by site-directed mutagenesis. The protein was then expressed in bacteria. Rather than purifying the protein to homogeneity, bacterial cell extracts were modified directly by 5-(iodoacetamido)-1,10-phenanthroline. A gel retardation assay with the high-affinity operator OR-3 as probe (labeled on either strand) was used to purify cro A66C–OP from excess 5-(iodoacetamido)-1,10-phenanthroline

(57) Rajagopal, P.; Feigon, J. *Nature* **1989**, *339*, 637–640.
(58) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859–7870.
(59) Sklenar, V.; Feigon, J. *Nature* **1990**, *345*, 836–838.
(60) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645–650.
(61) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967–971.
(62) Praseuth, D.; Perrouault, L.; Le Doan, T.; Chassignol, M.; Thuong, N.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349–1353.
(63) Francois, J.-C.; Saison-Behmoaras, T.; Helene, C. *Nucleic Acids Res.* **1988**, *16*, 11431–11440.
(64) Francois, J.-C.; Saison-Behmoaras, T.; Barbier, C.; Chassignol, M.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9702–9706.
(65) Helene, C.; Thuong, N. T.; Saison-Behmoaras, T.; Francois, J.-C. *Trends Biotechnol.* **1989**, *7*, 310–315.
(66) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927–7929.
(67) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73–75.
(68) Chen, C.-H. B.; Sigman, D. S. *Science* **1987**, *237*, 1197–1201.
(69) Gunsalus, R. P.; Yanofsky, C. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 7117–7121.
(70) Arvidson, D. N.; Bruce, C.; Gunsalus, R. P. *J. Biol. Chem.* **1986**, *261*, 238–243.
(71) Heatwole, V. M.; Somerville, R. L. *J. Bacteriol.* **1991**, *173*, 108–115.
(72) Sarsero, J. P.; Wookey, P. J.; Pittard, A. J. *J. Bacteriol.* **1991**, *173*, 4133–4143.
(73) Jue, R.; Lambert, J. M.; Pierce, L. R.; Traut, R. R. *Biochemistry* **1978**, *17*, 5399–5304.
(74) Ebright, R. H.; Ebright, Y. W.; Pendergast, P. S.; Gunasekera, A. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 2882–2886.
(75) Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 208–280.
(76) Anderson, W. F.; Ohlendorf, D. H.; Takeda, Y.; Matthews, B. W. *Nature* **1981**, *290*, 754–758.
(77) Brennan, R. G.; Matthews, B. W. *Trends Biochem. Sci.* **1989**, *14*, 286–290.
(78) Brennan, R. G.; Matthews, B. W. *J. Biol. Chem.* **1989**, *264*, 1903–1906.
(79) Matthew, J. B.; Ohlendorf, D. H. *J. Biol. Chem.* **1985**, *260*, 5860–5862.
(80) Leighton, P.; Lu, P. *Biochemistry* **1987**, *26*, 7262–7271.
(81) Metzler, W. J.; Arndt, K.; Tecza, E.; Wasilewski, J.; Lu, P. *Biochemistry* **1985**, *24*, 1418–1424.
(82) Brennan, R. G.; Roderick, S. L.; Takeda, Y.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8165–8169.
(83) Hubbard, A. J.; Bracco, L. P.; Eisenbeis, S. J.; Gayle, R. B.; Beaton, G.; Caruthers, M. H. *Biochemistry* **1990**, *29*, 9241–9249.
(84) Bruice, T. W.; Wise, J.; Sigman, D. S. *Biochemistry* **1990**, *29*, 2185–2186.
(85) Bruice, T. W.; Wise, J.; Rosser, D. S. E.; Sigman, D. S. *J. Am. Chem. Soc.* **1991**, *113*, 5446–5447.



and any other 1,10-phenanthroline-modified proteins present in the extract.

Cro A66C–OP retains high-affinity binding comparable to that of the wild-type protein. The scission chemistry was carried out by activating the chemical nuclease activity within the gel slice by adding cupric ion, 3-mercaptopropionic acid, and hydrogen peroxide, quenching the reaction after 10 min with 2,9-dimethyl-1,10-phenanthroline, eluting the scission products, and analyzing them on a sequencing gel. The sites of scission were *within* the recognition sequence and were exactly the nucleotides that were *protected* when the cro–OR-3 complex was footprinted within the gel matrix using 1,10-phenanthroline–copper[86] (Figure 3). The minor groove reactivity of cro A66C–OP, an essential feature of the design of the chimeric catalyst, can be inferred by the 3′-stagger of the sequence positions of the prominent scission sites (Figure 4). Since the two strands of B-DNA are antiparallel, the deoxyribose moieties closest to the OP–Cu within the minor groove are displaced two to three sequence positions with respect to one another in the 3′-direction.[87]

How effective is cro A66C–OP as a targeted nuclease? Let us consider the satisfactory aspects of the design, the deficiencies, and possible approaches for improvement. As an affinity cleaving reagent, the essential conclusions of the X-ray, NMR, and site directed mutagenesis experiments have been confirmed. Gel retardation assays indicate that cro A66C–OP binds to OR-3 with affinity comparable to that of the wild-type protein. The C-terminus of cro is clearly accessible to the minor groove and able to approach the reactive C-1 hydrogen of the deoxyribose on either strand. The observation of cleavage over the entire recognition sequence most likely reflects the flexibility of the C-terminal arm.

Roughly 40% of either the Watson or Crick strand is nicked within a 10-min incubation by OP–Cu linked to the C-terminal end of cro. At a concentration of the complex of 1×10^{-9} M, this is at least 10 times more rapid than the rate observed with the (nonsaturating) 30 μM 2:1 1,10-phenanthroline–copper complex used in the footprinting reaction in the presence of the protein. Linking the nuclease activity to the C-terminus of cro, therefore, increases the rate of scission within the recognition sequence by a factor of greater than 3×10^5. The efficiency of cro A66C–OP can also be expressed as the ratio of the turnover number (1.5×10^{-3} s^{-1}) to the dissociation constant (10^{-9} M). This value is 1.5×10^6 M^{-1} s^{-1}, indicating that cro A66C–OP is a respectable, if not perfect, catalyst.[88]

The principal problem with cro A66C–OP as a nuclease is not its rate of single-stranded scission but rather its inefficiency in causing double-stranded breaks. Under conditions where each strand is cleaved

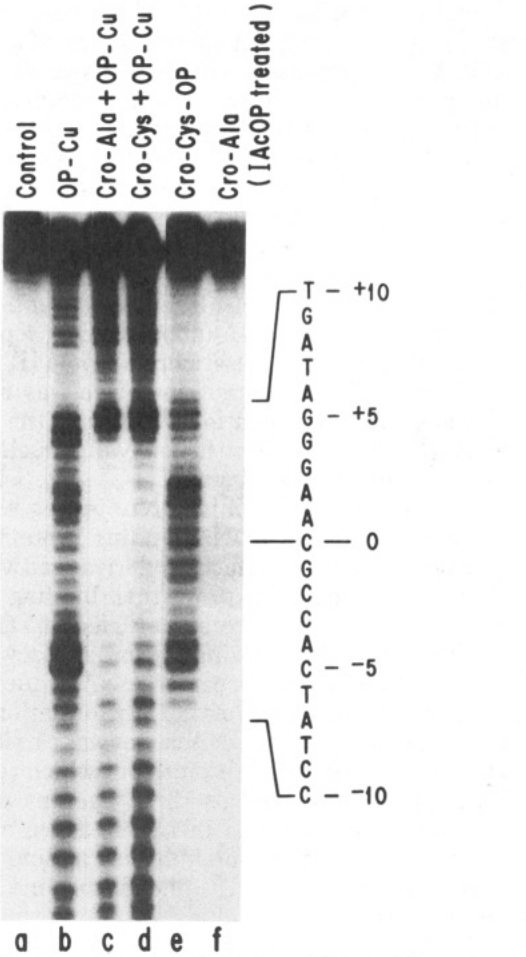


Figure 3. Cro–protein targeted scission of OR-3. All reactions were carried out in gel on DNA and protein–DNA complexes isolated using a gel retardation assay. See ref 78 for experimental details. Lane a is 5′ ^{32}P labeled OR-3 operator DNA. Lanes b–d are the OP–Cu digestions of OR-3. Lanes c and d contain wild-type cro and cro A66C footprints, respectively. Lane e shows the site-specific digestion pattern of cro A66C–OP. Lane f is a control of wild-type cro (cro-Ala) that has been treated with 5-(iodoacetamido)-1,10-phenanthroline. Note: Scission by cro A66C–OP is observed within the recognition sequence.

to the extent of 20%, the yield of double-stranded breaks is only about 5%. This suggests that the single-stranded nicks are independent of one another and that the probability of double-stranded breaks is just the product of the probabilities that each single-stranded break will occur within two nucleotides of each other.

Perhaps cro A66C–OP is poor at accomplishing double-stranded breaks because it nicks both strands throughout the recognition sequence. Two features of cro A66C–OP lead to this diffuse pattern of scission sites in addition to the possible diffusion of oxidative species responsible for strand scission. The first is the flexibility of the C-terminal arms. The second, and possibly more crucial, is their antiparallel orientation within the dimeric protein. This places the 1,10-phenanthroline–copper of each subunit at the outer limits of OR-3 and works against the clustering of the cleavage sites and the likelihood of double-stranded breaks. Our two goals in improving the design of protein-targeted nucleases therefore were to attach the phenanthroline (a) to a rigidly determined part of the protein and (b) near its 2-fold dyad axis.

Targeting Scission with *E. coli* Trp Repressor. These plans were executed with the Trp repressor

(86) Kuwabara, M. D.; Sigman, D. S. *Biochemistry* **1987**, *26*, 7234–7238.
(87) Drew, H. R. *J. Mol. Biol.* **1984**, *176*, 535–557.
(88) Hackney, D. *Enzymes* **1991**, *19*, 1.

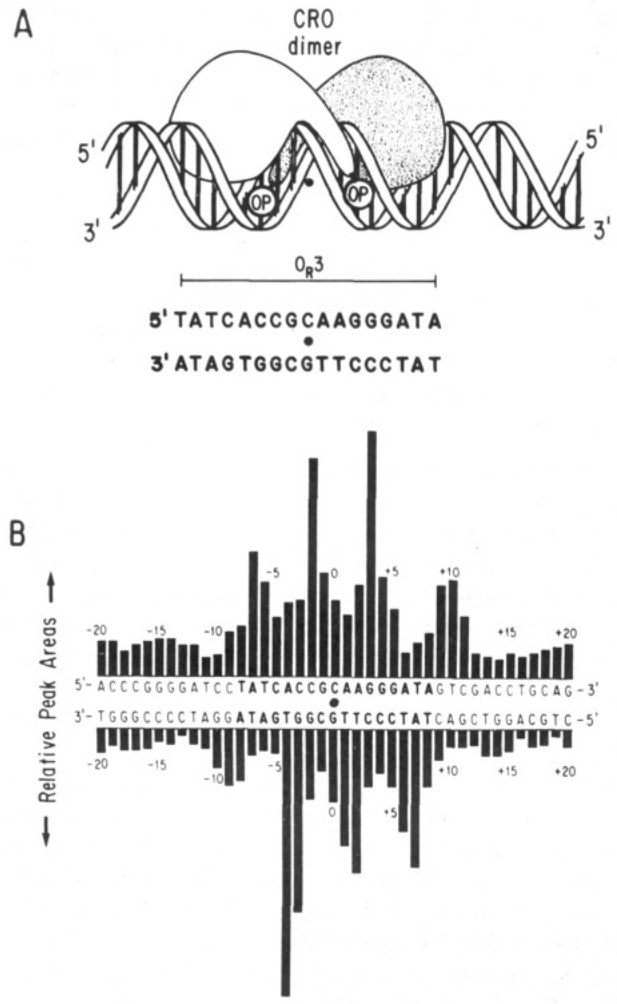


Figure 4. (A) Model of the cro dimer with the covalently attached OPs in the minor groove of the DNA. (B) Histogram of the single-stranded scission patterns of the cro A66C–OP nuclease on the OR-3 operator. The sequence position numbering scheme is dictated by the upper strand (corresponding to the sense strand of the cro gene in λ) by convention, with "O" increasingly positive and positions 5′ increasingly negative. Scission of both strands of OR-3 by cro A66C–OP was assayed as single-stranded nicks.

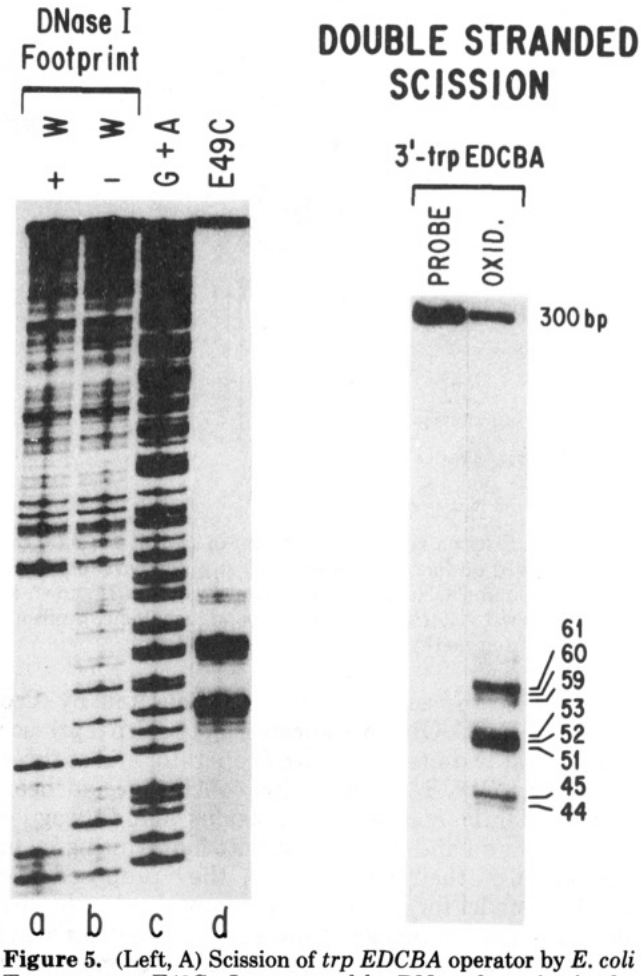


Figure 5. (Left, A) Scission of *trp EDCBA* operator by *E. coli* Trp repressor E49C. Lanes a and b: DNase footprint in the presence and absence of tryptophan. Lane c: G + A sequencing lane. Lane d: scission by *E. coli* Trp repressor upon addition of cupric, thiol, and hydrogen peroxide. (Right, B) Double-stranded scission of *trp EDCBA*.

because of the availability of a structure of the Trp repressor–*trp EDCBA* complex which suggested that preparation of the mutant Trp repressor E49C would achieve the dual design objectives cited above.[89,90] Since the cocrystal structure has been dismissed (and supported[91]) as a nonspecific protein–DNA complex,[92] its ability to predict scission sites would provide a test of the model. The controversial features of this model are its suggestion that subtle differences in the geometry of the recognition sequence rather than specific amino acid–nucleotide interactions are responsible for the high-affinity site-specific binding.[93] The structure of the cocrystal has indicated that the protein binds to DNA primarily by making hydrogen bonds to the phosphodiester backbone either directly or mediated through hydrogen bonds to water.[90]

Using a protocol parallel to that outlined above for the cro protein, Trp repressor E49C–OP was isolated by a gel retardation assay using a labeled DNA fragment containing either the *trp EDCBA* or *aro H* operator. Scission was activated within the gel slice in the presence of the corepressor L-tryptophan, and the products were isolated and assayed as single-stranded nicks (Figure 5A) or double-stranded breaks (Figure 5B). In contrast with the results obtained with cro A66C–OP, the scission sites are highly localized and the rate of formation of double-stranded breaks is within a factor of 2–3 of that for single-stranded nicks. The robust double-stranded scission catalyzed by Trp repressor E49C–OP is competitive with naturally occurring restriction enzymes with respect to rate and yield.[94]

As an affinity cleaving reagent, Trp repressor E49C–OP is highly informative with respect to the model of the protein's binding to DNA. Since amino acid 49 is adjacent to the dyad axis of the protein structure, Trp repressor E49C–OP should cut the target DNA at this symmetry position of the recognition sequence. Our results are fully supportive of the cocrystal structure because the major site of scission of *trp EDCBA* is at the principal dyad as anticipated. We conclude that the protein–DNA complex studied by X-ray crystallography is biologically relevant.

(89) Schevitz, R. W.; Otwinowski, Z.; Joachimiak, A.; Lawson, C. L.; Sigler, P. B. *Nature* **1985**, *317*, 782–786.
(90) Otwinowski, Z.; Schevitz, R. W.; Zhang, R. G.; Lawson, C. L.; Joachimiak, A.; Marmorstein, R. Q.; Luisi, B. F.; Sigler, P. B. *Nature* **1988**, *335*, 321–329.
(91) Carey, J.; Lewis, D. E. A.; Lavoie, T. A.; Yang, J. *J. Biol. Chem.* **1991**, *266*, 24509–24513.
(92) Staacke, D.; Walter, B.; Kisters-Woike, B.; Wilcken-Bergmann, B. v.; Muller-Hill, B. *EMBO J.* **1990**, *9*, 1963–1967.
(93) Matthews, B. W. *Nature* **1988**, *335*, 294–295.
(94) Sutton, C.; Mazumder, A.; Chen, C.-h. B.; Sigman, D. S. *Biochemistry*, submitted.

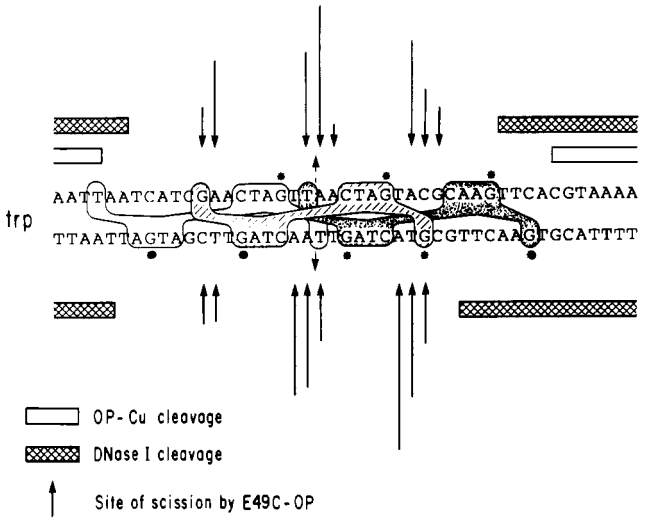


Figure 6. Scission sites (solid arrows) of *E. coli trp EDCBA* superimposed on footprints (open and stippled bars) obtained with OP–Cu and DNase I. The central dyad axis is shown as dashed arrows. Asterisks above correspond to sequence positions protected from methylation by dimethyl sulfate.

However, the scission patterns generated by Trp repressor E49C–OP also indicate that the Trp repressor can bind at two sites in the *aro H* operator and at three in the *trp EDCBA* locus. This conclusion had been independently reached by Gunsalus and colleagues using DNase I and dimethyl sulfate footprinting.[95] On the basis of their experiments, they proposed the binding model for *trp EDCBA* summarized in Figure 6. The guanines blocked from alkylation by dimethyl sulfate, the sequence positions protected from scission by DNase I and OP–Cu footprinting, and the scission sites of Trp repressor E49C–OP are all indicated. The nucleotides at the dyad axis are precisely those attacked with Trp repressor E49C–OP. These diverse experimental methodologies agree on the number and position of the binding sites for Trp repressor. Variation of the Trp repressor E49C–OP concentration in the binding reaction prior to a gel retardation assay has never resulted in more than one detectable protein–DNA complex. Our results are consistent with the isomerization of the protein between isoenergetic binding sites detected by the scission chemistry within the 1:1 complex isolated in the gel retardation assay.[96]

Trp repressor E49C–OP is the first example of an engineered protein capable of precise and *efficient* double-stranded scission of a target sequence. It is currently being used to dissect the circular *E. coli* chromosome which contains four known Trp repressor regulated operons. Four separable fragments with sizes consistent with the genetically and physically mapped loci should be obtained. This will be its first test as a rare cutting restriction enzyme that might be useful in chromosomal mapping. A variety of problems can be envisioned which might block this application. Oxidative chemistry always brings with it the specter of significant background chemistry that will only be apparent when very large molecular weight DNAs are analyzed. Ironically, Trp repressor may be too specific for chromosomal mapping since a recognition sequence 15 bp long may be present only once in 10^9 base pairs. This estimation is probably flawed since Trp repressor recognizes three different operators with related but distinct sequence, and it is possible that these sequences will have evolutionary cousins in other genomes capable of high-affinity binding to the wild-type Trp repressor. Moreover, selection and screening methods can be used to search for alternate DNA binding specificities of a mutagenized Trp repressor.[97,98]

Alternate Method Approaches for Selected Scission of DNA. Within the past two years, a powerful new method of selective and programmable DNA scission has been introduced that relies exclusively on biological reagents. This procedure, initially introduced by Szybalski and colleagues,[99] depends on the selective protection of a restriction enzyme site from methylation by a cognate methylase. Protection can be accomplished by the binding of a repressor,[99] triple-helix formation,[100] or RecA-mediated strand invasion.[101] All unliganded sites of this restriction enzyme are then methylated. When the protecting ligand is removed, its binding site is susceptible to hydrolysis by the cognate restriction endonuclease. The results with this procedure, particularly when the methylation reaction is protected by RecA-mediated strand invasion,[101] shows strong promise as a method for achieving sequence-specific scission of a selected target.

Conclusions

Linking a nucleolytic agent to a carrier ligand is a reliable method for inferring the stereochemistry and specificity of its binding to a nucleic acid. Affinity cleaving has provided insight into the mechanism of triplex formation, the binding of abiological nucleotides, and the interaction of proteins with nucleic acids. To be informative, the scission does not have to be robust. Parallel results have been obtained with either ferrous–EDTA or OP–Cu as the nucleolytic agent.

To be competitive with other methods in the scission of high molecular weight DNA, efficiency as well as specificity of targeted nucleases is desirable. Chimeric proteins are most efficient in the double-stranded scission of DNA if the nucleolytic agent is held rigidly near the DNA and directed toward a unique locus of the DNA-binding site. Engineered restriction enzymes designed like Trp repressor E49C–OP may prove useful in the analysis of high molecular weight DNAs especially if variants with altered recognition can be isolated. The enhancement of biological activity of a carrier by linkage to a chemical nuclease has not yet been achieved.

We thank Dr. C.-h. B. Chen for helpful conversations. Our research was supported by USPHS GM 21199 and ONR N00014-89J-3045.

(95) Kumamoto, A. A.; Miller, W. G.; Gunsalus, R. P. *Genes Dev.* **1987**, *1*, 556–564.
(96) Carey, J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 975–979.
(97) Benson, N.; Sugiono, P.; Bass, S.; Mendelman, L. V.; Youderian, P. *Genetics* **1986**, *114*, 1–14.
(98) Singh, H.; Clerc, R. G.; LeBowitz, J. H. *BioTechniques* **1989**, *7*, 252–261.
(99) Koob, M.; Szybalski, W. *Science* **1990**, *250*, 271–273.
(100) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **1991**, *254*, 1639–1642.
(101) Ferrin, L. J.; Camerini-Otero, R. D. *Science* **1991**, *254*, 1494–1497.