Template-Directed Interference Footprinting of Protein−Adenine Contacts

Changhee Min, Timothy D. Cushing,† and Gregory L. Verdine*

Contribution from the Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138

Received October 31, 1995[⊗]

Abstract: Noncovalent contacts between amino acid residues and DNA bases are the principal contributors to sequence-specific recognition in most protein−DNA complexes. We have developed a method, template-directed interference (TDI) footprinting, which not only identifies DNA bases that are contacted by a protein upon formation of a specific complex but also reveals the groove location of the contacts and provides a rough gauge of their energetics. Previously, we demonstrated TDI footprinting of guanine (TDI-G), cytosine (TDI-C), and thymine (TDI-T) residues in the major groove of DNA. Here we report the development of a procedure for TDI footprinting of the fourth and final DNA base, adenine (TDI-A). The base-analog 7-deaza-7-nitroadenine (A*), present as the corresponding 2′-deoxynucleoside 5′-triphosphate, was found to undergo efficient incorporation into DNA during template-directed enzymatic polymerization. The analog exhibits the same base-pairing preference as its native counterpart (A), and could be selectively degraded, leading to DNA strand scission upon treatment with aqueous piperidine. To validate the use of A* as a probe of specific protein−DNA contacts, we carried out TDI-A footprinting of the bacteriophage *434* repressor/O_R1 operator interaction. The observed interference by A* only at position-1 of the operator is consistent with the results of X-ray crystallographic analysis. Together with TDI-G, -C, and -T footprinting, the present TDI-A footprinting procedure now completes the series of four experiments required for the analysis of major groove contacts to all four DNA bases in solution.

Introduction

Noncovalent contacts between amino acid residues and DNA bases are the principal contributors to sequence-specific recognition in most protein−DNA complexes.[1] Determination of these contact partners and the strength with which they interact is therefore a necessary step toward understanding the regulation of gene expression in the cell. The most widely used biochemical method for determining base contacts in protein−DNA complexes, termed dimethyl sulfate (DMS) footprinting,[2] suffers from the disadvantage that it can only detect major groove contacts to guanine residues and minor groove contacts to adenine. DMS footprinting is not appropriate to contact analysis of C, T, or A in the major groove or to C, T, or G in the minor groove. To overcome these limitations of reagent-based contact analysis, we have developed a design-based approach, termed template-directed interference (TDI) footprinting,[3−6] which not only identifies DNA bases that are contacted by a protein upon formation of a specific complex but also reveals the groove location of the contacts and provides a rough gauge of their energetics.

TDI footprinting employs base analogs having non-native functionality that is designed to interfere with contacts ordinarily made to a protein (Figure 1a). These analogs are incorporated into DNA (via the corresponding 2′-deoxynucleotide 5′-triphosphates, dN*TPs) during enzymatic extension of a ^{32}P-labeled primer annealed to a single-stranded template containing a recognition sequence for the protein of interest (Figure 1b,c).[3] Because the conditions of template-directed polymerization are set such that approximately one analog is incorporated per full-length extension (∼1/200−300 base pairs), the product thus generated is a pool of duplex DNA molecules bearing on average a single analog moiety, the position of which is statistically distributed throughout the ^{32}P-end-labeled strand. The protein of interest is incubated with the DNA pool over a range of protein concentrations, and the mixture is subjected to nondenaturing polyacrylamide gel electrophoresis (PAGE), to size-separate the protein-bound DNA from naked DNA. The PAGE gel thus obtained represents a titration of the entire DNA pool with the protein, in which each member behaves according to its affinity for the protein. Most of the DNA molecules possess the analog in a base position that does not contact the protein directly; these bind the protein with comparable affinities which are also comparable to that of unmodified DNA, and thus partition similarly between the *unbound* and *bound* fractions at the same protein concentration. On the other hand, the very few members of the pool that possess the analog in a critical contact position have reduced affinity for the protein, relative to the majority, and are thus depleted in the *bound* fraction and enriched in the *unbound* fraction; this difference is typically most evident at high protein concentrations (see below). To read out the identity of positions at which the analog interferes with protein binding, DNA is recovered from the bound and unbound fractions and treated with reagents that cause selective strand scission at the position of the analog. Finally, the cleavage products are analyzed on a DNA sequencing gel.

* To whom correspondence should be addressed.

† Present address: Tularik Pharmaceuticals, 270 E. Grand Ave., South San Francisco, CA 94080.

⊗ Abstract published in *Advance ACS Abstracts,* June 1, 1996.

(1) (a) For a review see: Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 205−280. (b) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1992**, *61*, 1053−1095. (c) Larson, C. J.; Verdine, G. L. In *Bioorganic Chemistry: Nucleic Acids*; Hecht, S. M., Ed.; Oxford University Press: New York, 1996; pp 324−346. (d) Harrison, S. C. *Nature* **1991**, *353*, 715−719.

(2) Siebenlist, U.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 122−126.

(3) For a more detailed discussion of the TDI footprinting protocol, see refs 1c and 4.

(4) Hayashibara, K. C.; Verdine, G. L. *J. Am. Chem. Soc.* **1991**, *113*, 5104−5106.

(5) Hayashibara, K. C.; Verdine, G. L. *Biochemistry* **1992**, *31*, 11265−11273.

(6) Mascareñas, J. L.; Hayashibara, K. C.; Verdine, G. L. *J. Am. Chem. Soc.* **1993**, *115*, 373−374.

S0002-7863(95)03658-4 CCC: $12.00

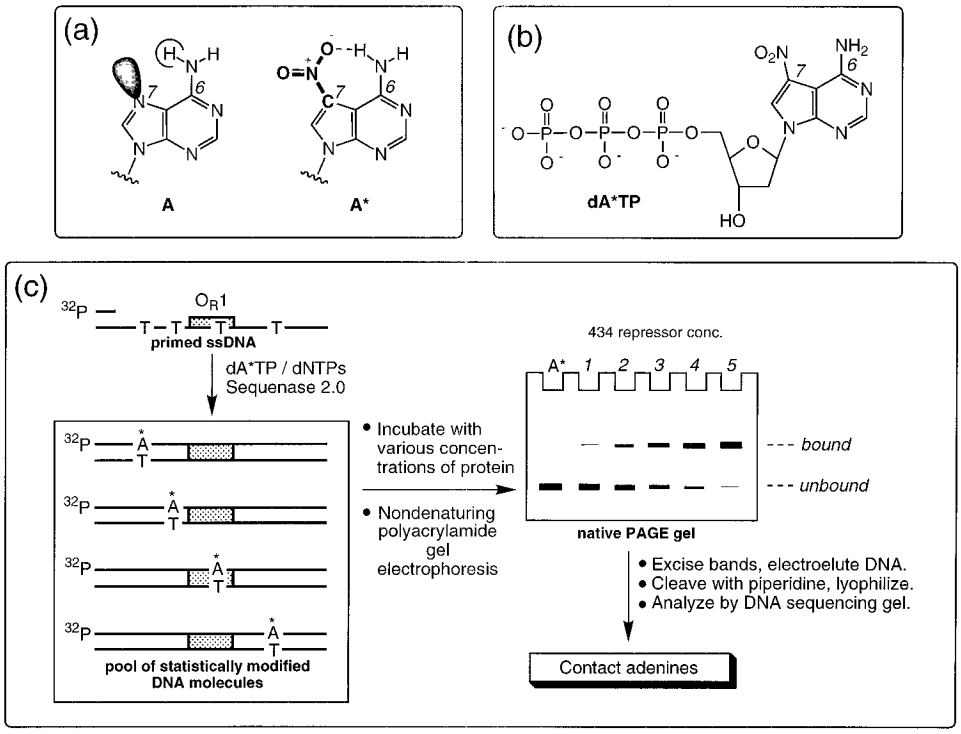

Figure 1. Design strategy, reagents, and procedure employed in TDI-A footprinting. (a) Structural comparison of adenine (A) with 7-deaza-7-nitroadenine (A*). Highlighted are the hydrogen bond acceptor (shaded) and hydrogen bond donor (encircled) contact sites on the major groove surface of A; these sites are blocked by the presence of the C-7-NO_2 group in A*. The C-7-NO_2 group is also intended to activate C-8 toward attack of hydroxide ion, leading to specific strand scission under basic conditions. (b) Structure of dA*TP, the reagent used in the enzymatic incorporation of A* into DNA. (c) Overview of the TDI-A footprinting procedure, which is described in detail in the text.

Analogs suitable for use in TDI footprinting must thus (i) possess structural alterations that alter contact functionality ordinarily employed by the corresponding native base, (ii) be utilized as a substrate by a DNA polymerase, and base-pair as their native counterpart, (iii) undergo specific chemical degradation, leading to DNA strand scission under conditions that leave the rest of the DNA unaffected, and (iv) minimally perturb the duplex DNA structure. Previously, we demonstrated TDI footprinting of guanine (TDI-G),[4] cytosine (TDI-C),[5] and thymine (TDI-T)[6] residues in the major groove of DNA. Here we report the development of a procedure for TDI footprinting of the fourth and final DNA base, adenine (TDI-A footprinting). The present advance now makes it possible to analyze contacts to the base surface of the entire major groove, the principal locus of sequence-specific interactions in protein−DNA complexes.

Materials and Methods

Materials. 7-Deaza-2′-deoxyadenosine (2′-deoxytubercidin)[7] was the generous gift of Boehringer Mannheim. Naturally occurring dNTPs were from Pharmacia (Piscataway, NJ). T*aq* polymerase was from Promega (Milwaukee, WI). T4 polynucleotide kinase was from Gibco BRL (Gaithersburg, MD), and glycogen was from Boehringer Mannheim (Indianapolis, IN). Sequenase 2.0 and the Sequenase 2.0 kit were from United States Biochemical (Cleveland, OH). [γ-^{32}P]ATP was from New England Nuclear. Full-length *434* repressor was expressed in *Escherichia coli* using XA90/pRW190 (gift of G. Koudelka and M. Ptashne) and purified as described.[8]

Synthesis of 4-Amino-7-(2-deoxy-3,5-diacetoxy-β-D-*erythro*-pentofuranosyl)pyrrolo[2,3-*d*]pyrimidine (2). To a stirred suspension of 4-amino-7-(2-deoxy-β-D-*erythro*-pentofuranosyl)pyrrolo[2,3-*d*]pyrimidine (2′-deoxytubercidin, **1**[7]) (164.3 mg, 0.656 mmol, 1.0 equiv) in pyridine (1.6 mL, 0.4 M) at 0 °C under N_2 was added acetic anhydride (0.33 mL, 3.5 mmol, 5.3 equiv). The mixture was allowed to stir overnight, reaching ambient temperature in that time. The solution was diluted with methanol (3 mL), stirred briefly, reduced to dryness, and placed on the vacuum line for a few moments. The residue was then subjected to a refluxing solution of methanol (3 mL) for 12 h. The solution was again reduced to dryness, yielding 240 mg of a crude solid. Chromatography: flash column (20 g of silica) 2% MeOH/CH_2Cl_2 to 5% MeOH/CH_2Cl_2. The product was obtained as a white foamy solid. Yield: 187 mg, 85%. ^{1}H NMR (300 MHz) ($CDCl_3$, relative to TMS) δ 2.09 (6H, s); 2.48 (1H, ddd, J = 2.1, 5.7, 14.0 Hz); 2.59−2.69 (1H, m); 4.27−4.37 (3H, m); 5.31−5.34 (1H, m); 5.61 (2H, br s, D_2O exchanged); 6.41 (1H, d, J = 1.0, 3.7 Hz); 6.68 (1H, dd, J = 5.7, 8.7 Hz); 7.11 (1H, d, J = 3.7 Hz); 8.29 (1H, s). ^{13}C NMR (126 MHz) ($CDCl_3$): δ 20.76, 20.87, 37,44, 64.02, 74.63, 81.54, 83.49, 99.44, 103.89, 120.89, 150.67, 151.98, 156.88, 170.38. IR (thin film, $CDCl_3$): 3152, 1740, 1734, 1684, 1653, 1233 cm^{-1}. MS (EI): *m/e* (rel intens) 334 (M^+, 0.1); 275 (0.5); 134 (100); 107 (0.15). EI HRMS (CI^+) (M + 1): 335.1355 ($C_{15}H_{18}N_4O_5$ requires 334.3328).

Synthesis of 4-Amino-5-nitro-7-(2-deoxy-3,5-diacetoxy-β-D-*erythro*-pentofuranosyl)pyrrolo[2,3-*d*]pyrimidine (3). To a vigorously stirred 0 °C solution of **2** (97.9 mg, 0.29 mmol, 1.0 equiv) in CH_2Cl_2 (1.9 mL, 0.15 M) was added a mixture of concentrated H_2SO_4 (320 mg) and fuming HNO_3 (260 mg) in a dropwise fashion. After 20 min the ice bath was removed and the biphasic solution stirred for an additional 20 min and then neutralized with a saturated $NaHCO_3$ solution until the pH reached 7.0. The mixture was extracted with CH_2Cl_2 (∼200 mL), and the bright yellow solution was washed with brine, dried over $MgSO_4$, and reduced to dryness. Crude yield: 66.3 mg, 60%. The resulting dark yellow solid was used without further purification. ^{1}H NMR (300 MHz) ($CDCl_3$): δ 2.09 (3H, s); 2.14 (3H, s); 2.49−2.58 (1H, m); 2.68 (1H, ddd, J = 3.1, 6.0, 17.2 Hz); 4.34 (3H, br s); 5.31 (1H, m); 6.40 (1H, br s, D_2O exchanged); 6.62 (1H, t, J = 6.4 Hz); 7.56 (1H, br s, D_2O exchanged); 8.26 (1H, s); 8.276 (1H, s). ^{13}C NMR (126 MHz) ($CDCl_3$): δ 20.69, 20.79, 38.82, 63.53, 73.83, 82.62, 84.62, 95.74, 95.78, 125.15, 129.55, 149.56, 154.12, 156.41, 170.23, 170.30.

(7) Cottam, H. B.; Kazimierrczuk, Z.; Geary, S.; McKernan, P. A.; Revankar, G. R.; Robins, R. K. *J. Med. Chem.* **1985**, *28*, 1461−1467.

(8) Anderson, J.; Ptashne, M.; Harrison, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1307−1311.

IR (thin film, $CDCl_3$): 3150, 1740, 1734, 1684, 1653, 1570, 1522, 1506, 1234 cm^{-1}. MS (CI): *m/e* (rel intens) 380 (M + 1, 0.2); 294 (0.60); 236 (100).

Synthesis of 4-Amino-5-nitro-7-(2-deoxy-*β*-D-*erythro*-pentofuranosyl)pyrrolo[2,3-*d*]pyrimidine (4). To a stirred solution of **3** (17.4 mg, 0.046 mmol, 1.0 equiv) under N_2 at 0 °C in methanol (1.3 mL) was added K_2CO_3 (13.9 mg, 0.10 mmol, 2.2 equiv). After 20 min the reaction mixture was diluted with brine (25 mL) and extracted with ethyl acetate (100 mL). The aqueous layer was then re-extracted with ethyl acetate (3 × 50 mL), and the organic fractions were combined, dried over $MgSO_4$, and evaporated to yield a bright yellow solid. Chromatography: PTLC 5% MeOH/CH_2Cl_2. Yield: 8.8 mg, 65%. ^{1}H NMR (300 MHz) (methanol-d_4): δ 2.42−2.48 (1H, m); 2.53−2.60 (1H, m); 3.73−3.87 (2H, m); 4.03 (1H, t, J = 2.9 Hz); 4.53 (1H, dd, J = 1.8, 2.9 Hz); 4.86 (2H, s, D_2O exchanged); 4.87 (2H, s, D_2O exchanged); 6.57 (1H, dd, J = 4.4, 6.2 Hz); 8.17 (1H, d, J = 2.5 Hz); 8.70 (1H, d, J = 2.5 Hz). ^{13}C NMR (126 MHz) (methanol-d_4): δ 41.73, 62.75, 72.14, 86.87, 89.22, 96.88, 128.70, 130.12, 150.12, 154.75, 157.96. IR (thin film, MeOH): 3445−3335 br, 1684, 1653, 1570, 1522, 1507, 1269 cm^{-1}. MS (EI): *m/e* (rel intens) 295 (M^+, 0.1); 200 (0.25); 179 (100). EI HRMS: *m/e* 295.0904 ($C_{11}H_{13}N_5O_5$ requires 295.0917). Mp: 186−190 °C.

Synthesis of 7-deaza-7-nitro-2′-deoxyadenosine 5′-Triphosphate (dA*TP, 5). To a stirred solution of **4** (4.1 mg, 14 μmol, 1.0 equiv) at 0 °C under N_2 in dry trimethyl phosphate (50 μL) was added $POCl_3$ (1.7 μL, 18 μmol, 1.3 equiv). After 1.5 h a solution of tris-(tributylammonium) pyrophosphate (0.14 mL, 1 M, in DMF) was added to the reaction flask, and stirred overnight (0 °C to room temperature). At this time Et_3N (43 μL) and H_2O-d_2 (612 μL) were added to the flask and stirred for 5 min. The mixture was lyophilized and the residue dissolved in 100 μL of H_2O-d_2 and passed through an ion exchange column (Sephadex A25, 1 g) equilibrated in 0.1 M aqueous triethylammonium bicarbonate (TEAB). The column was washed with 50 mL of H_2O-d_2 and then eluted with a 950 mL gradient from 0 to 2 M triethylammonium bicarbonate, with 10 mL fractions being collected. Fractions 11−27 were combined and lyophilized to dryness, and the remaining white powder was redissolved in 1 mL of TE buffer: 10 mM Tris−HCl (pH 8.0), 1 mM EDTA. This solution was used directly in extension reactions with Sequenase 2.0. The yield was estimated to be 5.7 μmol (41%) by spectrophotometric analysis using the reported extinction coefficient (λ_{max} = 365 nm, ϵ = 3900) for the related compound 5-nitrotubercidin.[9] ^{31}P NMR (202 MHz) (D_2O referenced to phosphoric acid): δ −5.4(d); −6.1(m); −19.5(m). UV (H_2O): λ_{max} = 356, 364 nm.

Preparation of Single-Stranded Template DNA by PCR/IMAC.[10] Complementary oligonucleotides containing the operator and flanking sequences used in the crystallographic study of the *434* repressor (N-terminal domain)/O_R1 complex were synthesized with 3′-*Pst*I-compatible ends, ligated into the *Pst*I site in the polylinker of the phagemid pBS+ (Stratagene), and transformed into *E. coli* strain XL1-blue (Stratagene). Insert-positive recombinants (designated pBS-434) were verified by sequencing of single-stranded phagemid DNA isolated from helper phage VCS-M13 (Stratagene).[11]

The polymerase chain reaction followed by immobilized metal affinity chromatography (PCR/IMAC)[10] was used to generate single-stranded DNA templates containing the operator sequence. Amplification was carried out in a 100 μL PCR reaction mixture containing 50 mM KCl, 10 mM Tris−HCl (pH 8.4), 1.5 mM $MgCl_2$, 0.25 mM (each) dATP, dCTP, dGTP, and TTP, 1 unit of T*aq* polymerase, 20 μg of pBS-434 as template, 50 pmol of H_6-tagged M13 (-48) primer [5′-d(GHHHHHHAGCGGATAACAATTTCACACAGG)-3′],[12] and 50 pmol of M13 (−47) primer [5′-dCGCCAGGGTTTTCCCAGTCAC-GA)-3′]. The reaction solution was overlaid with a 60 μL mineral oil cap to prevent evaporation and placed in a thermal cycler (MJ Research) programmed for 35 cycles set at 94 °C (1 min), 55 °C (1 min), and 72 °C (1 min).

The duplex PCR products were resolved by IMAC as follows: 1 mL of TE buffer was added to the crude reaction mixture, which was then concentrated to a total volume of 50 μL using a centrifugal dialysis cartridge (Centricon 30, Amicon). The DNA solution was transferred to an Eppendorf tube, to which was added 150 μL of binding buffer (6 M guanidine·HCl, 10 mM Tris, pH 8.2). This mixture was heated for 5 min at 90 °C. Separately, to a 1.5 mL Eppendorf tube was added 250 μL (bed volume) of Ni^{2+}-NTA-agarose resin and 1 mL of binding buffer at room temperature. The hot DNA solution was added to the suspension of resin, and the resulting mixture was mixed for 1−1.5 min by vigorous shaking. The mixture was transferred to an empty 5 mL fritted column (Qiagen), and the flow-through fraction, which contained the unmodified strand, was collected into an Eppendorf tube. After repeated pipetting to ensure complete mixing, the unbound fractions were aliquotted into four Eppendorf tubes (300 μL each) and held aside for further processing. Next, the resin was washed with 1 mL of washing buffer (10 mM Tris, 5 mM imidazole, pH 8.0), which is discarded. The H_6-containing strand was then eluted in 1.2 mL of 200 mM aqueous imidazole solution. The imidazole eluate was mixed thoroughly and aliquotted into four Eppendorf tubes (300 μL each), to each of which were added 30 μL of 100 mM ethanolic 1,10-phenanthroline and 30 μL of 3 M aqueous NaOAc. To each of the unbound and bound fractions was added 900 μL of absolute ethanol (stored at −20 °C), and then the tubes were vortexed briefly and chilled for 30 min on powdered solid CO_2. The tubes were microcentrifuged for 30 min at 16000*g*. The supernatant was removed and the pellet washed with 200 μL of 80% aqueous EtOH (−20 °C). Following removal of the ethanol solution, the tubes were evaporated to dryness by centrifugal lyophilization (SpeedVac, Savant). To each dry tube was added 50 μL of TE buffer, and the DNA concentration was quantified by UV spectrophotometry. This procedure typically yields 5−10 pmol of the H_6-tagged template and the unmodified template (the primers are used in ∼5−10-fold excess in the PCR reaction).

Template-Directed Extension. This procedure provides a sufficient template for 20 lanes on a DNA sequencing gel.

(a) End Labeling of Primer. A 20 μL volume of 0.2 pmol/μL M13 (−21) primer, 5'-d(GTAAAACGACGGCCAGT)-3', was end-labeled in a 100 μL volume by adding 10 μL of 400 mM Tris−HCl (pH 7), 200 mM $MgCl_2$, 20 units of T4 polynucleotide kinase, and 4 μL of 6000 Ci/mmol [γ-^{32}P]ATP. This mixture was incubated at 37 °C for 30 min, heated at 90 °C for 3 min, chilled on ice, and briefly microcentrifuged at 14000 rpm.

(b) Annealing. A 100 μL volume of the DNA template strand (approximately 0.04 pmol/μL in TE) was added to the end-labeled primer (100 μL) and heated to 65 °C for 5 min, and then allowed to reach room temperature while remaining in the bath. The condensate was recombined with the annealed primer/template solution by brief microcentrifugation.

(c) Primer Extension. A 200 sample μL of annealed primer/template (4 pmol) was combined with 70 μL of polymerization buffer [48 mM Tris−HCl (pH 7), 24 mM $MgCl_2$, 200 mM NaCl, 9 mM DTT, and 400 μg/μL BSA) and 40 μL of diluted Sequenase 2.0 (1.3 units/μL in a buffer containing 10 mM Tris−HCl (pH 7.5), 5 mM DTT, and 0.5 mg/mL BSA). The resulting solution was added to 220 μL of a mixture containing 80 μM each of dCTP, dGTP, and dTTP; an 80 μM (total nucleotide) 1:3 mixture of dA*TP and dATP; and 50 μM NaCl. The extension was allowed to proceed at 37 °C for 20 min and then stopped by addition of 220 μL of 3 M sodium acetate (pH 7.0), and precipitated at −70 °C with 20 μL of 20 mg/mL glycogen and 1800 μL of 100% ethanol. The pellet obtained after microcentrifugation at 14000 rpm for 30 min at 4 °C was washed with 80% aqueous ethanol and dried by lyophilization. The resulting DNA was dissolved in 80 μL of TE, 4 μL of which was then used for each subsequent incubation with protein.

Interference Binding Assay. The extended template DNA (4 μL or 0.2 pmol per incubation) was incubated for 30 min at room temperature with varying amounts of *434* repressor in 20 μL of buffer consisting of 50 mM KCl, 10 mM Tris−HCl (pH 7.8), 1 mM $MgCl_2$,

(9) Watanabe, S.-I; Ueda, T. *Nucleosides Nucleotides* **1983**, *2*, 113−125. Watanabe, S.-I; Ueda, T. *Nucleosides Nucleotides* **1982**, *1*, 191−203.

(10) Min, C.; Verdine, G. L. Submitted for publication. A long-standing problem in TDI footprinting, like other methods that require long, single-stranded DNA molecules, has been the inability to produce the template DNA consistently and in sufficient quantities for multiple experiments. This problem has been solved conclusively by the use of the PCR/IMAC procedure.

(11) Hayashibara, K. C. Ph.D. Thesis, Harvard University, Cambridge, MA, 1993.

(12) H = 6-histaminylpurine.

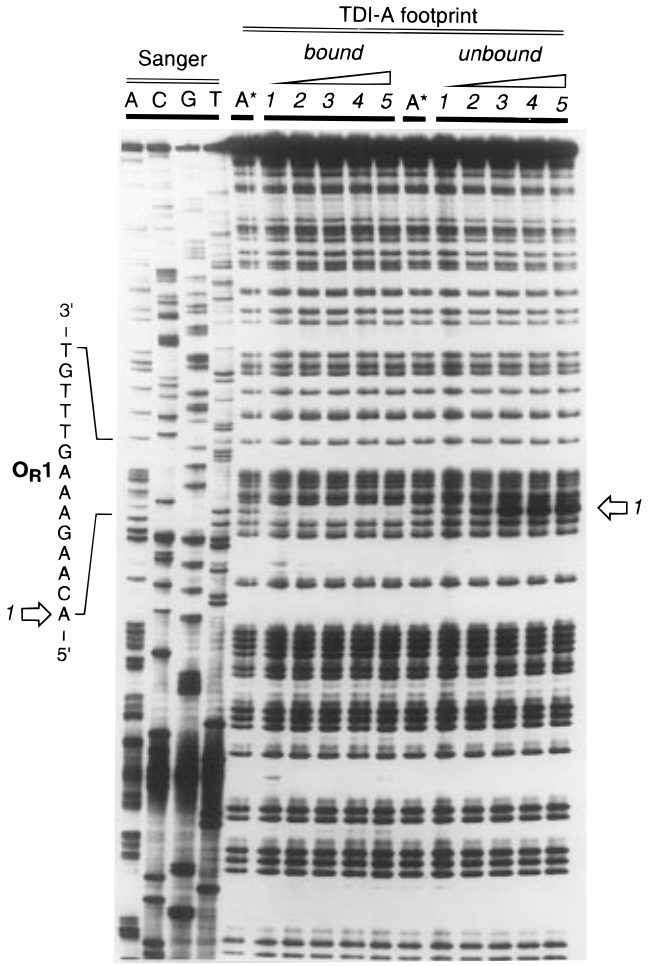

Figure 2. TDI-A footprinting of the *434* repressor/O_R1 operator complex. The four lanes at the far left (A, C, G, and T) are an authentic Sanger (dideoxy) sequence of the template, with the O_R1 sequence shown alongside. The remainder of the lanes shown represent data for a TDI-A footprint using A*-containing DNA. The two lanes denoted A* are controls in which the entire pool of A*-containing DNA was cleaved, without having been incubated with *434* repressor. Lanes *1*−*5* represent the respective *bound* and *unbound* fractions from experiments in which the pool of A*-containing DNA ($\sim 2.5 \times 10^{-9}$ M) was incubated with the following concentrations of *434* repressor (lane 1, 1×10^{-8} M; lane 2, 3×10^{-8} M; lane 3, 5×10^{-8} M; lane 4, 7×10^{-8} M; lane 5, 1×10^{-7} M). The arrow denotes the A1 position, the only one at which strong interference is observed, consistent with the results of X-ray structural analysis. The bands generated by piperidine cleavage of A*-containing DNA migrate slightly faster than the corresponding bands in the Sanger dideoxy-A sequencing lane, because the former DNA strands possess one fewer nucleoside unit at the 3′-end.

1 mM $CaCl_2$, 100 μg/mL BSA, and 10% (v/v) glycerol. Protein-bound and unbound DNA were separated using a 5% low-ionic-strength nondenaturing polyacrylamide gel[13] (37:1 monomer/methylenebisacrylamide). The gel was exposed to BioMAX film (Kodak) for 5 h. Bands were visualized by autoradiography (Figure 2) and excised from the gel using a razor blade. The gel slices were electroeluted into 130 μL of 10 M ammonium acetate using a V-channel apparatus (Harvard BioLabs) with 0.5× TBE [45 mM Tris−borate (pH 8.0), 0.5 mM EDTA] as running buffer. The DNA samples in 10 M ammonium acetate were precipitated directly by the addition of 3 volumes of 100% ethanol followed by precipitation as described above.

Cleavage and Sequencing Procedure. Each ^{32}P-labeled dsDNA sample was redissolved in 50 μL of 1 M aqueous piperidine. This solution was heated at 90 °C for 30 min, cooled to room temperature, and microcentrifuged briefly. Samples were lyophilized, twice resuspended in 30 μL of H_2O-d_2, and lyophilized, and then redissolved in 3 μL of TE and 2 μL of loading dye containing 95% formamide, 10 mM EDTA, 0.025% bromophenol blue, and 0.025% xylene cyanol. Samples were heated at 90 °C for 2 min and then loaded immediately onto a 7% polyacrylamide sequencing gel. The gel was visualized by autoradiography. The generation of an authentic sequence by Sanger sequencing made use of the Sequenase 2.0 kit (United States Biochemicals) according to the manufacturer's protocol. A volume of 2.5 μL of annealed [^{32}P]primer/template mixture was used in each dideoxy termination reaction.

Results and Discussion

Our aim in the present work was to design, synthesize, and validate an analog suitable for TDI footprinting of major groove contacts to adenine (TDI-A footprinting). The major groove contact surface of adenine consists of the aza N at position-7 and the N−H at position-6 (Figure 1a), both of which are widely involved in hydrogen-bonding contacts to proteins.[1] Taking the above criteria into consideration, we designed the adenine analog 7-deaza-7-nitroadenine (A*), in which the C-7-NO_2 function is envisioned to disrupt contacts not only at position-7 through steric exclusion but also at the nearby 6-NH through intramolecular hydrogen bonding (Figure 1a).[14] Furthermore, the strongly-electron-withdrawing 7-NO_2 group is expected to promote conjugate addition of hydroxide ion at C-8, ultimately leading to ring fragmentation and DNA strand scission.[15] To test these concepts, 7-deaza-7-nitro-2′-deoxyadenine (dA*) was synthesized by nitration of 2′-deoxytubercidin. A sample of dA* was dissolved in 1 M aqueous piperidine at 90 °C, and aliquots were periodically removed and spotted onto a TLC plate. Analysis of the resulting TLC plate (not shown) revealed virtually complete disappearance of the dA* chromophore within 30 min. These data indicated that the base moiety of dA* undergoes hydrolytic degradation under the conditions used in piperidine-catalyzed cleavage of DNA, suggesting it might be possible to cleave DNA selectively at positions containing the analog.

To assay for the base-pairing specificity of dA* and its selective cleavage in DNA, the free nucleoside was converted to the corresponding 5′-triphosphate (dA*TP), which was employed, together with the four natural 2′-deoxynucleoside 5′-triphosphates (dNTPs), in primer-extension and footprinting assays. The single-stranded template DNA used in these reactions was generated by a recently developed procedure that employs the polymerase chain reaction followed by immobilized metal affinity chromatography (PCR/IMAC).[10] A ^{32}P-radiolabeled primer was annealed to this template, and extension was initiated by the addition of modified T7 DNA polymerase (Sequenase 2.0), together with the four dNTPs and dA*TP in an amount sufficient to afford incorporation of approximately one dA* unit per full-length extension.[16] The extended DNA sample was then treated with 1 M aqueous piperidine. High-resolution electrophoretic analysis of the products arising from extension and cleavage is shown in the A* lanes of Figure 2. The pattern of bands observed in the A* lanes matches precisely that in the Sanger dideoxy-A sequencing lane, thus indicating that the analog shows the same base-pairing preference as its natural counterpart.[17]

(13) Chodish, L. A. In *Current Protocols in Molecular Biology*; Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., Struhl, K., Eds.; John Wiley and Sons: New York, 1989; Vol. 2, pp 12.2.1−12.2.10.

(14) The existence of this intramolecular hydrogen bond is consistent with the observation of distinct signals for each of the 6-amino protons in the ^{1}H NMR spectrum of dA*, whereas only one signal is observed for the 6-NH_2 protons of 7-deaza-dA.

(15) In nucleoside model studies, it was determined that treatment of dA* with 1 M aqueous piperidine for 5 min at 90 °C results in complete loss of the characteristic 364 nm 7-deaza-7-nitroadenine chromophore, consistent with base-catalyzed ring-opening.

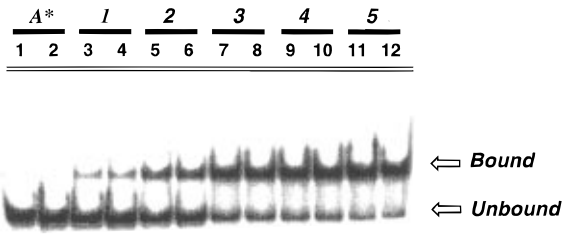

Figure 3. Nondenaturing polyacrylamide gel electrophoresis of mixtures containing a pool of A*-substituted DNA and various concentrations of *434* repressor. Each set of lanes (*e.g.*, 1/2, 3/4, *etc.*) represents a duplicate run. The italicized captions correspond to those in Figure 2. Concentrations of *434* repressor: (lane 1, 1×10^{-8} M; lane 2, 3×10^{-8} M; lane 3, 5×10^{-8} M; lane 4, 7×10^{-8} M; lane 5, 1×10^{-7} M). Concentration of DNA: 2.5×10^{-9} M.

To examine the ability of A* to interfere with specific recognition by a DNA-binding protein that is known to utilize A contacts, we chose the bacteriophage *434* repressor/O_R1 operator system, the structure of which has been determined at high resolution by X-ray crystallography.[18] The pool of A*-containing DNA, which contains a single, centrally located O_R1 sequence (Figure 2, refer to left-hand margin), was incubated with increasing concentrations of *434* repressor. The titration series was fractionated by nondenaturing polyacrylamide gel electrophoresis (PAGE), which separates naked DNA from protein-bound DNA on the basis of differences in size (Figure 3). The *bound* and *unbound* DNA bands were extracted form the native PAGE gel, cleaved with piperidine, and analyzed on a high-resolution denaturing polyacrylamide gel (Figure 2, lanes marked TDI-A footprint). Interference is evident in a TDI-A footprint as positions in the sequence at which a band is weak or absent from the *bound* lanes and enriched in the *unbound* lanes, relative to the A* control. Such behavior is clearly observed at one and only one position in the sequence: the 5′-A residue of O_R1 (position-A1; Figure 2, arrow).[19] A band at A1 is virtually absent from the *bound* lanes at all protein concentrations, thus indicating that binding by *434* repressor is strongly and adversely affected by the analog at this site. Conversely, the relative intensity of the A1 band in the *unbound* lanes increases as the protein concentration is increased, again suggesting that DNA molecules having the analog at position-1 bind *434* repressor poorly.

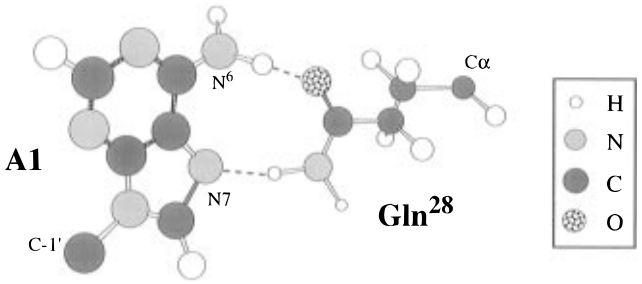

Figure 4. Ball-and-stick representation of the contact between position-A1 of the *434*/O_R1 operator site and the amino acid residue Gln^{28} of the *434* repressor.[18] The contact appears to involve A bidentate hydrogen bonding interaction between the side chain carboxamide function of Gln^{28} and the H bond donor and acceptor atoms of A1, N^6-H and N7, respectively. Both of these H bonding interactions are disturbed by substitution of N7 with $C-NO_2$ in the TDI-A footprinting analog A*. Interference is also observed at the symmetry-related position-A14 in O_R1, which makes an identical contact to Gln^{28} of the other *434* protomer in the DNA-bound protein dimer.

In two important respects, this result compares favorably with expectation on the basis of the crystal structure of the *434* repressor/O_R1 complex:[18] (i) a hydrogen-bonding contact is observed between A1 and the side chain of Gln^{28} (Figure 4) and (ii) with the exception of A14, the symmetry-related counterpart to A1, no other adenine contact is observed in this complex. Finally, the lack of interference at A positions that are base-paired to contacted thymines (A3 and A4) provides evidence that A* has little if any effect on the duplex DNA structure. We thus conclude that A* causes interference through localized disruption of specific A contacts, rather than through perturbation of the duplex DNA structure.

Conclusion

Here we have demonstrated the extension of the TDI footprinting technique to include analysis of major groove contacts to adenine residues. Together with TDI-G,[4] -C,[5] and -T[6] footprinting, the present TDI-A footprinting procedure now completes the series of four experiments required for the analysis of major groove contacts to all four DNA bases in solution. When interfaced with the PCR/IMAC procedure to facilitate template generation,[10] TDI footprinting represents a rapid and powerful procedure for studying specific recognition in protein−DNA complexes.

Acknowledgment. This work was supported by grants (to G.L.V.) from the National Science Foundation (Presidential Young Investigators Program), Searle Scholar Program, and the Sloan Foundation. We are grateful to Dr. Werner Wolf of Boehringer Mannheim for generously providing 2′-deoxytubercidin, which was used as a synthetic precursor in these studies, to Qing Fan, Dr. Chris Larson, and members of the Verdine group for experimental advice, and to Orlando Schärer and Huw Nash for a critical reading of the paper.

JA953658Z

(16) In a separate series of primer-extension/cleavage assays (not shown), the ratio of dA*TP to dATP used in the enzymatic extension was varied over a wide range. The extended samples were divided into two parts, one of which was treated with 1 M aqueous piperidine. The piperidine-treated and untreated samples were analyzed on a high-resolution DNA sequencing gel (not shown). By comparison of the two samples at each dA*TP/dATP ratio, it was determined that (i) dA*TP is fully capable of supporting enzymatic polymerization, even under conditions in which it completely replaces dATP, and (ii) DNA strand cleavage at dA* positions is almost quantitative in the presence of 1 M aqueous piperidine. The dA*TP:dATP ratio (1:3) that gives rise to roughly one dA* incorporation per extension was determined empirically by inspection of this gel.

(17) This series of bands is not present in samples that have been extended under the same conditions but not subjected to piperidine treatment. The bands generated by piperidine cleavage of A*-containing DNA migrate slightly faster than the corresponding bands in the Sanger dideoxy-A sequencing lane, because the former DNA strands possess one fewer nucleoside unit at the 3′-end.

(18) Aggarwal, A. K.; Rodgers, D. W.; Drottar, M.; Ptashne, M.; Harrison, S. C. *Science* **1988**, *242*, 899−907.

(19) *434* repressor binds DNA as a dimer, making symmetry-related contacts to the two pseudosymmetric half-sites of O_R1.[18] In TDI footprints of the complementary strand to that shown in Figure 2, interference is observed only at the one position−A14−which is related by symmetry to A1 (C. Min and G. L. Verdine, unpublished results).