Template-Directed Interference Footprinting of Protein–Thymine Contacts

Jose L. Mascareñas,† Kathleen C. Hayashibara,‡ and Gregory L. Verdine*

Department of Chemistry, Harvard University Cambridge, Massachusetts 02138

Received September 30, 1992

Understanding the structural basis of transcriptional control has emerged as a principal goal of life science; at the root of this problem lies the issue of sequence discrimination in DNA-binding proteins.[1-3] Toward realization of this goal we have developed a method, termed template-directed interference (TDI) footprinting, for determining DNA bases contacted by a protein upon formation of a specific protein–DNA complex.[4,5] In TDI footprinting, analogs of the naturally occurring DNA bases are incorporated into DNA during enzymatic polymerization and then tested for site-specific interference of binding by a protein. We have previously reported TDI footprinting of contact guanine[4] and cytosine[5] residues; in this report we now demonstrate TDI footprinting of contact thymines.

For TDI footprinting of contact thymines (TDI-T footprinting), we designed[6] the analog 5-hydroxy-2′-deoxyuridine (h^5dU, Figure 1) as follows. The overwhelming majority of known regulatory proteins binds DNA in the major groove;[1-3] hence, the major groove surface of thymine was selected for modification. X-ray crystallographic studies have indicated that proteins recognize thymine by hydrophobic interactions with the C5-methyl group or hydrogen bonding to the carbonyl oxygen (O^4).[1-3] h^5dU is designed to interfere with either kind of contact by replacing the nonpolar C5-methyl with a polar hydroxyl group and by providing for competitive intramolecular hydrogen bonding to O^4 (Figure 1). Furthermore, we reasoned that the presence of an electron-donating group at C5 would activate the nucleoside toward electrophilic oxidation, which in DNA would then allow for selective strand cleavage at positions occupied by the analog.[6]

The substrate required for enzymatic incorporation of h^5dU into DNA, 5-hydroxy-2′-deoxyuridine 5′-triphosphate (h^5dUTP), was synthesized and found to replace thymidine triphosphate (TTP) in DNA synthesis reactions using several polymerases.[7,8] Moreover, duplex DNA was found to undergo selective cleavage at h^5dU sites upon treatment with dilute permanganate followed by aqueous piperidine; control DNA containing T and no h^5dU was completely unaffected.[9] Using this cleavage assay, a 1:3 mol ratio of TTP to h^5dUTP was found to direct incorporation of approximately one h^5dU moiety per extension. Under these

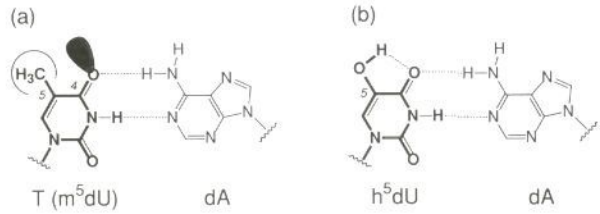

Figure 1. Base pair formed between dA and T (a) or h^5dU (b). Highlighted in a are the hydrogen-bond acceptor (shaded) and hydrophobic (encircled) contact sites on the major groove surface of T; both of these are altered in h^5dU.

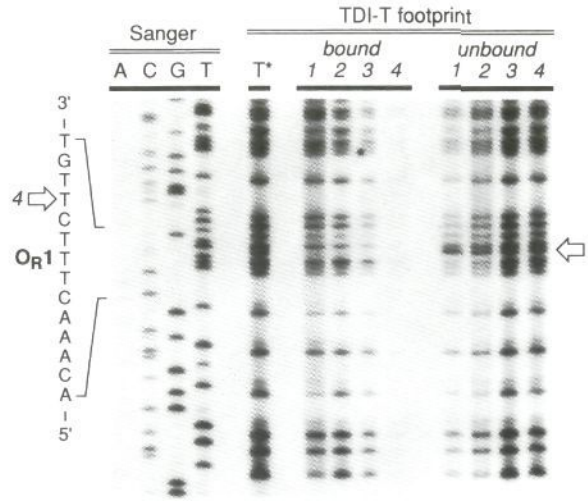

Figure 2. TDI-T footprinting of the 434 operator upper strand. A, C, G, and T lanes are an authentic Sanger (dideoxy) sequence (radioactivity in the A lane was weak). The T* is a control in which the entire pool of h^5dU-containing DNA was cleaved. The protein concentrations used in binding experiments to generate bound and unbound fractions were as follows: lane 1, 1×10^{-7}; lane 2, 5×10^{-8}; lane 3, 3×10^{-8}; lane 4, 1×10^{-8} M. The DNA concentration was $\sim 10^{-8}$ M. Arrows denote the T4 position, at which strong interference is observed.

Figure 3. View of a three base pair section of the 434 repressor–O_R1 complex.[10] In the DNA (blue), the A3–T3 base pair is at the top, the A4–T4 is in the middle, and G5–C5 is at the bottom. Shown in red is an α-carbon tracing of the recognition helix of the 434 repressor, residues 36 (front) to 28 (back); Gln residues are in yellow: Gln-33 (front) and Gln-29 (back). The dashed line indicates a hydrogen bond between the O^4 of T4 and the NH_2 of Gln-33; replacement of the methyl group of T4 (immediately underneath the dashed line) with a hydroxyl group disrupts this interaction. Similar replacement of the methyl group of T3 (immediately above the dashed line) causes little or no interference of a putative hydrophobic contact to the β-CH_2 of Gln-29.

* Author to whom correspondence should be addressed.
† Current address: Departamento de Quimica Organica, Universidad de Santiago, 15706 Santiago de Compostela, Spain.
‡ Current address: Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

(1) Harrison, S. C. *Nature* **1991**, *353*, 715.
(2) Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 205.
(3) Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1992**, *61*, 1053.
(4) Hayashibara, K. C.; Verdine, G. L. *J. Am. Chem. Soc.* **1991**, *113*, 5104.
(5) Hayashibara, K. C.; Verdine, G. L. *Biochemistry* **1992**, *31*, 11265.
(6) TDI footprinting analogs must fulfill four requirements. They must (i) possess the same base-pairing characteristics as their natural counterpart; (ii) minimally perturb DNA secondary structure; (iii) possess some change in functionality that disrupts a specific amino acid contact; and (iv) be capable of chemical modification leading to DNA strand scission under conditions that do not lead to cleavage of normal DNA.
(7) Roy-Burman, S.; Roy-Burman, P.; Visser, D. W. *Biochem. Pharmacol.* **1970**, *19*, 2745. A modified version of this procedure was used in the present study.[8]
(8) Detailed experimental procedures are deposited in the supplementary material, which can be obtained directly by FAX (617) 495-8755.
(9) Permanganate is known to cleave thymines in single-stranded DNA while leaving those in double-stranded DNA unaffected: McCarthy, J. G. *Nucleic Acids Res.* **1990**, *18*, 7541. We estimate that, in double-stranded DNA at room temperature, h^5dU is at least 10^4 more reactive toward permanganate than T.

0002-7863/93/1515-373$04.00/0

conditions the product of enzymatic extension is a pool of singly-h^{5}dU-substituted DNA molecules, in which the sequence position of the analog varies throughout the pool.

As a test for specific interference by h^{5}dU, we chose the 434 repressor–O_R1 operator complex, for which the structure has been determined at high resolution by X-ray crystallography.[10] h^{5}dU-substituted DNA containing the 434 O_R1 sequence was incubated with varying concentrations of 434 repressor, the protein-bound DNA was separated from unbound DNA, and the DNA was recovered, cleaved, and then analyzed on a DNA sequencing gel. On the left of this gel (Figure 2) is an authentic dideoxy sequence, with the region corresponding to O_R1 denoted. The T* lane is a control in which the entire pool of h^{5}dU-substituted DNA was cleaved; its 1:1 correspondence with the dideoxy-T lane reveals that the analog replaces only T residues.[11] In the TDI-T footprint, interference is evident as a T position at which a band is depleted or missing in the *bound* lanes and enriched relative to others in the *unbound* lanes, especially at high protein concentration (increasing from lane 4 → 1). Such behavior is clearly evident at one position—T4—in O_R1, at which a band is absent from all four *bound* lanes, and the relative intensity is substantially enriched in lanes 1 and 2 of the *unbound* lanes. This result is consistent with crystallographic evidence that T4 makes a hydrogen-bonding contact with Gln-33 of the 434 repressor (Figure 3).

A hydrophobic contact between T3 and Gln-29 has also been inferred from the X-ray structure (Figure 3).[10] However, h^{5}dU at this position interferes weakly at best, suggesting that either (i) h^{5}dU does not interfere strongly with hydrophobic contacts or (ii) the hydrophobic contact between Gln-29 and T3 contributes weakly to the overall stability of the protein–DNA complex. We favor the latter explanation for the following reasons: (i) 2′-deoxyuridine—which lacks the hydrophobic methyl group of T—does not interfere detectably at T3,[12] and more importantly, (ii) in the Jun–Fos complex, we have observed potent interference by h^{5}dU at thymines suggested by crystallographic analysis[13] to make hydrophobic contacts.[12,14]

Supplementary Material Available: Detailed experimental procedures for synthesis of h^{5}dU and h^{5}dUTP, UV and ^{1}H and ^{31}P NMR spectra, and details of the procedure for TDI-T footprinting (14 pages). Ordering information is given on any current masthead page.

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

(11) The bands in the Sanger T lane migrate more slowly than the corresponding bands in the TDI-T lanes, because the latter possess one fewer nucleoside unit.

(12) Hayashibara, K. C.; Mascareñas, J. L.; Verdine, G. L. Unpublished results.

(13) A crystal structure of the related protein GCN4 complexed with DNA has been obtained. Ellenberger, T. E.; Brandl, C.; Struhl, K.; Harrison, S. C. *Cell* **1992**, in press.

(14) This research was supported by grants (to G.L.V.) from the Institute of Chemistry in Medicine, the Chicago Community Trust (Searle Scholars Program), and the National Science Foundation (Presidential Young Investigators Program). G.L.V. is a Sloan Fellow, a Lilly grantee, and a Dreyfus Teacher–Scholar. J.L.M. was supported by a visiting professorship from the Universidad de Santiago. K.C.H. was partially supported by a predoctoral fellowship from the National Science Foundation. We thank Ronen Marmorstein for assistance with graphics and Stephen Harrison and Tom Ellemberger for providing X-ray coordinates of the GCN4–DNA complex prior to publication.