Structural flexibility in transcription complex formation revealed by protein–DNA photocrosslinking

Michele A. Cleary*†, P. Shannon Pendergast*‡, and Winship Herr*§

*Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724; and ‡Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724

Communicated by James D. Watson, Cold Spring Harbor Laboratory, Long Island, NY, May 29, 1997 (received for review April 4, 1997)

ABSTRACT The Oct-1 POU domain binds diverse DNA-sequence elements and forms a higher-order regulatory complex with the herpes simplex virus coregulator VP16. The POU domain contains two separate DNA-binding domains joined by a flexible linker. By protein–DNA photocrosslinking we show that the relative positioning of the two POU DNA-binding domains on DNA varies depending on the nature of the DNA target. On a single VP16-responsive element, the POU domain adopts multiple conformations. To determine the structure of the Oct-1 POU domain in a multiprotein complex with VP16, we allowed VP16 to interact with previously crosslinked POU-domain–DNA complexes and found that VP16 can associate with multiple POU-domain conformations. These results reveal the dynamic potential of a DNA-binding domain in directing transcriptional regulatory complex formation.

The human transcriptional activator protein Oct-1, which contains a DNA-binding POU domain, displays diversity in DNA-sequence recognition and coregulator association. The POU domain is a bipartite DNA-binding structure made up of two helix–turn–helix-containing DNA-binding domains: a POU-specific (POUS) domain joined by a flexible linker to a carboxy-terminal POU-homeo (POUH) domain (1–3). The Oct-1 POU domain displays exceptional sequence recognition flexibility and covalent formation with viral and cellular coregulators, including the herpes simplex virus transactivator VP16 (Vmw65, TIF) (reviewed in ref. 4). DNA-binding studies of POU domains have revealed a high degree of structural flexibility in their interaction with DNA. For example, the x-ray crystal structures of the Oct-1 and Pit-1 POU domains bound to Oct-1- and Pit-1-responsive sites, respectively, differ: The Oct-1 POU domain binds to the octamer sequence ATGCAAAAT as a monomer, with the POUS and POUH domains bound to opposite faces of the DNA (2), whereas the Pit-1 POU domain binds its site as a dimer with the POUS and POUH domains of each POU domain bound to the same face of the DNA and with the POUH domain in an inverted orientation to that found with Oct-1 (3). Furthermore, combined mutagenesis of the Oct-1 or related Brn-2 POU domains and their respective DNA targets has provided indirect evidence that even the same POU domain adopts different conformations on different cis-regulatory sites, providing a mechanism for divergent sequence recognition (5, 6).

To probe directly the conformational flexibility of the Oct-1 POU domain bound to DNA and in a higher-order complex, we adapted a site-specific protein–DNA photocrosslinking protocol (7) to map, individually, the locations of the POUS and POUH domains on DNA and in complex with VP16. Our results show that the Oct-1 POU domain adopts multiple conformations on a single binding site and in association with VP16.

MATERIALS AND METHODS

In Vitro Mutagenesis. The site-specific protein–DNA photocrosslinking method we used (7) involves the attachment of the photoactivatable crosslinker 4-azidophenacyl bromide to an appropriately positioned unique cysteine in the DNA-binding protein. We removed the existing Oct-1 POU domain cysteines at positions 61 of the POUS domain and 50 of the POUH domain and introduced unique cysteines by oligonucleotide-mediated site-directed mutagenesis (6). To create unique cysteine (UC) derivatives, we made the following POUS mutations: L23C (POUSUC23), G24C (POUSUC24), F25C (POUSUC25), G28C (POUSUC28), D29C (POUSUC29), and N54C (POUSUC54); and the following POUH mutations: S7C (POUHUC7), E9C (POUHUC9), and T10C (POUHUC10).

4-Azidophenacyl-Bromide Modification. Oct-1 POU-domain variants were prepared in Escherichia coli as glutathione S-transferase (GST)-fusion proteins essentially as described (8). Purified POUS domain was cleaved from the GST moiety (3.5 μM in 500 μl) with thrombin and treated for 6 hr at room temperature in the dark with 140 μM 4-azidophenacyl bromide (Sigma) and 1.7% dimethylformamide as described (7). Reaction mixtures were then dialyzed twice against 1 liter of 20 mM Tris-HCl, pH 8.0/200 mM KCl/5% glycerol/0.1 mM EDTA for a total of 14 hr.

POU Domain–DNA Photocrosslinking and Site-Specific Cleavage. POUS domain–DNA binding reactions (100 μl in flat-bottomed microtiter plate wells) contained 0.28 μM Oct-1 POU domain protein, 3 × 105 cpm of the appropriate singly end-labeled DNA probe generated by PCR amplification (see ref. 6) in 8 mM NaHepes, pH 7.9/60 mM NaCl/4 mM spermidine/2 mM EDTA/0.2 mM DTT/0.03% Nonidet P-40/0.1 mg/ml bovine serum albumin. Reaction mixtures were incubated at room temperature for 30 min in the dark before being irradiated for 1.5 min in a UV Strataligner 2400 (Stratagene) with 312-nm UV bulbs. The irradiation time was determined empirically for optimal crosslinking. After UV irradiation, 5 μl of each binding reaction mix was removed, added to 5 μl of SDS loading buffer, heated to 90°C for 10 min, and subjected to 10% SDS/PAGE in the absence of a stacking gel. Levels of crosslinking (typically 3%) were determined by phosphorimager analysis (Fuji BAS1000). The remainder of the sample was used to map the sites of crosslinking as described (7). Purine cleavage patterns for each probe were generated as described (9). DNA cleavage products were analyzed on 20% polyacrylamide/7 M urea denaturing gels.

Abbreviation: UC, unique cysteine.

†Present address: Howard Hughes Medical Institute and Department of Molecular Biology, Princeton University, Princeton, NJ 08544.
§To whom reprint requests should be addressed. e-mail: herr@csbl.org.
**RESULTS**

Experimental Design. Pendergrast et al. (7) have described a site-specific protein–DNA photocrosslinking method in which the photactivatable crosslinking agent 4-azidophenacyl bromide is covalently attached to a unique cysteine side chain on the surface of a DNA-binding protein. The modified protein is allowed to bind to DNA, and the binding reaction is then subjected to UV irradiation, enabling the modified cysteine side chain to form a covalent bond with nearby DNA. Crosslinked protein–DNA complexes are then analyzed directly by SDS/PAGE or the sites of crosslinking mapped by site-specific crosslink-induced DNA cleavage (7). We created Oct-1 POU domains carrying either POU S- or POU H-domain-specific crosslinkers. These domain-specific crosslinkers allowed us to probe independently the positions of the two Oct-1 POU DNA-binding domains on cis-regulatory sites.

We examined Oct-1 POU-domain binding to three previously described sites (see ref. 6): the high-affinity octamer sequence ATGCAAT from the human histone H2B promoter [for which the structure of the bound Oct-1 POU domain is known (2)] and two VP16-responsive TAATGARAT (R = purine) elements derived from the herpes simplex virus ICP0 [called (OCTA')TAATGARAT] and ICP4 [called (OCTA')TAATGARAT] immediate-early (IE) promoters. These two sites represent the two different kinds of TAATGARAT elements found in herpes simplex virus IE promoters (see ref. 11); only (OCTA')TAATGARAT elements contain an overlapping octamer sequence (shown in bold: ATGCTA-ATGARAT).

Fig. 1A shows a representation of the crystal structure of the Oct-1 POU domain bound to the histone H2B octamer site (2), with the POU S domain considered 5’ of the POU H domain. The structure of the Oct-1 POU domain on a TAATGARAT site is not known. Mutational analysis has suggested that on an (OCTA')TAATGARAT site, the POU S domain lies 3’ of the POU H domain, and high-resolution hydroxy-radical footprinting of an Oct-1 POU domain–(OCTA')TAATGARAT complex revealed only protection of the TAAT sequence, suggesting an unusual POU-DNA conformation on this site (12).

To generate Oct-1 POU domains with unique cysteines suitable for protein–DNA crosslinking, we created the cysteine-free double mutant POU<sub>S</sub>UC54(POU<sub>H</sub>C50S) by replacing the wild-type POU<sub>S</sub> domain with an alanine and the wild-type POU<sub>H</sub> domain cysteine with a serine. This cysteine-free Oct-1 POU domain binds like wild-type POU domain to the octamer, and (OCTA') and (OCTA') TAATGARAT sites (13). Directed by the structure of the Oct-1 POU domain–octamer site complex (2), we placed unique cysteines at a number of POU<sub>S</sub>- and POU<sub>H</sub>-domain positions that (i) were likely not to be important for DNA binding and (ii) whose α carbons were ~9–13 Å from the DNA. The ability of these mutant proteins to crosslink to DNA was analyzed by SDS/PAGE. Of six POU<sub>S</sub>-domain derivatives tested (see Materials and Methods), only POU<sub>S</sub>UC54 crosslinked the to the octamer or (OCTA')TAATGARAT sites effectively, and, of the four POU<sub>H</sub>-domain derivatives tested (see Materials and Methods), only POU<sub>H</sub>UC7 and POU<sub>H</sub>UC50 crosslinked effectively (data not shown).

Fig. 1B shows the results of crosslinking POU<sub>S</sub>UC54 and POU<sub>H</sub>UC7 to the octamer and the (OCTA')TAATGARAT sites as analyzed by SDS/PAGE in the absence of a stacking gel. The slower migrating complexes probably represent POU-
domain crosslinked DNAs because formation of these species depended on (i) added POU domain (Fig. 1B; a nonspecific slower migrating species in lanes 7 and 10 formed in the absence of POU domain), (ii) UV irradiation, (iii) modification of the protein with 4-azidophenacyl bromide, and (iv) an appropriately positioned cysteine (data not shown). Some crosslinked samples produced more than one band (lanes 3, 6, 8, 9, 11, and 12), which may reflect POU-domain crosslinking at different positions in the DNA.

The Oct-1 POU Domain Adopts Different Conformations on the Octamer and (OCTA⁻)TAATGARAT Sites. Although both strands of the octamer and (OCTA⁻)TAATGARAT sites could be crosslinked to the POU domain (Fig. 1B), only the upper strands of these probes were good substrates for site-specific DNA cleavage (data not shown). Fig. 2 shows the cleavage patterns that map sites of POU₈ and POU₁₄-domain crosslinking to the upper strands of these two sites. On the upper strand of the octamer probe, the primary site of crosslink-induced cleavage by the modified POU₈ULE54 protein (lane 2) maps to the G at the third position of the octamer sequence (ATGCAAAT) and is consistent with the POU domain–octamer site crystal structure (2). The POU₁₄ULE7 protein generated two cleavage products (Fig. 2, lane 3). These products map to the G and C at positions 3 and 4 (ATGCAAAT) and are also consistent with the POU domain–octamer site crystal structure (2).

On the (OCTA⁻)TAATGARAT site (Fig. 2, lanes 4–6), the primary POU₁₄ULE7 protein crosslink-induced cleavage (lane 6) maps to the G just 5' of the TAATGARAT sequence (GTAATGAGAT). The location of this crosslink suggests that the TAAT sequence serves as the POU₁₄-domain-binding site the same way the AAAT sequence does in the octamer sequence. In contrast, the major POU₈ULE54 cleavage product maps to the opposite side of the TAAT POU₁₄-domain-binding site, at an A immediately 3' of the GARAT sequence (TAATGAGAT). Together with the POU₁₄-domain cleavage product, this predominant POU₈-domain cleavage product provides physical evidence that the POU domain can adopt a conformation on an (OCTA⁻)TAATGARAT sequence that is very different from its conformation on an octamer sequence.

In addition to the major 3' POU₈-domain cleavage product, there are several minor cleavage products including one resulting from a crosslink to the G four positions 5' of the TAAT sequence (GCCGTAATGAGAT). This crosslink does not map to the same 5' position as on the octamer site (see Fig. 2 Lower) and may result from binding of the POU₈ domain in an inverted orientation as seen with Pit-1 (3) and suggested for Brn-2 (5). The relative levels of cleavage 5' and 3' of the TAATGARAT sequence are consistent with previous chemical modification interference analyses (6, 14, 15) and mutational analysis (6), which indicate that the POU₈ domain is principally positioned over the GARAT element when the Oct-1 POU domain binds to this (OCTA⁻)TAATGARAT sequence.

The Oct-1 POU Domain Adopts Multiple Conformations on an (OCTA⁺)TAATGARAT Site. The (OCTA⁺)TAATGARAT site contains both octamer and (OCTA⁻)TAATGARAT sequences. Differences in Oct-1 POU domain footprinting patterns of (OCTA⁺)TAATGARAT sites with wild-type and mutant GARAT sequences have suggested that the Oct-1 POU domain adopts an unusual conformation on the wild-type site (12). We used domain-specific crosslinking to probe this unusual conformation.

Fig. 3 shows the cleavage patterns resulting from crosslinking to both strands of the (OCTA⁺)TAATGARAT site. The POU₁₄ULEC5 protein crosslinked to both strands of this site just upstream of the TAAT sequence (lanes 3 and 6), suggesting that, on this site, the POU₁₄ domain binds uniquely to the TAAT sequence in the same orientation as it binds to the octamer and (OCTA⁻)TAATGARAT sites. In contrast, the POU₈ domain crosslinked to the (OCTA⁺)TAATGARAT site at several locations both 5' as on an octamer site, and 3', as on an (OCTA⁻)TAATGARAT site, of the TAAT sequence (Fig. 3, lanes 2 and 5, and Lower), suggesting that the Oct-1 POU domain adopts at least two very different conformations on this site. In the region 5' of the TAAT POU₁₄-domain-binding site, the POU₈ domain generates two crosslink-induced cleavages on each strand (Fig. 3, lanes 2 and 5, and Lower). The TAAT-proximal cleavages probably reflect Oct-1 binding in an octamer-site-like conformation; the other two cleavages may reflect an inverted POU₁₄-domain orientation similar to that described for the Pit-1 POU domain (3).
To examine whether the POU S domain can bind independently either 5′ or 3′ of the POUH domain-binding site of the (OCTA)TAATGARAT sequence, we compared crosslinking of the POU S-UC54 protein bound to the wild-type (OCTA)TAATGARAT sequence with crosslinking to two mutant forms of the site in which either the 5′ or 3′ POU domain-binding site has been disrupted. As indicated in Fig. 4 by dots, we mutated residues in the POU domain-binding sites but not the precise sites of crosslinking. Consistent with independent binding of the POU S domain to the 5′ ATGC element and the 3′ GARAT element, when the GARAT region of the (OCTA)TAATGARAT sequence is mutated, we only detect cleavages in the upstream ATGC portion of the octamer site (compare lane 3 with lane 2). Likewise, when the upstream ATGC portion of the octamer site is mutated, we only detect cleavages near the GARAT sequence (compare lane 4 with lane 2). These results demonstrate that two functionally independent, but overlapping, POU domain-binding sites make up the (OCTA)TAATGARAT element.

These results may also explain why hydroxy-radical footprinting reveals a POUH-domain TAAT protection pattern but not a POU S-domain protection pattern when the POU domain is bound to an (OCTA)TAATGARAT site (12). Whereas the POU H domain apparently binds stably to the TAAT sequence, the POU S domain can bind either upstream or downstream of the POU H domain, thus serving as an ineffective inhibitor of hydroxy-radical attack of the sequences flanking the POUH domain-binding site. Consistent with this hypothesis, when the GARAT sequence is mutated, the Oct-1 POU domain effectively protects the 5′ ATGC element (12).

Together, the footprinting (12) and crosslinking (this study) results suggest that the Oct-1 POU domain structure is dynamic on an (OCTA)TAATGARAT site.

FIG. 3. The POU domain adopts multiple conformations on an (OCTA)TAATGARAT site. Shown are the cleavage products generated from POU domain–DNA crosslinking to both strands of the (OCTA)TAATGARAT site. Lanes 2 and 3, cleavage pattern generated on the upper strand. The POUH doublet may be due to an incomplete cleavage reaction. Lanes 5 and 6, cleavage pattern generated on the lower strand. Lanes 1 and 4 (A+G), purine-cleavage pattern of the strand indicated. In lanes 5 and 6, the uppermost bands represent uncleaved DNA. Brackets, arrows, and asterisk are as in Fig. 2. (Lower) Summary of positions of cleavage. Dashed lines indicate lower-yield cleavage products.

To examine whether the POU S domain can bind independently either 5′ or 3′ of the POUH domain-binding site of the (OCTA)TAATGARAT sequence, we compared crosslinking of the POU S domain to wild-type (lane 2) and mutant (lanes 3 and 4) (OCTA)TAATGARAT sites. Lane 1 (A+G), purine-cleavage pattern of the wild-type probe; lane 3, (OCTA)TAATGARAT site with the mutated 3′ GARAT sequence (underlined), . . . ATTAGCTTC . . . (mutations shown in bold); lane 4, (OCTA)TAATGARAT site with the mutated 5′ OCTA sequence (underlined), . . . GCCAGCTA . . . (mutations shown in bold). Brackets and vertical arrows are as in Fig. 2. (Lower) Sequence with sites of mutations marked by dots and crosslinking sites indicated. Dashed line indicates lower-yield cleavage product.

FIG. 4. The POU S domain binds to the wild-type half-site in mutant forms of the (OCTA)TAATGARAT site. Shown are the cleavage products generated from crosslinking of the POU S domain to wild-type (lane 2) and mutant (lanes 3 and 4) (OCTA)TAATGARAT sites. Lane 1 (A+G), purine-cleavage pattern of the wild-type probe; lane 3, (OCTA)TAATGARAT site with the mutated 3′ GARAT sequence (underlined), . . . ATTAGCTTC . . . (mutations shown in bold); lane 4, (OCTA)TAATGARAT site with the mutated 5′ OCTA sequence (underlined), . . . GCCAGCTA . . . (mutations shown in bold). Brackets and vertical arrows are as in Fig. 2. (Lower) Summary of positions of cleavage. Dashed lines indicate lower-yield cleavage products.

VP16 Associates with Multiple Oct-1 POU-Domain Conformations. The multiple conformations of the Oct-1 POU domain on the (OCTA)TAATGARAT site raise the issue of how the Oct-1 POU domain is structured in a VP16-induced complex. The structure of the VP16-induced complex is unclear, principally because VP16 binds DNA poorly in the absence of Oct-1 (16, 17). Thus, although the GARAT and neighboring 3′ sequences are specifically important for forming a VP16-induced complex (18–22), it is not known whether
these sequences induce an Oct-1 POU-domain conformation compatible with recognition by VP16 (12) or create a VP16-binding site (see ref. 4).

To probe the conformation of the Oct-1 POU domain in a VP16-induced complex, we devised a crosslinking strategy analogous to chemical modification interference analysis. We first crosslinked either the POUS or POUH domain to the (OCTA<sup>+</sup>)TAATGARAT site and then selected those crosslinked species that could still form a complex with VP16 (see Materials and Methods). If a particular conformation of the POU domain is incompatible with VP16-induced complex formation, crosslinks resulting from that conformation should interfere with VP16-induced complex formation and be missing in the DNAs selected for VP16-induced complex formation. We refer to this strategy as “crosslinking interference analysis.”

Fig. 5 shows the results of the crosslinking interference analysis with POUS<sup>+</sup> (POUS<sub>UC54</sub>, lanes 2–4) and POUH<sup>+</sup> (POUH<sub>UC50</sub>, lanes 5–7) specific crosslinkers. Because the level of POUH<sub>UC50</sub> cleavage on the (OCTA<sup>+</sup>)TAATGARAT site was not sufficient for the crosslinking interference analysis (see Fig. 3), we used the natural POU<sub>H</sub>-domain cysteine in POUH<sub>UC50</sub> for this analysis. Cysteine 50 lies in the DNA-recognition helix and contacts the DNA major groove in the octamer cocrystal structure (2); we did not use this crosslinker in the analyses described above to avoid crosslinker-induced changes in POU<sub>H</sub>-domain binding to the DNA.

Because cysteine 50 is directed 3′ of the octamer sequence (2) in the Oct-1 POU-domain–octamer site crystal structure, we expected POUH<sub>UC50</sub> crosslinking to occur within the GARAT sequence. The observed crosslink-induced cleavage, however, maps 5′ of this region at the G residue in the C:G base pair at the fourth position within the octamer sequence (ATGCTAAT; Fig. 5, lane 5). This difference may reflect an alteration of the precise conformation of the POUH<sub>UC50</sub> domain on DNA induced by (i) the different binding sites [H2B octamer vs. (OCTA<sup>+</sup>)TAATGARAT] or (ii) the crosslinking moiety. Whichever the case, the POUH<sub>UC50</sub> crosslink interferes very strongly with VP16-induced complex formation (compare lanes 6 and 7). These results demonstrate that it is possible to interfere with VP16-induced complex formation by crosslinking and suggest that the conformation, precise position, or flexibility of the POU<sub>H</sub> domain on the (OCTA<sup>+</sup>)TAATGARAT site is critical for association of the Oct-1 POU domain with VP16.

In striking contrast to the strong interference caused by the POUH<sub>UC50</sub> crosslink, the POUS<sub>UC54</sub> crosslinks, either at or near the octamer ATGC sequence (OCTA X-links) or near the GARAT sequence (GARAT X-link), are compatible with VP16-induced complex formation (Fig. 5, compare lanes 3 and 4). An apparent slight decrease in crosslink-induced cleavages at the octamer ATGC sequence (compare lanes 3 and 4) may suggest a preference for POUS<sub>D</sub>-domain binding to the GARAT sequence in a VP16-induced complex. Nevertheless, crosslinking of the POUS<sub>D</sub> domain to the ATGC sequence is compatible with VP16-induced complex formation, demonstrating that VP16 can associate with multiple conformations of the Oct-1 POU domain.

**DISCUSSION**

We have probed the structure of the Oct-1 POU domain on different cis-regulatory sites and in a multiprotein complex. The structures of DNA-bound POU domains have been previously studied by x-ray crystallography (2, 3) and by mutagenesis (5, 6); crystallography, however, inherently provides only a static view of macromolecular interactions, and mutational analyses are indirect. Here, we have used protein–DNA photocrosslinking in a domain-specific manner to probe directly the conformation of the Oct-1 POU domain on DNA and in the VP16-induced complex. The photocrosslinking method we used (7) was particularly well suited for this analysis because the crosslinker is placed at a unique position on the protein and multiple sites of crosslinking to the DNA can be easily mapped with end-labeled probes. The results of our crosslinking studies of the two DNA-binding domains of the Oct-1 POU domain on different regulatory sites revealed the dynamic nature of Oct-1 binding to a single regulatory site—the (OCTA<sup>+</sup>)TAATGARAT site—and crosslinking interference revealed flexibility of the Oct-1 POU domain in VP16-induced complex formation.

We used three different Oct-1-binding sites in our studies—the high-affinity histone H2B octamer sequence and two VP16-responsive elements, an (OCTA<sup>+</sup>) and (OCTA<sup>+</sup>) TAATGARAT site—and each has revealed different arrangements of the Oct-1 POU domain as illustrated in Fig. 6. The pattern of crosslinking to the octamer site agrees with that predicted by the crystal structure (2), suggesting that, in solution, the POU-domain structure adopts the unique conformation elucidated crystallographically. Crosslinking to the (OCTA<sup>+</sup>)TAATGARAT site provides physical evidence for the unexpected structure of the Oct-1 POU domain revealed on this site by mutational analysis (6), with the POUS<sub>D</sub> domain principally positioned over the GARAT element and thus to
The crosslinking interference analysis described here with an (OCTA\(^+\))TAATGARAT site shows that, whether the POUS domain lies over the OCTA element 5’ of the POUH domain or the GARAT element 3’ of the POUH domain, the Oct-1 POU domain can have a selective effect on VP16-induced complex formation on an (OCTA\(^+\))TAATGARAT site (6).

The crosslinking interference analysis described here with an (OCTA\(^+\))TAATGARAT site shows that, whether the POUS domain lies over the OCTA element 5’ of the POUH domain or the GARAT element 3’ of the POUH domain, the Oct-1 POU domain can have a selective effect on VP16-induced complex formation on an (OCTA\(^+\))TAATGARAT site (6).

In conclusion, these studies reveal an extraordinary level of structural versatility and adaptability in how two macromolecules—a protein and DNA—interact with each other. As the DNA changes structure, as in the different sequences of the cis-regulatory sites, the protein—the Oct-1 POU domain—responds by also changing its structure.

We thank N. Hernandez, L. Joshua-Tor, B. Stillman, and W. Tansey for critical readings of the manuscript, and J. Duffy and P. Renna for artwork. P.S.P. thanks N. Hernandez for support and encouragement. These studies were supported by Public Health Service Grant CA13106 from the National Cancer Institute.