Biochemistry

 Volume 33, Number 9 March 8, 1994

Accelerated Publications

Bipartite DNA Recognition by the Human Oct-2 POU Domain: POU_S-Specific Phosphate Contacts Are Analogous to Those of Bacteriophage λ Repressor†

Martyn C. Botfield‡,§ and Michael A. Weiss*,‡,∥,⊥

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, and Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts 02114

Received November 1, 1993®

ABSTRACT: The POU motif defines a family of eukaryotic transcription factors broadly involved in tissue-specific gene expression and developmental regulation. The motif contains two DNA-binding domains: an N-terminal POU-specific domain (POU_S) and C-terminal homeodomain (POU_{HD}). Surprisingly, POU_S has recently been found to be similar in structure to helix–turn–helix (HTH) domains of phage repressor and Cro proteins [Assa-Munt, N., Mortishire-Smith, R., Aurora, R., Herr, W., & Wright, P. E. (1993) *Cell 73*, 193–205; Dekker, N., Cox, M., Boelens, R., Verrijzer, C. P., van der Vliet, P. C., & Kaptein, R. (1993) *Nature 362*, 852–855]. Because POU_{HD} and POU_S are expected to bind DNA differently, we have used "methylphosphonate interference" to investigate the alignment of their HTH elements in a specific DNA complex. This neutral phosphate analogue, originally developed for applications in antisense drug design [Miller, P. S., & Ts'o, P. O. P. (1987) *Anti-Cancer Drug Des. 2*, 117–128], is shown to provide a sensitive probe for sites of backbone-specific protein–DNA interaction. Inferred POU_S–phosphate contacts are in striking accord with cocrystal structures of bacteriophage repressor and Cro proteins. Alignment of POU_{HD} and POU_S in successive major grooves in each case predicts unique HTH–adenine contacts. This prediction is verified using DNA base analogues to effect interchange of AT functional groups by the method of 2′-deoxyinosine/5-methyl-2′-deoxycytosine substitution [McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238–7245]. Our results strongly support the hypothesis that the DNA-binding properties of POU_S are analogous to those of bacteriophage λ repressor.

Protein–DNA recognition is mediated by families of structural motifs (Pabo & Sauer, 1984, 1992). Of particular importance is the helix–turn–helix (HTH), first recognized in crystal structures of prokaryotic regulatory proteins (Anderson et al., 1981; McKay & Steitz, 1981; Pabo & Lewis, 1982; Steitz et al., 1982; Sauer et al., 1982). Crystallographic studies of protein–DNA complexes demonstrated that the prokaryotic HTH binds in the major groove of B-DNA (Anderson et al., 1985, 1987; Jordan & Pabo, 1988; Aggarwal et al., 1988; Wolberger et al., 1988) with characteristic base and phosphate contacts (Pabo et al., 1990; Harrison & Aggarwal, 1990; Pabo & Sauer, 1992). The N-terminus of the first α-helix contacts the DNA backbone whereas the N-terminus of the second α-helix (the "recognition helix"; Wharton & Ptashne, 1985) projects into the major groove.

Eukaryotic HTH elements were first recognized in sequences of homeodomains (Laughon & Scott, 1984; Shepherd et al., 1984). The homeodomain defines a conserved family of transcription factors broadly involved in regulation of gene expression (Scott et al., 1989). Its solution structure verified

† This work was supported in part by grants from the NIH and Council for Tobacco Research to MAW and by a Cancer Research Institute Postdoctoral Fellowship to MCB.

* Address correspondence to this author at the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School.

‡ Harvard Medical School.

§ Present address: ARIAD Pharmaceuticals, Inc. Cambridge, MA 02139.

∥ Massachusetts General Hospital.

⊥ Present address: Center for Molecular Oncology and Departments of Biochemistry and Molecular Biology and of Chemistry, The University of Chicago, Chicago, IL 60637-5419

® Abstract published in *Advance ACS Abstracts*, February 1, 1994.

0006-2960/94/0433-2349$04.50/0

the predicted HTH in an otherwise novel fold (Qian et al., 1989). Structures of homedomain–DNA complexes revealed that the HTH effects DNA recognition by a mechanism different from that of prokaryotic HTH domains (Otting et al., 1990; Kissinger et al., 1990; Wolberger et al., 1991). The center of the second HTH α-helix (rather than its N-terminus) binds in the major groove, leading to an altered pattern of side chain–base contacts and reorientation of the first HTH α-helix relative to the DNA backbone. In addition to its HTH, the homeodomain contains an N-terminal arm, which contacts the DNA minor groove.

The POU motif defines a subgroup of homeodomain-containing transcription factors (Bodner et al., 1988; Ingraham et al., 1988; Strum et al., 1988; Clerc et al., 1988; Ko et al., 1988; Schneidereit et al., 1988; Finney et al., 1988; Herr et al., 1988). The motif consists of an N-terminal POU-specific subdomain (POU_S) and C-terminal homeodomain (POU_{HD}). A large family of homologous sequences has been cloned from eukaryotic libraries [for review, see Verrijzer and Van der Vliet, 1993)]. A well-studied example is provided by Pit-1, which specifies the identity of lactotrophic and somatotrophic cells in the anterior pituitary of placental mammals (Bodner et al., 1988; Ingraham et al., 1988). A mutation in the Pit-1 1 POU_S subdomain has been identified in association with human dwarfism (Pfaffle et al., 1992). The present study focuses on human Oct-2, which participates in the regulation of immunoglobulin genes in B-cells (Clerc et al., 1988; Ko et al., 1988; Müller-Immerglück et al., 1988; Scheidereit et al., 1988).

POU_S and POU_{HD} each contribute base-specific contacts and are jointly required for DNA recognition (LeBowitz et al., 1988; Kristie & Sharp, 1990; Verrijzer et al., 1990, 1992). The two domains fold independently (Botfield et al., 1992). Surprisingly, the structure of POU_S is similar to the HTH domains of phage repressor and Cro proteins (Assa-Munt et al., 1993; Dekker et al., 1993). This similarity, which was not anticipated by sequence comparisons, predicts an analogous mechanism of DNA recognition. This hypothesis is consistent with the strong conservation of solvent-exposed residues in the putative recognition α-helix (Assa-Munt et al., 1993) and with the results of site-directed mutagenesis (Sturm & Herr, 1988). In the present study the orientations of POU_{HD} and POU_S HTH elements are defined in successive DNA major grooves by analysis of phosphate and adenine-specific contacts. A novel methylphosphonate interference assay is used to demonstrate that POU_S exhibits DNA-binding properties analogous to those of bacteriophage repressor and Cro proteins.

MATERIALS AND METHODS

Synthetic Methods. 5-Methyl-2′-deoxycytidine (5mC), kindly provided by Dr. L. C. Sowers (City of Hope Medical Center, Duarte, CA), was synthesized from thymidine as described (Divakar & Reese, 1982) and converted to 5′-*O*-(4,4′-dimethyoxytrityl)-3′-*O*-(2-cyanoethyl)-*N*,*N*-diisopropyl-phosphoramidite by standard methods; the structure and purity of the phosphoramidite was verified by ^1H and ^{31}P NMR. The presence of the modified base in oligonucleotides was verified by GC/MS following acid hydrolysis and by analysis of limit dinucleotides following enzymatic digestion with nuclease P1 and bacterial alkaline phosphatase (Sigma, Inc.). Oligonucleotides, including methylphosphonate derivatives (Miller & Ts'o, 1987), were otherwise obtained from Oligos, Etc. (Ridgefield, CT). The methylphosphonate synthesis yielded both *R* and *S* diastereomers.

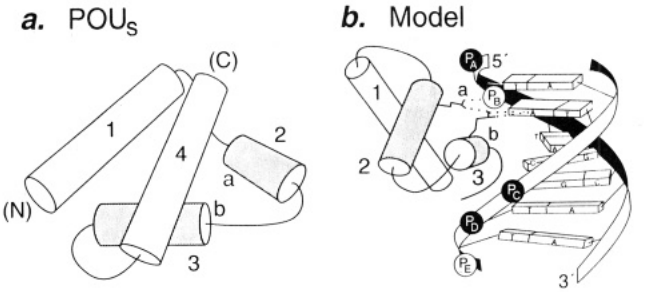

FIGURE 1: (a) Cylinder model of Oct-1 POU_S (Assa-Munt et al., 1993; Dekker et al., 1993). The putative HTH element is shaded (α-helices 2 and 3); Gln is invariant at N-termini of HTH α-helices (positions a and b). (b) Model of Oct-1 POU_S–DNA complex based on N-terminal domain of λ repressor (Jordan & Pabo, 1988). Positions of phosphate contacts in the λ cocrystal structure (designated P_{A-E}) are shown (see Figure 2A). Gln a and b (residues 33 and 44 in λ repressor) are proposed to participate in network of hydrogen bonds (dotted lines). The 5′ sequence of the octamer motif is shown. Panel b is adapted from Pabo and Sauer (1992).

Protein Expression and Purification. Oct-2 POU domain and POU_{HD} fragment were expressed and purified as described (Botfield et al., 1992).

DNA-Binding Assays. Methylphosphonate interference studies use 14 base pair duplexes containing a consensus octamer (5′-GGTATGCAAATGGT). Each modified strand was annealed with an unmodified complementary strand. The terminal 3′ and 5′ OH on each strand were not phosphorylated in the syntheses; in each case the 5′ OH of the unmodified strand was enzymatically phosphorylated with [^{32}P]ATP to provide a radioactive label. The DNA site used in I-5mC studies is the 12-mer 5′-GTATGCAAATGG. Gel-retardation assays (GRA; Fried & Crothers, 1981) were done at 4 °C under low-salt conditions as described (Botfield et al., 1992). Shifted and unshifted counts were quantified using a PhosphoImager (Molecular Dynamics, Inc., Sunnyvale, CA).

MODEL

The structure of Oct-1 POU_S (Figure 1a) contains four α-helices whose orientation is similar to α-helices 1–4 of the DNA-binding domain of λ repressor (Assa-Munt et al., 1993; Dekker et al., 1993). α-Helices 2 and 3 (shaded in figure) comprise a putative HTH. Its proposed alignment in the major groove, obtained by analogy to λ repressor (Jordan & Pabo, 1988), is shown in Figure 1b. This alignment predicts a specific pattern of contacts with DNA phosphates as identified in the λ repressor–operator complex (Johnson, 1980) by ethylation interference (Siebenlist & Gilbert, 1980). Sites of interference (designated P_{A-E} and $P_{A'-E'}$; Figure 2A) are symmetrical in each half-site:

```
    ∇ ∇
5'  T A T C A C C G C
    1 2 3 4 5 6 7 8 9
3'  A T A G T G G C G
        Δ Δ Δ
```

The structural basis of the ethylation footprint is shown in Figure 2B.

(i) Upper Strand. The Gln33 main chain NH and side chain NH_2 donate hydrogen bonds to P_A and P_B, respectively. Gln33, the N-terminal residue of the first HTH α-helix (position a in Figure 1), is conserved among phage repressor and Cro proteins (Sauer et al., 1982; Pabo & Sauer, 1984); similar contacts are observed in structures of homologous complexes (Harrison & Aggarwal, 1990). In each case the Gln side chain is also hydrogen bonded to a conserved Gln at

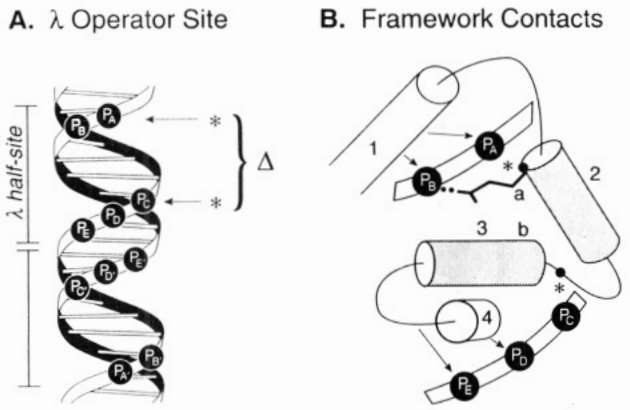

FIGURE 2: (A) λ repressor–operator phosphate contacts inferred from ethylation interference (Johnson, 1980). Sites of interference in upper and lower half-sites are symmetric (P_{A-E} and $P_{A'-E'}$, respectively). (B) Phosphate contacts in cocrystal structure of N-terminal domain of λ repressor (Jordan & Pabo, 1988). Spacing between contacts on opposite strands (Δ in panel A) reflects orientation of HTH (shaded) in major groove. Asterisks indicate hydrogen bonds between DNA (P_A and P_C) and peptide NH (Gln33 and Gly43). Arrows in B indicate side chain interactions from α-helix 1 (P_A and P_B), the turn between α–helices 3 and 4 (P_E), and α-helix 4 (P_D). Panel B is adapted from Pabo and Sauer (1992).

the N-terminus of the second HTH α-helix (Gln44 in λ repressor; position b in Figure 1). The Gln44 carboxamide is coplanar with an adenine in the major groove (A_2; boldface above) and forms bidentate hydrogen bonds with N7 and 6-NH_2 (Figure 1b); no other adenines are contacted. In the λ repressor complex (Jordan & Pabo, 1988) P_A and P_B are also contacted by side chains in α-helix 1 (arrows in Figure 2B).

(ii) Lower Strand. P_C receives a hydrogen bond from the main chain NH of the residue immediately preceding the second HTH α-helix (Gly43; asterisk in Figure 2). This contact is also conserved among homologous cocrystal structures. P_{A-C} contacts reflect the orientation of the HTH element in the major groove. P_D is contacted by a residue in α-helix 4 and P_E by a residue in the turn between helices 3 and 4.

PREDICTIONS

As among phage repressor and Cro proteins, Gln is conserved at the N-terminus of each putative POU_S HTH α-helix (positions a and b in Figure 1). Extension of the λ model to an Oct POU_S–DNA complex (Assa-Munt et al., 1983; Dekker et al., 1993) predicts the following contacts:

```
   ∇ ∇
5' T A T G C A A A T N N
   1 2 3 4 5 6 7 8 9
3' A T A C G T T T A N N
         Δ Δ? Δ?
```

As in the λ repressor–operator complex, a unique adenine contact (boldface) is also predicted. On the opposite strand contact P_C is required by the proposed HTH orientation; contacts P_D and P_E (designated Δ?), not made by the HTH element itself, are possible but not required by the model.

Binding of the intact POU motif would introduce additional contacts by POU_{HD}. These are presumably similar to those of the Engrailed homeodomain–DNA complex (Kissinger et al., 1990) as shown in Figure 3. Cocrystal structures of

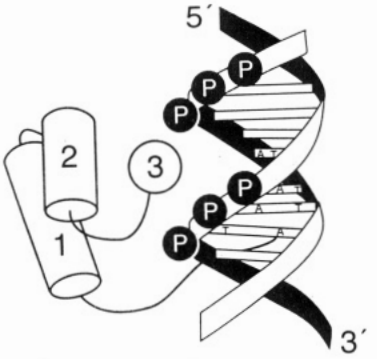

FIGURE 3: Cylinder model and phosphate contacts of Engrailed homeodomain–DNA complex (Kissinger et al., 1991). Both upper and lower phosphates are contacted by side chains in helix 3 (W48, R53, K55, and K57); additional contacts are provided by side chains in α-helix 2 (R31), preceding loop (Y25), and N-terminal arm (T6).

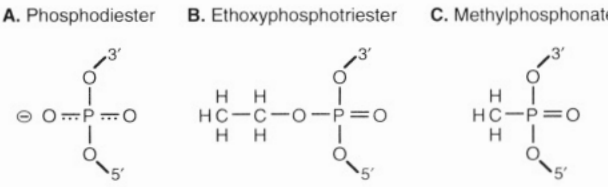

FIGURE 4: Structures of (A) native phosphodiester linkage, (B) ethoxyphosphotriester (Siebenlist & Gilbert, 1980), and (C) methylphosphonate (Miller & Ts'o, 1987), an uncharged and almost isosteric analogue.

Engrailed and MATα2 (Wolberger et al., 1991) predict the following additional contacts (■):

```
   ∇ ∇          ■ ■ ■
5' T A T G C A A A T N N N N
   1 2 3 4 5 6 7 8 9
3' A T A C G T T T A N'N'N'N'
         Δ Δ? Δ?     ■ ■ ■
```

An additional major groove adenine–carboxamide contact (A_8; boldface above) is predicted to involve an invariant Asn in the POU_{HD} recognition α-helix. This interaction is analogous to that predicted between A_2 and an invariant Gln in POU_S. Phosphate contacts of the homeodomain (unlike those of prokaryotic HTH domains) are mediated entirely by side chains; the orientation of the homeodomain HTH does not permit peptide NH–phosphate hydrogen bonds (Kissinger et al., 1990; Wolberger et al., 1991).

RESULTS

Twenty-six 14-mer oligonucleotides were prepared, each containing a single methylphosphonate (Miller & Ts'o, 1987). The structure of this modification is shown in Figure 4. At the site of modification the negative phosphate charge is removed and the local hydrophobic character of the backbone is increased; in addition, a small steric perturbation is introduced.[1] The probe sequence (5′-GGTATCGAAATGGT) contains a consensus octamer (5′-ATGCAAAT; underlined). Previous studies of Oct-1 POU fragments have demonstrated that the octamer motif contains overlapping binding sites for POU_S (5′-GGTATGCAAATGGT) and POU_{HD} (5′-GGTATGCAAATGGT); overlapping base pairs appear to be contacted in the major groove by POU_S and in the minor groove by POU_{HD} (Verrijzer et al., 1990, 1992). The 14-mer truncates potential

[1] Ethylation interference introduces a ~4.5-Å steric perturbation (Figure 4b); methylphosphonate introduces a ~0.5-Å steric perturbation (Figure 4c).

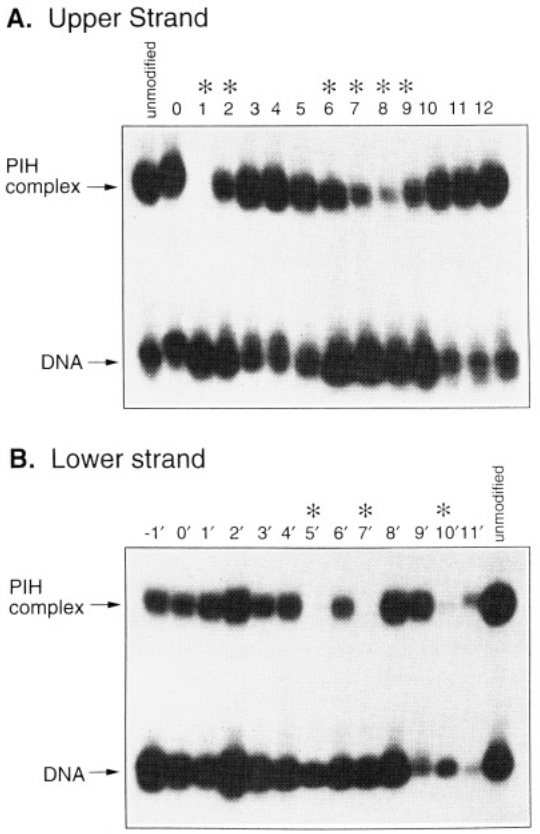

FIGURE 5: GRA autoradiogram of Oct-2 POU binding to variant octamer sites containing methylphosphonate at specific sites in upper strand (A) and lower strand (B). Asterisks indicate sites of interference numbered as given in text. The protein concentration was 50 nM.

POU_{HD} phosphate contacts on the lower strand. Its affinity for the Oct-2 POU domain (K_d) is ca. 25 nM under conditions used.

Relative affinities of native and modified duplex oligonucleotides were tested by GRA as shown in Figure 5 and in quantitative form in Figure 6. Complete or partial interference[2] is observed at the following sites (asterisks in Figures 5 and 6).

```
  ▽ ▽       ■ ■ ■ ■
5' T A T G C A A A T N N N
   1 2 3 4 5 6 7 8 9
3' A T A C G T T T A N'N'N'
          Δ   Δ     ■
```

Presumed POU_S contacts (▽) are in accord with the prediction of the λ model. These contacts do not significantly affect the low-affinity binding of the isolated POU_{HD} domain. Assignment of upper-strand contacts 6–9 (■) to POU_{HD} is supported by observation of corresponding effects on POU_{HD} binding at these positions (asterisks in Figure 7).

(i) Upper Strand. Modification of the phosphate 5′ to T_1 (corresponding to predicted main chain contact P_A; asterisk in Figure 2B) leads to the almost complete absence of an observable complex whereas modification of the adjacent phosphate (a predicted side chain contact; P_B-Gln a in Figure 2B) leads only to a 3-fold reduction in protein binding. POU_{HD} phosphate contacts in the upper strand are also observed as partial interferences. The most severe are 5′ to A_7 and A_8; the first corresponds to a contact with the invariant Trp side chain in the HTH recognition α-helix, and the second to a contact with a side chain in the POU_{HD} "arm" (Kissinger et al., 1990; Wolberger et al., 1991). An additional interference is seen 5′ to T_9, which is not predicted by the Engrailed or MATα2 cocrystal structures.

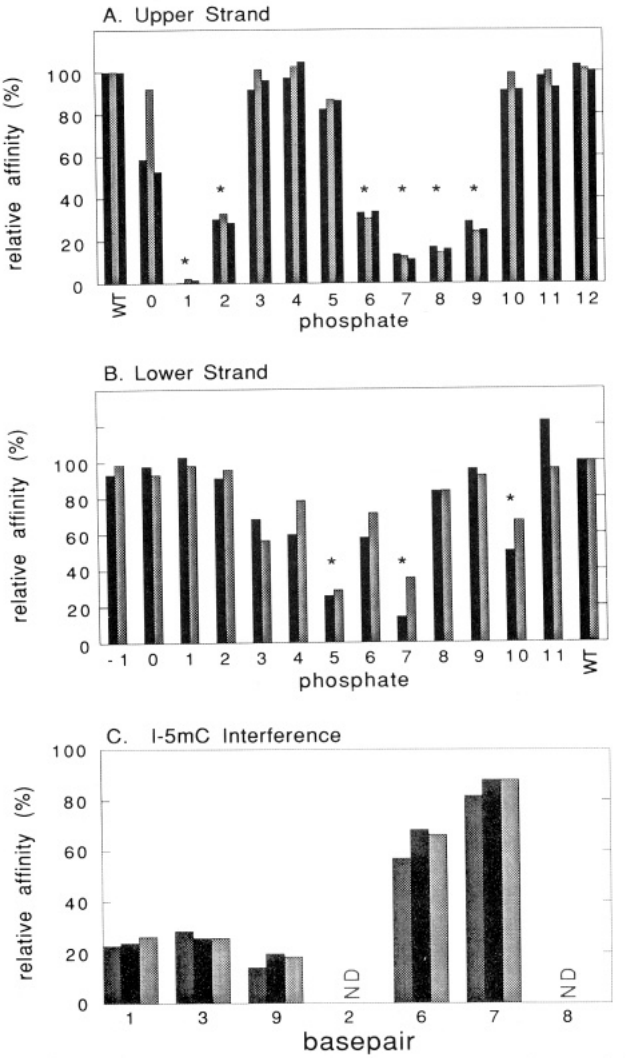

FIGURE 6: Quantitative analysis of gel-retardation experiments. (A) Effects of methylphosphonate in upper strand (see Figure 5A); values are shown in triplicate. (B) Effects of methylphosphonate in lower strand (see Figure 5B); values are shown in duplicate. (C) Effects of functional group interchange (I-5mC interference; see Figure 8A); values are shown in triplicate. In each panel the vertical axis represents the percent DNA shifted relative to the native DNA site (defined to be 100%) under conditions in which approximately 50% of the native site is shifted. In panels A and B asterisks indicate sites of interference, and "WT" indicates the native DNA site; in panel C "ND" indicates no shifted complex was detectable (≪1% shifted).

(ii) Lower Strand. Modification of phosphates 5′ to G5 and T7 (predicted P_C and P_E contacts; Figure 2) leads to partial interference. No evidence is obtained for a contact at P_D; in phage repressor cocrystal structures this contact is made by a side chain in α-helix 4, which is not conserved among POU_S sequences. Weak interference is observed following modification of the phosphate 5′ to G_{10} and ascribed in POU_{HD}. No interference is observed 5′ to the terminal base (it is possible that this contact is perturbed by the short length of the oligonucleotide).

[2] *R* and *S* diastereomers, although each uncharged, would be expected to influence differentially stereoselective hydrogen bonds between the protein and a nonester oxygen. Accordingly, partial interference in the present assay can in principle represent strong interference by one diastereomer and weak interference by the other. Absence of stereoselectivity also limits interpretation of partial ethylation interference (Siebenlist & Gilbert, 1980).

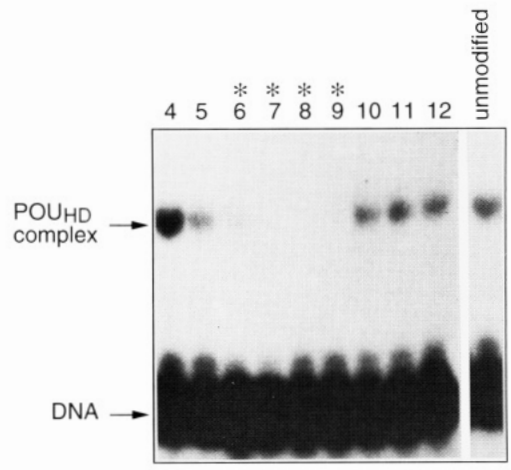

FIGURE 7: GRA autoradiogram of Oct-2 POU_{HD} fragment binding to selected methylphosphonate analogues in the upper strand. Interferences at positions 6–9 (asterisks) are analogous to those observed with the intact POU motif (Figure 5). The POU_{HD} concentration was *ca* 4 μM. Corresponding lower-strand data reveal no significant sites of interference (Supplemental Material).

Consistency of the observed pattern of POU phosphate contacts with cocrystal structures of repressors and homeodomains in turn requires that A_2 and A_8 (boldface above) be essential major-groove contacts. The N7 and 6-NH_2 of each is predicted to form bidentate hydrogen bonds with a side chain carboxamide (see Predictions). The inferred alignment of POU_S and POU_{HD} HTH elements further suggests that no other adenine would be critical. These predictions are tested using the functional group interchange method (McLaughlin et al., 1987). Substitution of an AT base pair by I-5mC interchanges the order of hydrogen bond donor (adenine 6-NH2) and acceptor (thymidine 4 C=O) in the major groove; functional groups in the minor groove are unchanged (Figure 8C). Such interchange eliminates the carboxamide binding site on the purine (Seeman et al., 1976). Effects of I-5mC substitution on POU binding are shown in Figure 8A; quantitative results are shown in Figure 6C and summarized in Figure 8B. As predicted, substitution at positions 2 or 8 prevents formation of a detectable protein–DNA complex (asterisks in Figure 8A); the sensitivity of the assay implies that the association constant has in each case been decreased by at least 100-fold. Assignment of the A2 contact to POU_S and the A8 contact to POU_{HD} is supported by comparison of the sequence specificities of the Oct-2 POU and POU_{HD} fragments at these positions. The two proteins have identical specificities at position 8 (A $\gg$ T, C, or G); in contrast, at position 2 the POU domain is specific for A whereas POU_{HD} lacks specificity (Supplemental Material) in accord with the previous studies of Oct-1 POU fragments (Verrijzer et al., 1990, 1992).

I-5mC substitutions at positions 1, 3, and 9 reduce complex formation somewhat (Figure 6C). Such effects may reflect either perturbation of adjoining thymine contacts (as would be expected, for example, at T9; Kissinger et al., 1990) or indirect effects of changes in DNA structure.

DISCUSSION

POU transcription factors recognize DNA using distinct HTH scaffolds. Whereas POU_{HD} is a homeodomain, POU_S represents a novel sequence motif. The structure of Oct-1 POU_S has recently been determined and, surprisingly, is similar to that of the DNA-binding domains of phage repressor and Cro proteins (Assa-Munt et al., 1993; Dekker et al., 1993). The putative POU_S HTH element exhibits strong conservation among eukaryotes, including invariant Gln residues at the N-terminus of each α-helix. Analogous Gln residues are conserved among phage repressor and Cro proteins (but not more generally among prokaryotic HTH domains; Sauer et al., 1982; Sauer & Pabo, 1984) and are required for high-affinity DNA binding (Nelson & Sauer, 1986).

Because the structural similarity between POU_S and prokaryotic HTH domains predicts a shared mechanism of DNA binding, we have investigated the structure of a POU–DNA complex by study of DNA analogues. Methylphosphonate interference (like ethylation interference footprinting; Siebenlist & Gilbert, 1980; Bushman et al., 1985) identifies phosphate contacts by neutralizing the backbone charge (Figure 4C). The methylphosphonate introduces a significantly smaller steric perturbation than ethylation (which yields the ethoxyphosphotriester derivative; Figure 4B),[1] and its incorporation in individual oligonucleotides permits quantitative assessment of changes in protein affinity.[2] Methylphosphonate oligonucleotides were originally developed for applications in antisense drug design (Miller & Ts'o, (1987). Molecular dynamic simulations suggest that the analogue is well-tolerated in B-DNA, primarily affecting the local electrostatic potential and counterion distribution (Hausheer et al., 1990). To our knowledge, the present study is its first application as a probe of protein–DNA interactions. We have also used functional group interchange (substitution of AT by I-5mC) to test the sensitivity of protein binding to an altered pattern of major-groove hydrogen-bond donors and acceptors as previously described (McLaughlin et al., 1987).

Our results, summarized in Figure 9, are in accord with the combined predictions of cocrystal structures of prokaryotic HTH domains (Pabo et al., 1990; Harrison & Aggarwal, 1990) and homedomain–DNA complexes (Kissinger et al., 1990; Wolberger et al., 1991). The figure provides a common alignment of DNA target sites for phage repressors (upper two sequences), Oct POU domains (center), and classical homeodomains (lower two sequences). In boldface are shown analogous adenine contacts. In cocrystal structures each is recognized by bidentate carboxamide–adenine hydrogen bonds, mediated by Gln in the phage recognition helix and by Asn in the homeodomain recognition helix. The octamer motif contains two essential adenines at corresponding positions in the POU_S and POU_{HD} binding sites (A_2 and A_8, respectively). The pattern of POU–phosphate contacts relative to these adenines is strikingly similar to that expected on the basis of the two sets of cocrystal structures.

Of particular interest is the almost complete absence of detectable complex formation following methylation 5′ to T_1 (upper strand). This contact is analogous to P_A in the λ repressor-operator complex (Figure 2A), which interacts with the HTH mainchain (Figure 2B). We speculate that its modification perturbs the *global* alignment of the HTH and hence leads to dramatic reduction in protein binding. In contrast, the pattern of partial interferences in the POU_{HD} binding site suggests that *local* displacement of individual side chain–phosphate contacts can be well tolerated.

In summary, the consistent alignment of the bipartite octamer motif with DNA binding sites of phage repressors (POU_S) and homeodomains (POU_{HD}) supports the hypothesis that POU_S is functionally analogous to phage HTH domains. If the products of convergent evolution, this arrangement of α-helices in the major groove must represent a particularly successful strategy for DNA recognition.

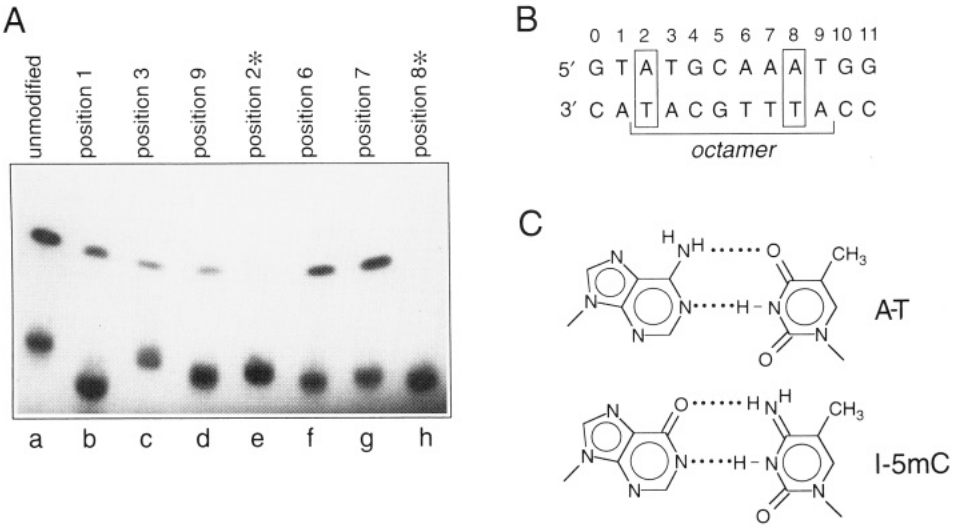

FIGURE 8: (A) GRA autoradiogram of Oct-2 POU binding to variant DNA sites containing I-5mC in place of AT. (B) Sequence of 12-mer probe; base pairs 2–9 contain the consensus octamer motif. Positions sensitive to I-5mC substitution are boxed. (C) Structures of AT and I-5mC base pairs. The protein concentration was 50 nM.

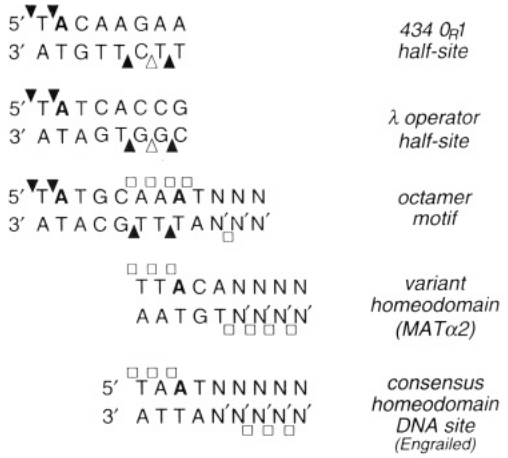

FIGURE 9: Alignment of DNA target sites for phage 434 and λ repressors (top), Oct-2 (middle), and homeodomains (bottom). Observed or proposed adenine–carboxamide contacts are shown in boldface. Phosphate contacts in repressor–operator complexes (▽, ▼, ▲, and △) and in homeodomain–DNA complexes (□) were obtained from cocrystal structures (Jordan & Pabo, 1988; Anderson et al., 1985; Aggarwal & Harrison, 1990; Kissinger et al., 1991; Wolberger et al., 1992). Phosphate contacts in the octamer motif are as obtained in Figure 5.

ADDED IN PROOF

The proposed alignment of POU_S and POU_{HD} HTH elements in successive major grooves is supported by analysis of "loss of contact" mutations in their respective recognition α-helices (Li et al., 1993; Jancso et al., 1994; Botfield et al., 1994). In particular, relaxed sequence specificity is observed at positions 2 or 8 following mutagenesis of the predicted HTH contracts (Gln 238 and Asn 347, respectively, in the Oct-2 POU motif).

ACKNOWLEDGMENT

We thank Prof. P. A. Sharp and J. LeBowitz (MIT, Cambridge, MA) for Oct-2 cDNA and advice, Dr. L. C. Sowers (City of Hope Medical Canter, Durate, CA) for oligonucleotide synthesis, A. Jancso for collaborative interactions, J. Englert and E. Collins for technical assistance, and A. Hochschild, C. Richardson, H. Singh, and K. Struhl for helpful discussion.

SUPPLEMENTARY MATERIAL AVAILABLE

Three figures showing autoradiogram of Oct-2 POU_{HD} binding to mutant octamer sites, quantitative comparison of POU and POU_{HD} sequence specificities, and binding to POU_{HD} to lower-strand methylphosphonate sites (4 pages). Ordering information is given on any current masthead page.

REFERENCES

Aggarwal, A. K., Rogers, D. W., Drottar, M., Ptashne, M., & Harrison, S. C. (1988) *Science 242*, 899–907.
Anderson, J. E., Ptashne, M., & Harrison, S. C. (1985) *Nature 316*, 596–601.
Anderson, J. E., Ptashne, M., & Harrison, S. C. (1987) *Nature 326*, 846–852.
Anderson, W. F., Ohlendorf, D. H., Takeda, Y., & Matthews, B. W. (1981) *Nature 298*, 754–758.
Assa-Munt, N., Mortishire-Smith, R., Aurora, R., Herr, W., & Wright, P. E. (1993) *Cell 73*, 193–205.
Beamer, L. J., & Pabo, C. O. (1992) *J. Mol. Biol. 227*, 177–196.
Bodner, M., Castrillo, J. L., Theill, L. E., Deerinck, T., Ellisman, M., & Karin, M. (1988) *Cell 55*, 505–518.
Botfield, M. C., Jancso, A., & Weiss, M. A. (1992) *Biochemistry 31*, 5841–5848.
Botfield, M. C., Jancso, A., & Weiss, M. A. (1994) *Biochemistry* (submitted).
Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986) *J. Biol. Chem. 261*, 7270–7278.
Bushman, F. D., Anderson, J. E., Harrison, S. C., & Ptashne, M. (1985) *Nature 316*, 651–653.
Clerc, R. G., Corcoran, L. M., LeBowitz, J. H., Baltimore, D., & Sharp, P. A. (1988) *Genes Dev. 2*, 1570–1581.
Dekker, N., Cox, M., Boelens, R., Verrijer, C. P., van der Vliet, P. C., & Kaptein, R. (1993) *Nature 362*, 852–855.
Divakar, K. J., & Reese, C. B. (1982) *J. Chem. Soc., Perkin Trans. 1*, 1171–1176.
Finney, M., Ruvkun, G., & Horvitz, H. R. (1988) *Cell 55*, 757–769.
Fried, M., & Crothers, D. M. (1981) *Nucleic Acids Res. 9*, 6505–6525.
Harrison, S. C., & Aggarwal, A. K. (1990) *Annu. Rev. Biochem. 59*, 533–569.
Hausheer, F. H., Singh, U. C., Palmer, T. C., & Saxe, J. D. (1990) *J. Am. Chem. Soc. 112*, 9468–9474.
Ingraham, H. A., Chen, R., Mangalam, H. J., Elsholtz, H. P., Flynn, S. E., Lin, C. R., Simmons, D. M., Swanson, L., & Rosenfeld, M. G. (1988) *Cell 55*, 519–529.

Jancso, A., Botfield, M. C., Sowers, L. C., & Weiss, M. A. (1994) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
Johnson, A. D. (1980) Ph.D. Thesis, Harvard University, Cambridge, MA.
Jordan, S. R., & Pabo, C. O. (1988) *Science 242*, 893–899.
Kissinger, C. R., Liu, B., Martin-Blanco, E. Kornberg, T. B., & Pabo, C. O. (1990) *Cell 63*, 579–590.
Ko, H.-S., Fast, P., McBride,W., & Staudt, L. M. (1988) *Cell 55*, 135–144.
Kristie, T. M., & Sharp, P. A. (1990) *Genes Dev. 4*, 2383–2396.
Laughon, A., & Scott, M. P. (1984) *Nature 310*, 25–31.
LeBowitz, J. H., Kobayashi, T., Staudt, L., Baltimore, D., & Sharp, P. A. (1988) *Genes Dev. 2*, 1227–1237.
Li, P., He, X., Gerrero, M. R., Mok, M., Aggarwal, A., & Rosenfeld, M. G. (1993) *Genes Dev. 7*, 2483–2496.
McKay, D. B., & Steiz, T. A. (1981) *Nature 290*, 744–749.
McLaughlin, L. W., Benseler, F., Graeser, E., Piel, N., & Scholtissek, S. (1987) *Biochemistry 26*, 7238–7245.
Miller, P. S., & Ts'o, P. O. P. (1987) *Anti-Cancer Drug Des. 2*, 117–128.
Nelson, H. C. M., & Sauer, R. T. (1986) *J. Mol. Biol. 192*, 27–38.
Otting, G., Qian, Y.-Q., Billeter, M., Müller, M., Affolter, M., Gehring, W., & Wüthrich, K. (1990) *EMBO J. 9*, 3085–3092.
Pabo, C. O., & Sauer, R. T. (1984) *Annu. Rev. Biochem. 53*, 293–321.
Pabo, C. O., & Sauer, R. T. (1992) *Annu. Rev. Biochem. 61*, 1053–1095.
Pabo, C. O., Sauer, R. T., Sturtevant, J. M., & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 59*, 1259–1263
Pabo, C. O., Aggarwal, A. K., Jordon, S. R., Beamer, L. J., Obeysekare, U. R., & Harrison, S. C. (1990) *Science 247*, 1210–1213.
Pfaffle, R. W., DiMattia, G. E., Parks, J. S., Brown, M. R., Wit, J. M., Jansen, M., Van der Nat, H., Van den Brande, J. L., Rosenfeld, M. G., & Ingraham, H. A. (1992) *Science 257*, 1210–1213.
Qian, Y. Q., Billeter, M., Otting, G., Müller, M., Gehring, W. J., & Wüthrich, K. (1989) *Cell 59*, 573–580
Sauer, R. T., Yocum, R. R., Doolittle, R. F., Lewis, M., & Pabo, C. O. (1982) *Nature 298*, 447–451.
Scheidereit, C., Cromlish, J. A., Gerster, T., Kawakami, K., Balmaceda, C.-G., Currie, R. A., & Roeder, R. G. (1988) *Nature 336*, 551–557.
Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 253*, 1001–1007.
Scott, M. P., Tamkun, J., & Hartzell, G. W., 3rd (1989) *Biochim. Biophys. Acta 989*, 25–48.
Seeman, N. C., Rosenberg, J. M., & Rich, A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 804–808.
Shepherd, J. C. W., McGinnis, W., Carrasco, A. E., DeRobertis, E. M., & Gerhing, W. J. (1984) *Nature 310*, 70–71.
Siebenlist, U., & Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 122–126.
Singh, H., Sen, R., Baltimore, D., & Sharp, P. A. (1986) *Nature 319*, 154–158.
Steitz, T. A., Ohlendorf, D. H., McKay, D. B., Anderson, W. F., & Matthews, B. W. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3097–3100.
Sturm, R. A., & Herr, W. (1988) *Nature 336*, 601–604.
Verrijzer, C. P., & Van der Vliet, P. C. (1993) *Biochim. Biophys. Acta 1173*, 1–21.
Verrijzer, C. P., Kal, A. J., & van der Vliet, P. C. (1990) *Genes Dev. 4*, 1964–1974.
Verrijzer, C. P., Alkema, M. J., Van Weperen, W. W., Van Leeuwen, H. C., Strating, M. J. J., & Van der Vliet, P. C. (1992) *EMBO J. 11*, 4993–5003.
Wharton, R. O., & Ptashne, M. (1985) *Nature 316*, 601–605.
Wolberger, C., Dong, Y. C., Ptashne, M., & Harrison, S. C. (1988) *Nature 335*, 789–795.
Wolberger, C., Vershon, A. K., Liu, B., Johnson, A. D., & Pabo, C. O. (1991) *Cell 67*, 517–528.