Oct-1 POU domain–DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage

Juli D. Klemm and Carl O. Pabo

Department of Biology and 1Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139 USA

Structural and biochemical studies of Oct-1 POU domain–DNA interactions have raised important questions about cooperativity and the role of the linker connecting the POU-specific domain and the POU homeo domain. To analyze these interactions, we have studied binding of the isolated domains. Surprisingly, we find that two unlinked polypeptides corresponding to the POU-specific domain and the POU homeo domain bind cooperatively to the octamer site and have a coupling energy of 1.6 kcal/mole. We suggest that overlapping DNA contacts near the center of the octamer site may be the source of this cooperativity, as there are no protein–protein contacts between the domains in the crystal structure of the Oct-1 POU domain–DNA complex. These studies also have allowed us to describe the thermodynamic contribution of the linker (present in the intact POU domain) in terms of an effective concentration (3.6 mM). The broader implications for understanding cooperativity in protein–DNA recognition and gene regulation are discussed.

[Key Words: POU domain, Oct-1, cooperativity, effective concentration, protein–DNA interactions, homeo domain]

Received June 20, 1995; revised version accepted November 8, 1995.

The crystal structure of an Oct-1 POU domain–octamer (ATGCAAAT) complex has recently been determined and has raised interesting questions about cooperativity in protein–DNA interactions. In the crystal structure of the complex, the two globular domains are positioned on opposite sides of the double helix and each makes specific base contacts to half of the octamer site: The POU-specific domain contacts the 5' portion of the ATG-CAAAAT site and the POU homeo domain contacts the 3' portion of this site. The two domains make overlapping contacts near the center of the site; however, there are no protein–protein contacts between the two domains, and the crystal structure suggests that the linker region is very flexible and would be long enough to allow binding with other spacings or other orientations of the domains [for review, see Herr and Cleary 1995].

In spite of the apparent structural autonomy of the two subdomains, biochemical evidence indicates that their binding is coupled. Thus, binding-site selection experiments have shown that Oct-1 strongly prefers the wild-type octamer arrangement of POU-specific domain and POU homeo domain subsites [Verrijzer et al. 1992]. In agreement with this result, DNA binding competition experiments with the Oct-1 POU domain have shown that inserting 2 or 3 bp near the center of the octamer site significantly decreases the affinity for the Oct-1 POU domain [Klemm et al. 1994].

To determine why the Oct-1 POU-specific domain and POU homeo domain prefer to bind the wild-type arrangement of octamer subsites, we have performed binding studies with purified polypeptides corresponding to the isolated domains. We find that the individual domains bind cooperatively to the octamer site, and consideration of these results in conjunction with the cocrystal structure suggests that coupling may be mediated through overlapping DNA contacts near the center of the octamer site. In particular, we show that an A/T base pair near the center of the site is contacted by both domains, and we propose that these contacts—along with flanking phosphate contacts—may help mediate cooperative binding. Our data also allow us to describe the contribution of the linker that connects the POU-specific domain and the POU homeo domain in terms of an effective concentration. These results may have important implications for understanding cooperativity in protein–DNA recognition and for understanding the mechanisms of combinatorial control of gene expression.

Results

Individual polypeptide chains containing the Oct-1 POU-specific domain and Oct-1 POU homeo domain form a ternary complex with the octamer site

Our strategy for analyzing cooperativity in POU domain–DNA interactions has involved studies of a ternary complex that contains the octamer site and two
Figure 1. The individual POU-specific domain and POU homeo domain polypeptides form a ternary complex with the octamer site, giving a “linkerless” POU domain–DNA complex. (A) Gel mobility-shift assay showing that the POU-specific domain and the POU homeo domain form a ternary complex with the octamer site. The wild-type octamer probe (lanes 1–4) is 26 bp in length and contains the sequence ATGCAAAT. In the mutant octamer probe (lanes 5–8), 2 bp in the POU-specific subsite have been changed [ATGCAAAT → CGGCAAAT]. In the experiment with the mutant [CGGCAAAT] probe, addition of the POU-specific domain to the POU homeo domain binding reaction does not produce a supershifted band, but the apparent affinity of the POU homeo domain for the DNA is increased (cf. lanes 6 and 7). Given the very high concentration of protein used in these experiments, we suspect that the POU-specific domain may still associate weakly with the mutant site. We believe that this facilitates homeo domain binding but that the interactions of the POU-specific domain with the mutant half site are weak enough that this domain falls off during electrophoresis. (B) Comparison of the footprinting patterns of the POU domain, the POU homeo domain, and a mixture containing the POU homeo domain and the POU-specific domain. The DNA probe used in this and subsequent footprinting experiments (Fig. 3) is 115 bp in length and contains the octamer sequence ATGCAAAT in the top strand. The protection patterns shown are for the bottom strand of the probe. The autoradiograph represents sections of the same gel that have been rearranged for the clarity of presentation. The footprint of the POU homeo domain (lane 3) includes four of the five prominent bands of the intact POU domain footprint (lane 2). The region of the DNA protected from DNase I cleavage by a mixture containing the POU homeo domain and the POU-specific domain corresponds precisely to the region protected by the intact POU domain (lane 4). Note that addition of the POU-specific domain protects one prominent band (asterisk, lane 4) that is not protected by the isolated homeo domain. The protein concentrations in the binding reactions are POU domain, 1 × 10⁻⁷ M; POU-specific domain, 1 × 10⁻⁸ M; POU-homeo domain, 2.5 × 10⁻⁹ M; POU-specific domain, 1 × 10⁻⁵ M.

To that of the POU domain–DNA complex, suggests that the isolated domains can, under appropriate conditions, form a ternary complex with the octamer site. DNase I footprinting provided further evidence that the complex we obtained with the isolated domains resembles the intact POU domain–DNA complex. The region of the octamer-containing DNA probe that is protected by the POU domain contains five strong DNase I bands (Fig. 1B, lane 2), whereas the region protected by the POU homeo domain includes four of these bands (Fig. 1B, lane 3). When the POU-specific domain and the POU homeo domain are combined, the footprinting pattern is extended to include all five bands that are characteristic of the intact POU domain footprint (Fig. 1B, lane 4). Thus, mixing the POU-specific domain and the POU homeo domain polypeptides can protect the same positions from DNase I cleavage that are protected by the intact POU domain. Both the footprinting and the gel mobility shift results suggest that this ternary complex resembles a “linkerless” POU domain–DNA complex, and we have used this ternary complex for dissecting the cooperative interactions.

Cooperative binding of the Oct-1 POU-specific domain and the Oct-1 POU homeo domain

Further binding studies with the isolated domains show that the POU-specific domain and the POU homeo do-
main can bind cooperatively even in the absence of the polypeptide linker. We measured the cooperativity in this system by comparing [1] the affinity of the POU-specific domain for an unoccupied octamer site with [2] its affinity for an octamer site that has a POU homeo domain bound to the AAAT subsite (Fig. 2). The difference in the free energies of binding under these two conditions reflects the cooperative interaction between the POU-specific domain and the POU homeo domain.

Comparing the DNase I protection pattern of the POU homeo domain with the combined protection pattern of the POU homeo domain and the POU-specific domain showed one clear band that could be used to monitor binding of the POU-specific domain. [This is the key difference between the footprint of the intact POU domain and the footprint of the isolated POU homeo domain.] To measure the $K_d$ of the POU-specific domain, we quantitated the disappearance of this band as we increased the concentration of this domain. We measure a $K_d$ of $2.2 \pm 0.14 \times 10^{-7}$ M for the POU-specific domain–octamer complex, and this corresponds to a $\Delta G$ of $-6.3$ kcal/mole at room temperature ($24^\circ$C) (Fig. 3A, B).

In parallel experiments, we measured binding of the POU-specific domain to a POU-homeo domain–octamer complex. In these experiments, the concentration of POU homeo domain was kept constant at $5 \times 10^{-6}$ M and this concentration of POU homeo domain should give 97% saturation of the AAAT subsite. [The $K_d$ we have measured for the POU homeo domain on the octamer sequence is $1.5 \pm 0.17 \times 10^{-7}$ M.] Variation of the concentration of the POU-specific domain and use of DNase I footprinting to monitor binding shows that the $K_d$ of the POU-specific domain under these conditions is $1.7 \pm 0.88 \times 10^{-6}$ M, corresponding to a $\Delta G$ of $-7.9$ kcal/mole (Fig. 3C, D). The higher affinity of the POU-specific domain for the POU homeo domain–octamer complex demonstrates that these subdomains bind cooperatively with a favorable interaction of $-1.6$ kcal/mole.

Cooperative binding is disrupted by mutation of the POU homeo domain binding site or by separation of the subsites

Control experiments confirm that the observed cooperativity requires both an intact homeo domain-binding site and the wild-type arrangement of the subsites that are recognized by the individual domains. One of the first control experiments was performed to confirm that occupancy of the POU homeo domain subsite was required for cooperative binding of the POU-specific domain. In this experiment, we repeated our cooperativity assay using a footprinting probe in which the POU homeo domain subsite and the flanking base pairs just outside of the octamer site (which are also contacted by the POU homeo domain in our crystal structure) were changed from AAATaa to AACGag (flanking bases in lowercase letters). No binding by the POU homeo domain could be detected on this probe, and we found that POU-specific binding to the probe was the same in the presence or absence of $5 \times 10^{-6}$ M POU homeo domain (data not shown). As expected, cooperative binding of the POU-specific domain clearly requires binding of the POU homeo domain to the AAAT subsite.

Previous studies had shown that the binding constant of the POU domain was reduced when two or three base pairs were inserted near the center of the octamer site [Klemm et al. 1994]. In parallel experiments with the isolated domains, we find that changing the spacing of the subsites abolishes the cooperative interactions. In this set of experiments, we repeated the cooperativity assay (as sketched in Fig. 2) with a footprinting probe in which two base pairs were inserted at the center of the site. We find that this altered spacing does not affect the binding constants of the isolated domains. Cooperativity is lost, however, and now the $K_d$ of the POU-specific domain is unaffected by binding of the POU homeo domain [Fig. 4]. These experiments confirm that the wild-

\[ \Delta G_{\text{cooperativity}} = \Delta G_2 - \Delta G_1 \]

Figure 2. Sketch summarizing the strategy used to measure cooperative DNA-binding by the POU-specific domain and the POU homeo domain. The sketch on the left depicts the POU-specific domain binding to the octamer site, and this reaction has a free energy $\Delta G_1$. The sketch on the right depicts the POU-specific domain binding to a POU homeo domain–octamer complex, and this reaction has a free energy $\Delta G_2$. [As described in the legend to Fig. 3, binding was monitored by quantitative DNase I footprinting experiments.] The difference in the free energies of these reactions represents the free energy of cooperativity $\Delta G_{\text{cooperativity}}$ for the POU-specific domain and the POU homeo domain.
Figure 3. The POU-specific domain and the POU homeo domain bind cooperatively even in the absence of a linker. (A) Quantitative DNase I footprinting experiment measuring the binding of the POU-specific domain to the octamer site. The concentration of the POU-specific domain was increased in twofold increments from $9.76 \times 10^{-8}$ to $2 \times 10^{-4}$ M. Disappearance of the band (asterisk) identified as specifically protected by the POU-specific domain (see legend to Fig. 1) was quantitated as a function of the concentration of this domain. With increasing POU-specific domain concentrations, we observe that some regions of the DNA probe become increasingly sensitive to cleavage by DNase I. We do not know the source of this hypersensitivity, but it should not significantly affect our measurements. All measurements of band intensity in footprinted regions are normalized to a standard region in that lane (see Materials and methods); we have normalized our measurements to the strong doublet near the top of each lane for all of our calculations. The sensitivity of this region to DNase I cleavage is not affected by increasing POU-specific concentrations, therefore, our intensity measurements should be independent of this hypersensitivity. (B) Scatchard plot for binding of the POU-specific domain to the octamer site. (C) Quantitative DNase I footprinting experiment measuring binding of the POU-specific domain to a POU homeo domain-octamer complex. The POU homeo domain concentration is $5 \times 10^{-6}$ M in all reactions, which should be sufficient to give an occupancy of 97%. The concentration of the POU-specific domain was increased in twofold increments from $9.76 \times 10^{-8}$ M to $2 \times 10^{-4}$ M. Disappearance of the band (asterisk) that is specifically protected by the POU-specific domain was quantitated as a function of the concentration of this domain. (D) Scatchard plot for binding of the POU-specific domain to the POU homeo domain-octamer complex.

$K_d = 2.2 (\pm 0.14) \times 10^{-5}$ M
$\Delta G_1 = -6.3$ kcal/mol
$K_d = 1.7 (\pm 0.88) \times 10^{-6}$ M
$\Delta G_2 = -7.9$ kcal/mol
Cooperativity in POU domain–DNA interactions

Leu-55–thymine interaction has several interesting implications. (1) Because the homeo domain also makes a contact with base pair 5 (in the minor groove), it seems more appropriate to describe the Oct-1 complex in terms of overlapping subsites rather than distinct half sites. Inclusion of this additional base contact shows that the POU-specific domain contacts an ATGCA subsite and that the POU homeo domain contacts an AAAT subsite [Fig. 5]. This revised definition of the octamer subsites does not affect any other aspects of our previous description of the complex but it is directly relevant for the current study because contacts near the center of the site may mediate cooperative binding. (2) This contact may explain why the Oct-1 POU domain prefers a site with the sequence ATGCAAAT, even though the isolated homeo domain prefers a site with the sequence TAAT [Verrijzer et al. 1992]. (3) It is also interesting that this Leu-55–thymine interaction extends the structural analogy between the λ repressor and the POU-specific domain: the positions of this residue and this base make them analogous to the Asn-55–guanine contact seen in the λ repressor–operator complex (Jordan and Pabo 1988, Beamer and Pabo 1992).

Discussion

Cooperative binding by the POU-specific domain and POU homeo domain may be mediated by subtle changes in DNA structure

Our studies have focused on understanding at a structural and physical level how the Oct-1 POU domain recognizes the octamer site. Our goal has been to understand why this bipartite DNA-binding protein prefers to bind the wild-type octamer site and why binding is less favorable if 2 bp are added near the center of the site. We also have been interested in understanding what contribution the linker makes to cooperative binding of the subdomains.

We have analyzed the mechanism of cooperativity using separate polypeptide chains for the POU-specific domain and the POU homeo domain, and we have shown that these domains bind cooperatively to the octamer site. Gel mobility-shift assays and DNase I footprinting experiments indicate that the isolated domains can mimic binding by the intact POU domain. Quantitative DNase I footprinting experiments have shown that there is a cooperative interaction energy of 1.6 kcal/mole. [Our studies compared POU-specific binding to the octamer site in the presence and absence of the homeo domain, but cooperativity should be reciprocal. The converse experiment would be technically more difficult because very high concentrations of POU-specific domain would be needed, but titrating the octamer site with the POU-specific domain should cause the POU homeo domain to bind more tightly.]

We find cooperative binding of the isolated subdomains is dependent on the wild-type spacing of subsites: Insertion of 2 bp near the center of the site abolishes the cooperative effect. We believe that cooperative binding

Figure 4. Scatchard plot for the POU-specific domain binding to an octamer site in which two base pairs have been inserted near the center of the sequence: ATGCA__CCAAAT. Open boxes indicate the POU-specific domain alone binding to this site; solid boxes indicate the POU-specific domain binding to this site in the presence of 5×10⁻⁶ M POU homeo domain. The K₅ values (derived from the slopes of the plots) are both ~2.1×10⁻⁶ M, thus, there is no cooperativity between the subdomains when their subsites are separated. Shown for comparison (broken line) is the Scatchard plot for binding of the POU-specific domain to a POU homeo domain–wild-type octamer complex [data shown in Fig. 3D].

The POU-specific domain and the POU homeo domain make overlapping contacts at the center of the octamer site

Because cooperativity may be mediated by contacts near the center of the octamer site, we reexamined the role of all residues in this region of the complex. Although we did not include this contact in our original description of the POU domain–DNA structure [Klemm et al. 1994], we now have evidence that Leu-55 of the POU-specific domain contacts the methyl group of the thymine at the fifth base pair of the octamer site [ATGCAAAT]. Leu-55, in the loop connecting helices 3 and 4 of the POU-specific domain, is 4.3 Å from the thymine C₅ methyl group. To test the energetic significance of this putative contact, we measured binding of the Oct-1 POU domain to an octamer site in which Thy-5 was replaced with uracil. Binding of the POU domain to this uracil-substituted octamer was reduced approximately fourfold relative to binding the wild-type octamer, indicating that Leu-55 does interact with the methyl group of the thymine.

Our evidence for an energetic contribution from the
Figure 5. Overall structure of the Oct-1 POU domain–octamer complex (Klemm et al. 1994) illustrating the overlapping sub-sites near the center of the octamer sequence. The POU homeo domain is shown as a red ribbon and the POU-specific domain is shown as a yellow ribbon. The DNA is blue, but bases and phosphates contacted by the POU homeo domain have been highlighted with a red surface and those contacted by the POU-specific domain have been highlighted with a yellow surface. The contacts illustrated with these colored surfaces correspond to those indicated in Fig. 5 of Klemm et al. (1994), and we include the contact that Leu-55 of the POU-specific domain makes with Thy-5. Bases along one strand of the octamer sequence are labeled. Because there are no protein–protein contacts between the domains (Klemm et al. 1994), we believe that the overlapping base and phosphate contacts near the center of the site may mediate cooperative binding of the isolated domains. (The 24-amino-acid linker connecting the carboxyl terminus of the POU-specific domain with the amino terminus of the POU homeo domain is not shown because it is not ordered in the crystal structure and because it is not present when we assay cooperative binding of the isolated domains.)

of the isolated domains may involve the same energetic factors that cause the intact POU domain to prefer a contiguous octamer site. Both the POU domain and the isolated subdomains prefer a normal octamer site and there are similar energetic changes when this site is disrupted. Specifically, the binding constant of the intact POU domain is reduced by 2.1 kcal/mole when 2 bp are added near the center of the site. Corresponding studies with the isolated subdomains show that addition of these two base pairs near the center of the site eliminates the 1.6 kcal/mole cooperative interaction. We infer that cooperative interactions between the subdomains are the main reason that the POU domain prefers a contiguous octamer site.

Evaluation of our binding data in light of the known crystal structure of the Oct-1 POU domain–DNA complex suggests that the cooperativity may be mediated by overlapping DNA contacts near the center of the octamer site. Specifically, the crystal structure of the Oct-1 POU domain–DNA complex shows that there are no protein–protein contacts between the domains. Because the linker is flexible in the crystal and because our current experiments show that the isolated domains bind cooperatively, we have focused on ways that the cooperativity could be mediated via the DNA. In this regard, we note that there are a set of overlapping contacts near the center of the octamer site (Fig. 5). As described above, the A/T base pair at the fifth position of the octamer site (ATGCAAAT) is contacted in the major groove by a leucine from the POU-specific domain and in the minor groove by an arginine from the POU homeo domain. The overlap is even more striking when we also consider phosphate contacts: Both the POU-specific domain and the POU homeo domain contact phosphates that flank the A/T base pair at position 5. Given these extensive contacts, it certainly seems plausible that binding of the homeo domain could affect the conformation or flexibility of the DNA in this region and that subtle changes might facilitate binding of the POU-specific domain. Although we cannot rigorously rule out other mechanisms, our data suggest that cooperative binding by the POU-specific domain and the POU homeo domain is mediated via subtle changes in the DNA structure.

**Contribution of the linker to cooperative binding**

In the intact POU domain, the linker facilitates binding by a chelate effect that maintains a high local concentration of the domains. Covalent linkage between the POU-specific domain and the POU homeo domain reduces the overall binding reaction from a three-body problem to a two-body problem, and this reduces the entropic cost of complex formation. Now that we can separate out the other cooperative interactions, our new data allow us to quantitatively evaluate the contribution of the linker. Knowing the binding energies of the intact
POU domain and of the isolated subdomains allows us to describe the contribution of the linker in terms of an effective concentration. In simplest terms, the effective concentration provides a way of describing the chelate effect. Whenever one domain of this bipartite protein is bound to DNA, the linker will tether the other domain to a nearby region of space and hold it—at a high local concentration—near the DNA. A value for the effective concentration in this system can be calculated as:

\[ C_{\text{eff}} = \frac{(K_d)^{\text{POU}}(K_d)^{\text{POU,ln}}}{(K_d)^{\text{POU}}} \]

where \((K_d)^{\text{POU}}\) is the \(K_d\) of the POU-specific domain in the presence of the POU homeo domain \((1.7 \pm 0.88) \times 10^{-6} \text{ M}\), \((K_d)^{\text{POU,ln}}\) is \(1.5 \pm 0.17 \times 10^{-7} \text{ M}\), and \((K_d)^{\text{POU}}\) is \(7.1 \pm 1.9 \times 10^{-7} \text{ M}\) (see Creighton 1984, and Materials and methods for a further discussion of effective concentration). Substitution of the values obtained in our experiments shows that the effective concentration of either domain in an intermediate complex (when one half of the POU domain is bound to the DNA but the other is still free) is \(3.6 \times 10^{-3} \text{ M}\). Because this effective concentration is significantly larger than the binding constant of either domain, it appears that any intermediate complex will be thermodynamically unstable and can only be present as a transient intermediate during binding or release of the POU domain. The flexible linker clearly makes a significant contribution to binding.

**DNA binding by other POU domain proteins**

Our studies of cooperativity in Oct-1 POU domain–DNA interactions may provide a model system for analyzing cooperativity in other protein–DNA interactions, but the precise nature of these interactions will be specific to each system. In this regard, it is important to realize that there may be fundamental differences in cooperativity and subsite spacing with other POU domain proteins (Herr and Cleary 1995). Studies of DNA recognition by the POU domain proteins Brn-2, Tst-1, Unc-86, and Brn-3 have shown that they prefer to bind sites in which (1) the POU-specific domain subsite is inverted relative to its orientation in the octamer sequence \([\text{ATGCA} \rightarrow \text{TGCAT}]\) and (2) the POU-specific domain and POU homeo domain subsites are separated by two base pairs [for Brn-2 and Tst-1] or 3 bp [for Unc-86 and Brn-3] [Xue et al. 1992, Li et al. 1993].[Brn-2 also is capable of binding to half-sites spaced by 0 or 3 bp.] The different orientations and spacings of the subdomains in these other POU domain–DNA complexes presumably involve differences in the mechanisms of coupling between the subdomains. In each system, we expect that the covalent linkage between the POU-specific domain and the POU homeo domain will provide a favorable chelate effect, but the precise magnitude of this effect will depend on the length and sequence of the linker and the relative arrangement of the binding sites. Our studies suggest that cooperative binding of the isolated domains has provided another mechanism for the evolutionary diversification of the POU domains, and differences in these cooperative interactions may help us understand how the POU domain has adopted a range of roles in protein–DNA recognition and gene regulation.

**Relevance to other studies of protein–DNA interactions**

Studies of covalently linked DNA-binding domains involve a new level of complexity in understanding protein–DNA interaction, and our experimental strategies should be applicable to studies of other DNA-binding proteins that have two or more covalently linked domains. Corresponding studies of TFIIIA-like zinc fingers could be particularly interesting. The linkers clearly play an important role in recognition [isolated fingers do not bind effectively [Frankel et al. 1987]], but it would be very interesting to see whether there was any cooperativity left in the absence of a linker. In analogy to the POU domain, features of TFIIIA-type zinc finger–DNA complexes hint that DNA binding by adjacent fingers might be thermodynamically and structurally coupled. Crystal structures of these complexes show that each finger makes base contacts in a short 3- to 4-bp subsite, but there typically are overlapping phosphate contacts and there can be overlapping base contacts (Pavletich and Pabo 1991, 1993, Fairall et al. 1993). For instance, finger 4 of GLI contacts three phosphates in the finger 5 subsite [Pavletich and Pabo 1993]. In the case of the zinc fingers, there is also evidence for a hydrogen bond between neighboring fingers, usually involving an arginine in the \(\alpha\) helix of one finger and a carbonyl oxygen in the turn of a neighboring finger. Finally, an analysis of the DNA structure in zinc finger–DNA complexes reveals an enlarged major groove [Nekludova and Pabo 1994], and thus subtle changes in DNA structure may be involved in binding. Given these possibilities, it would be very interesting to remove one linker from a zinc finger protein and see whether the terminal \((n\text{th})\) finger would bind cooperatively in the presence of a polypeptide containing the other fingers 1 through \(n-1\). Measuring binding constants for these fragments would also give important information about the contribution of the linker to recognition.

Our data suggest that cooperativity in the Oct-1 POU domain–DNA interactions, which is seen in binding of the isolated domains, may be mediated by overlapping DNA contacts near the center of the octamer site and may involve subtle changes in the DNA structure. It is widely accepted that binding of a protein to DNA may affect the overall structure of its binding site [for review, see Travers 1992] but it has not been clear how these changes might affect the interactions of other proteins with neighboring sites. In some complexes, such as the CAP–DNA complex [Schultz et al. 1991] and the TBP–DNA complex [Kim et al. 1993a,b], the DNA double helix is drastically bent or distorted by the protein. In other protein–DNA complexes, however, the changes in DNA structure are less pronounced. In this report, we have
shown that the POU subdomains can bind cooperatively to overlapping subsites, in the absence of protein–protein interactions and in the absence of the normal covalent linker. We think this involves subtle, interdependent effects on the structure and/or flexibility of the DNA at the juncture of the subdomains. This type of indirect interaction between proteins bound to neighboring DNA sites may be one mechanism by which multiple transcription factors interact cooperatively in the combinatorial control of transcription. While many regulatory interactions have been shown to involve direct protein–protein contacts, the contribution from subtle changes in DNA structure has been much more difficult to assess. The Oct-1 POU domain provides an important model system for studying how DNA-mediated effects may lead to cooperativity in protein–DNA interactions.

Materials and methods

Protein constructs

The POU domain, the POU homeo domain and the POU-specific domain from Oct-1 were each expressed as carboxy-terminal fusions with glutathione S-transferase (GST). The POU domain and POU-specific domain constructs were gifts of Reggie Aurora and Winship Herr (Cold Spring Harbor Laboratory). Construction of the POU domain expression vector is described in Aurora and Herr (1992); the POU-specific domain vector is a derivative of this. The POU homeo domain expression vector was constructed by PCR amplification of appropriate region from the POU domain construct and cloning of the fragment into the BamHI and XhoI sites of pGEX-4T-2 (Pharmacia). To facilitate cleavage of the GST tag from the POU homeo domain, three glycine residues were included after the thrombin cleavage site (Guan and Dixon 1991). Therefore, the resulting protein has the five additional amino-terminal amino acids GSGGG. The construct was verified by dideoxy sequencing.

Soluble proteins were expressed, purified on glutathione–agarose (Sigma) and cleaved as described in Aurora and Herr (1992). The POU-specific domain was concentrated by use of Centricon and Microcon concentrators (Amicon). The POU domain was further purified on a DNA–cellulose column and concentrated as described previously [Klemm et al. 1994]. Further purification of the POU homeo domain was carried out by use of Sep-Pak C18 reversed-phase cartridges (Millipore). Sep-Paks were primed with 5 ml of acetonitrile plus 0.1% trifluoroacetic acid (TFA), then equilibrated with 5 ml of 4 M guanidine hydrochloride (GuHCl) in 25 mM Tris (pH 7.5). Then loaded onto an 18% 0.5x Tris-borate-EDTA (TBE) gel and electrophoresed at 300 V at 4°C. To detect a supershift, relatively high protein concentrations were necessary. The POU homeo domain concentration in this experiment was 5 x 10^-6 M and the POU-specific domain concentration was 1 x 10^-4 M.

Footprinting probe construction, purification, and labeling

Footprinting probes were prepared by use of synthetic oligonucleotides cloned into the EcoRI and BamHI restriction sites of pBluescript KS(+). The sequence of the top strand of the synthetic oligonucleotide is AACCTAAGGCCTATGCCCCATGCCCCAAGGCGTGAATTAAGGAGGCTTGCGGAGGGCGGCGGGCC (octamer sequence shown in bold). The sequence of the bottom strand of the synthetic oligonucleotide is: GATGCGGGCCCTCGGCACCCCCATGGGCTAGGCGCATGCCCCAAGGCGTGAATTAAGGAGGCTTGCGGAGGGCGGCGGGCC (octamer sequence shown in bold). The oligonucleotides for studies with altered subsite spacing are similar but have 2 bp inserted between the octamer half sites: ATGCAACAAAT. The synthetic duplexes were designed such that the 5’ and 3’ ends were ready for cloning into the EcoRI/BamHI-cut plasmid. Constructs were confirmed by dideoxy sequencing. DNA probes for footprinting were prepared by digestion of the appropriate plasmid with EcoRI and SalI, the resulting fragments are 115 [wild-type octamer] or 117 [altered subsite spacing] bp in length. After restriction digestion, the fragments for footprinting were purified by electrophoresis through 2% low melting point agarose gels. Excised bands were digested with Gelase (Epicentre Technologies) according to the manufacturer’s protocol. The probe was radiolabeled by filling in the EcoRI 5’ overhang by use of the Klenow enzyme (Boehringer Mannheim) in the presence of [α-32P]dATP and [α-32P]dCTP.

DNase I footprinting

Binding reactions were carried out at room temperature in a 200-μl volume with a buffer containing 8 mM HEPES (pH 7.8), 100 mM NaCl, 0.03% NP-40, 5 mM MgCl2, 1 μg/ml of BSA, 1% Ficoll, and the appropriate amount of protein. The DNA probe was present at concentrations of less than 5 pm. For measuring POU domain Kd values, reactions were incubated at room temperature for 30 min and loaded onto 10% 0.5x Tris-borate–EDTA (TBE) gels and electrophoresed at 150 V. The dissociation constants measured for the POU domain on the wild-type octamer and on an octamer in which 2 bp have been inserted in the center of the octamer are as follows: wild-type octamer: 7.1 ± (±0.9) x 10^-11 M octamer with 2-bp insert: 2.3 ± (±0.37) x 10^-9 M. The DNA probe used in the experiment with the separated subites is analogous to the wild-type probe (above) but has the sequence ATGCCAACAAAT.

For supershift assays with the POU homeo domain and the POU-specific domain, reactions were incubated on ice for 30 min, then loaded onto an 18% 0.5x TBE gel and electrophoresed at 300 V at 4°C. To detect a supershift, relatively high protein concentrations were necessary. The POU homeo domain concentration in this experiment was 5 x 10^-6 M and the POU-specific domain concentration was 1 x 10^-4 M.
100 mM NaCl, 1 mM dithiothreitol, 0.1 μg/ml of BSA, 0.03% NP-40, 5 mM MgCl₂, 1 mM CaCl₂ and 20,000 cpm of probe DNA. The concentration of probe DNA in the reactions was <0.1 nm. After a 30-min incubation, 40 μl of 2.5 ng/μl of DNase I (Worthington) was added and the reaction was left at room temperature for 1 min. The reaction was stopped by addition of 200 μl of a buffer containing 2.5 M NH₄COOH, 1 μg/ml of sonicated salmon sperm DNA, and 20 mM EDTA. The DNase I-treated probe was purified by extraction with phenol and chloroform and then precipitated with ethanol in the presence of 20% NP-40, 5 mM MgCl₂, 1 mM CaCl₂ and 20,000 cpm of probe DNA. The concentration of probe DNA in the reactions was roform and then precipitated with ethanol in the presence of 20% NP-40, 5 mM MgCl₂, 1 mM CaCl₂ and 20,000 cpm of probe DNA.

Quantitation of dissociation constants

All gels were quantitated by use of a Molecular Dynamics 492I PhosphorImager. For quantitation of footprinting gels, fractional protection of a site [θ] was calculated as

\[ \theta = \frac{I_{\text{bound}}}{I_{\text{free}} + I_{\text{bound}}} \]

where \( I_{\text{site}} \) is the integrated intensity within a defined region, \( I_{\text{std}} \) is the integrated intensity of a standard region on the DNA fragment, \( I_{\text{ref,site}} \) and \( I_{\text{ref, std}} \) are analogous parameters derived from a reference lane with no protein present [Brenowitz et al. 1993]. The fractional protection was assumed to represent the fraction [θ] of bound DNA. For gel shift experiments, θ was calculated as

\[ \theta = \frac{I_{\text{bound}}}{I_{\text{free}} + I_{\text{bound}}} \]

Equilibrium dissociation constants \( [K_d] \) were determined by linear regression using the Scatchard equation

\[ \frac{[P]}{[P]} = K_d - \frac{[P]}{K_d} \]

where \([P]\) represents the free protein concentration. For \( K_d \) calculations, we used points for which 0.2 ≤ θ ≤ 0.8. Because the total DNA concentration used in our binding experiments was always well below the \( K_d \), the free protein concentration and the total protein concentration are essentially equal. Reported \( K_d \) values were derived from the slope of the Scatchard plot and represent the mean of at least of three independent experiments.

Effective concentration

We have used the term effective concentration to describe the contribution the flexible linker makes to cooperative binding of the POU subdomains. This provides a quantitative way of describing the chelate effect. In the calculation of effective concentration described in the text, we use the dissociation constant for the POU-specific domain that is observed in the presence of the homeo domain, \( [K_d]P\text{OU} \). Identical results would be obtained by calculating

\[ C_{\text{eff}} = \frac{|K_dP\text{OU}|[K_dP\text{OU}]_\text{std}}{K_dP\text{OU}} \]

where

\[ K_{\text{cooperativity}} = e^{-\frac{\Delta G_{\text{cooperativity}}}{RT}} \]

Thinking of the effective concentration that we have calculated (3.6 nm) in terms of the volume accessible to a single molecule shows that this corresponds to having the free domain constrained to a spherical volume of radius 48.1 Å. A rigorous geometric and energetic interpretation of the effective concentration would be extremely difficult, but we note that this radius falls well within the distance that could be spanned by the 24-amino-acid Oct-1 POU domain linker (which might cover 24 × 3.5 Å = 84 Å in a fully extended linker).

Acknowledgments

We thank Winship Herr and Reggie Aurora for the Oct-1 POU domain and POU-specific domain expression plasmids; Tracy Smith for advice on DNase I footprinting; Ernest Fraenkel and Harvey Greisman for critical reading of the manuscript; and Robert Sauer, Phillip Sharp, and Karyn Cepek for helpful discussions. This work was supported by National Institutes of Health grant GM31471, and we used equipment purchased with support from the Howard Hughes Medical Institute. C.O.P. is in the Howard Hughes Medical Institute.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Aurora, R. and W. Herr. 1992. Segments of the POU domain influence one another's DNA-binding specificity. Mol. Cell. Biol. 12: 455-467.

Beamer, L.J. and C.O. Pabo. 1992. Refined 1.8 Å crystal structure of the λ repressor-operator complex. J. Mol. Biol. 227: 177–196.

Brenowitz, M., D. Senear, E. Jamison, and D. Dalma-Weiszhausz. 1993. Quantitative DNase I footprinting. In Footprinting of nucleic acid-protein complexes [ed. A. Revzin], pp. 1–43. Academic Press, San Diego, CA.

Creighton, T.E. 1984. Proteins: Structures and molecular principles, pp. 360–365. W.H. Freeman and Company, New York, NY.

Fairall, L., J.W.R. Schwabe, L. Chapman, J.T. Finch, and D. Rhodes. 1993. The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition. Nature 366: 483–487.

Frankel, A.D., J.M. Berg, and C.O. Pabo. 1987. Metal-dependent folding of a single zinc finger from transcription factor IIIA. Proc. Natl. Acad. Sci. USA 84: 4841–4845.

Guan, K. and J.E. Dixon. 1991. Eukaryotic proteins expressed in Escherichia coli: An improved thrombin cleavage and purification procedure of fusion proteins with glutathione S-transferase. Anal. Biochem. 182: 262–267.

Herr, W. and M.A. Cleary. 1995. The POU domain: Versatility in transcriptional regulation by a flexible two-in-one DNA-binding domain. Genes Dev. 9: 1679–1693.

Jordan, S.R. and C.O. Pabo. 1988. Structure of the lambda co-
plex at 2.5 Å resolution: Details of the repressor-operator interactions. Science 242: 893–899.

Kim, J.L., D.B. Nikolov, and S.K. Burley. 1993a. Co-crystal structure of TBP recognizing the minor groove of a TATA element. Nature 365: 520–527.

Kim, Y., J.H. Geiger, S. Hahn, and P.B. Sigler. 1993b. Crystal structure of a yeast TBP/TATA-box complex. Nature 365: 512–520.

Klemm, J.D., M.A. Rould, R. Aurora, W. Herr, and C.O. Pabo. 1994. Crystal structure of the Oct-1 POU domain bound an octamer site: DNA recognition with tethered DNA-binding modules. Cell 77: 21–32.

Li, P., X. He, M.R. Gerrero, M. Mok, A. Aggarwal, and M.G. Rosenfeld. 1993. Spacing and orientation of bipartite DNA-binding motifs as potential functional determinants for POU domain factors. Genes & Dev. 7: 2483–2496.

Nekludova, L. and C.O. Pabo. 1994. Distinctive DNA conformation with enlarged major groove is found in Zn-finger–DNA and other protein–DNA complexes. Proc. Natl. Acad. Sci. 91: 6948–6952.

Pavletich, N.P. and C.O. Pabo. 1991. Zinc finger–DNA recognition: Crystal structure of a zif268-DNA complex at 2.1 Å. Science 252: 809–817.

——. 1993. Crystal structure of a five-finger GLI-DNA complex: New perspectives on zinc fingers. Science 261: 1701–1707.

Schultz, S.C., G.C. Shields, and T.A. Steitz. 1991. Crystal structure of a CAP–DNA complex: The DNA is bent by 90°. Science 253: 1001–1007.

Travers, A.A. 1992. DNA conformation and configuration in protein–DNA complexes. Curr. Opin. Struct. Biol. 2: 71–77.

Verrrriter, C.P., M.J. Alkema, W.W. van Wepener, H.C. van Leeuwen, M.J. Strating, and P.C. van der Vliet. 1992. The DNA binding specificity of the bipartite POU domain and its subdomains. EMBO J. 11: 4993–5003.

Xue, D., M. Finney, G. Ruvkin, and M. Chalfie. 1992. Regulation of the Mec-3 gene by the C. elegans homeoproteins unc-86 and Mec-3. EMBO J. 11: 4969–4979.
Oct-1 POU domain-DNA interactions: cooperative binding of isolated subdomains and effects of covalent linkage.

J D Klemm and C O Pabo

Genes Dev. 1996, 10:
Access the most recent version at doi:10.1101/gad.10.1.27