MOLECULAR AND CELLULAR BIOLOGY, Feb. 1992, p. 455–467
0270-7306/92/020455-13$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 12, No. 2

Segments of the POU Domain Influence One Another's DNA-Binding Specificity†

RAJEEV AURORA[1,2] AND WINSHIP HERR[1]*

Cold Spring Harbor Laboratory, P.O. Box 100, 1 Bungtown Road, Cold Spring Harbor, New York 11724-2206,[1] and Department of Molecular Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794[2]

Received 5 September 1991/Accepted 21 October 1991

The ubiquitously expressed mammalian POU-domain protein Oct-1 specifically recognizes two classes of *cis*-acting regulatory elements that bear little sequence similarity, the octamer motif ATGCAAAT and the TAATGARAT motif. The related pituitary-specific POU protein Pit-1 also recognizes these two motifs but, unlike Oct-1, binds preferentially to the TAATGARAT motif. Yet in our assay, Pit-1 still binds octamer elements better than does the octamer motif-binding protein Oct-3. The POU domain is responsible for recognizing these diverse regulatory sequences through multiple DNA contacts that include the two POU subdomains, the POU-specific region, and the POU homeodomain. The DNA-binding properties of 10 chimeric POU domains, in which different POU-domain segments are derived from either Oct-1 or Pit-1, reveal a high degree of structural plasticity; these hybrid proteins all bind DNA well and frequently bind particular sites better than does either of the parental POU domains. In these chimeric POU domains, the POU-specific A and B boxes and the hypervariable POU linker can influence DNA-binding specificity. The surprising result is that the influence a particular segment has on DNA-binding specificity can be greatly affected by the origin of other segments of the POU domain and the sequence of the binding site. Thus, the broad but selective DNA-binding specificity of Oct-1 is conferred both by multiple DNA contacts and by dynamic interactions within the DNA-bound POU domain.

Transcriptional regulation relies on sequence-specific interactions between transactivator proteins and *cis*-acting elements. In contrast to many prokaryotic transcriptional regulators, eukaryotic DNA-binding transcription factors frequently display the potential to bind to diverse sets of DNA sequences. As examples, the yeast transcription factor HAP-1, a zinc finger protein, and the mammalian transcription factors C/EBP, a leucine zipper protein, and Oct-1, a POU-domain protein, all display the ability to bind to very different sequences (reviewed in reference 18). The ability of site-specific transactivators to recognize diverse DNA sequences is important for transcriptional regulation because it permits greater flexibility in promoter structure and transactivator function, but relatively little is known about the mechanisms by which this flexibility is brought about. In this study, we have examined how Oct-1 recognizes a set of functional *cis*-acting elements differing greatly in DNA sequence.

Oct-1 (also referred to as NF-A1, OTF-1, NFIII, and OBP100) was originally discovered (44) because it recognizes the highly conserved octamer motif ATGCAAAT (33), which is found in both ubiquitously active and lymphoid-specific promoters (reviewed in reference 39) and which is also referred to as the decamer motif because of a conserved A residue in position 10 (ATGCAAATNA [10]). Oct-1, however, can also bind on its own to sequences that bear little resemblance to an octamer motif (1, 2, 20, 48). For example, Oct-1 can bind to the TAATGARAT (R=purine) motif, the *cis* target for transactivation by the herpes simplex virus (HSV) transactivator of immediate-early (IE) promoters called VP16 (Vmw65, VF65, α-TIF) (reviewed in reference 12), even though in some cases this motif bears little resemblance to an octamer motif (1, 2). A study of the ability of Oct-1 to bind to very dissimilar DNA sequences, including two Oct-1 binding sites within the simian virus (SV40) enhancer, called sites I and II, suggested that few, if any, DNA residues are obligatory for Oct-1 to recognize DNA but rather that it is the sum of many individual Oct-1–DNA contacts that results in effective binding (2).

How does Oct-1 realize such flexibility in DNA sequence recognition? Oct-1 binds DNA through its POU domain; the POU domain is a bipartite structure of 155 to 162 amino acids that contains two conserved motifs, a C-terminal POU-type homeodomain and an N-terminal POU-specific region separated by a 15- to 27-amino-acid hypervariable linker (14). The 75- to 82-amino-acid POU-specific region was further subdivided into two smaller segments called A and B based on the pattern of sequence conservation among the four founding members of the POU family: the ubiquitous and lymphoid octamer-binding proteins Oct-1 and Oct-2 (NF-A2, OTF-2), the pituitary transcription factor Pit-1 (GHF-1), and the product of the nematode gene *unc-86* (14). At least four other complete POU domains have since been described (5, 19, 25, 31, 37, 42, 51; see also reference 13). To date, the POU-specific and POU-homeodomain motifs have always been found associated with one another as a complete POU domain, suggesting a functional link between the two motifs.

Consistent with a functional link, both the POU-specific region and POU homeodomain are involved in DNA binding (17, 22, 50, 57). The homeodomain itself is a member of the helix-turn-helix class of DNA-binding motifs (21, 24, 32, 36, 43), and at relatively high concentrations, a POU homeodomain alone can bind to DNA (17, 22, 57). In contrast, little is known about the structure of the POU-specific region, and DNA binding by the POU-specific region alone has not been demonstrated.

* Corresponding author.
† Dedicated to the memory of our colleague William Clouston.

Mutational analyses of the POU domain have, nevertheless, revealed an involvement of the POU-specific region in DNA binding. Point mutations in the Oct-1 homeodomain, or in either the A or B segment of the POU-specific region, severely affect DNA binding, whereas a six-amino-acid insertion within the linker has little or no effect on DNA-binding activity (50). Deletion of the Oct-1 POU-specific region (50, 56, 57) and extensive mutagenesis of the Pit-1 POU-specific region, including an Oct-1/Pit-1 POU-specific region exchange (17), also revealed a critical role of this region in high-affinity sequence-specific DNA binding. Interestingly, although Pit-1 exists as a monomer in solution, it binds to its DNA response element as a dimer, exhibiting DNA dependent protein-protein interactions that also require the POU-specific region (17). The relative contributions of the Pit-1 POU-specific region to sequence-specific binding through interprotein dimer contacts and through direct protein-DNA contacts have not been determined.

To understand the mechanism of divergent DNA sequence recognition by Oct-1, we have assayed the relative contributions of the different regions of the Oct-1 POU domain for binding as a monomer to a series of different octamer- and TAATGARAT-related DNA-binding sites. We took advantage of the different DNA-binding specificities of the Pit-1 POU domain by constructing a series of Oct-1/Pit-1 POU-domain chimeras and examining their DNA-binding specificities. Our studies show that through DNA contacts that involve both the POU-specific and POU-homeodomain regions, the different segments of the POU domain contribute to the DNA-binding specificity, but their influence is dependent on both the origin of the other segments within the POU domain and the sequence of the binding site.

MATERIALS AND METHODS

Expression constructs. The Pit-1 expression construct for in vitro translation was made by using Pit-1-specific primers that contained *Xba*I and *Bam*HI sites (5′ primer, GAGTC TAGAATGAGTTGCCAACCTTTC; 3′ primer, CTGGGAT CCTTATCTGCACTCAAGATG) to amplify, by the polymerase chain reaction (PCR) (27), the *pit-1* (4, 16) sequences from cDNAs prepared with total GH3 cell RNA and cloning the amplified fragment, digested with *Xba*I and *Bam*HI, into pBSM13$^+$ to produce pBS*pit-1*. To create the POU-domain exchange chimeras, a silent mutation was introduced in pBS$^{B\text{-}X}$-*oct*-1(2)1(3)1 (46) in the sequences encoding the amino acids PSLEE, upstream of the Oct-1 POU domain (see Fig. 6), to create a *Hin*dIII site (CCa AGC TTG; the mutation is shown in lowercase) by site-directed mutagenesis, thus creating pBS$^{B\text{-}X}$-*oct*-1(H)1(2)1(3)1. The designations (2) and (3) represent silent *Xho*I and *Sal*I sites (52), respectively, that were not made use of here. The Oct-1 POU-domain sequences can be excised from pBS$^{B\text{-}X}$-*oct*-1(H)1(2)1(3)1 by digestion with *Hin*dIII and *Pfl*MI at a naturally occurring site immediately downstream of the Oct-1 POU domain. To exchange POU domains, two primers, an upstream primer containing a *Hin*dIII site and a downstream primer containing the *Pfl*MI site, were used to amplify the regions corresponding to the POU domains of Oct-3 and Pit-1 by PCR. The Oct-3 POU-domain sequence was amplified from cDNA prepared from F9 cell RNA, using the *Hin*dIII oligonucleotide CCAAGCTTGGAGGAGTCCC AGGAC and the *Pfl*MI oligonucleotide AACCACTGCTTG GTGGAATACTTGATCTTTT. The Pit-1 POU domain was amplified from pBS*pit-1* by using the *Hin*dIII oligonucleotide AAAAGCTTGGAAGAGCCAATAG and the *Pfl*MI oligonucleotide AACCACTGCTTGGCGTTTTCACCCGTTTT. The amplified fragments were digested with *Hin*dIII and *Pfl*MI and ligated into the large fragment of *Hin*dIII-*Pfl*MI-digested pBS$^{B\text{-}X}$-*oct*-1(H)1(2)1(3)1 to create pBS*oct*-1(H)3(Pf)1 (for Oct-1.3.1) and pBS*oct*-1(H)P(Pf)1 (for Oct-1.P.1), where Pf signifies the natural Oct-1 *Pfl*MI site used to recombine fragments. *oct*-1(H)P(Pf)1 differs from *oct*-1(2)P(3)1 described previously (46) by containing the entire Pit-1 POU domain. The sequences encoding the entire POU domains of Oct-3 and Pit-1 were sequenced by the dideoxy method (38). The templates for expression of Oct-2 (pCG*oct*-2) and Oct-1.2.1 [pCG*oct*-1(2)2(3)1] have been described previously (52).

Oct-1/Pit-1 chimeras. The precise segment swaps between the Oct-1 and Pit-1 POU domains in Oct-1.1.1 and Oct-1.P.1 were prepared by fusing the two POU domains in tandem and then deleting intervening sequences to generate the precise POU-domain recombinant. For example, to prepare an Oct-1.1.1 construct with the Pit-1 POU-specific A box (called Oct-1.1.1[A_p]), we placed the N terminus of Oct-1 together with the Pit-1 POU-domain sequences upstream of the Oct-1 POU domain and C terminus in the construct Oct-1.(P.1).1. Then we precisely deleted the sequences between the junction of the Pit-1 A and B boxes (see Fig. 6) and the junction of the downstream Oct-1 A and B boxes by oligonucleotide-directed deletion (9) with T7 DNA polymerase (3). The origin of the segments in the chimeras is indicated by using a subscript o for Oct-1 and a subscript p for Pit-1. The chimeras A_o, AB_o, and H_p were made with Oct-1.(1.P).1, and A_p, AB_p, and H_o were prepared with Oct-1.(P.1).1. To make the chimeras B_o, B_p, L_o, and L_p, a second round of tandem POU-domain constructs was prepared to create 1.(P.AB_o).1, for [B_o]; 1.(P.ABL_o).1, for [L_o]; 1.(1.AB_p).1, for [B_p]; and 1.(1.ABL_p).1, for [L_p]. The oligonucleotide primers to make in-frame deletions were as follows (the vertical bar indicates the POU-segment junction):

A_o/B_p: GATGTTGGGCTCGCT|CTGGCCGCTGTCCAC
A_p/B_o: AACGTGGGCGAAGCT|ATGGGGAAACTATATGG
AB_o/L_p: CTAAATGATGCAGAG|CAGGTCGGAGCTTTG
AB_p/L_o: CTGGAGGAAGCTGAG|AACCTCTCATCTGATT
ABL_p/H_o: GGAGCAAACGAAAGG|AGGAGGAAGAAACGC
ABL_o/H_p: GAGGGCTTGAGCCGT|AAGAGGAAACGGAGG

Preparation of in vitro-translated proteins. All chimeras and POU swap-encoding DNA templates for SP6 RNA polymerase were prepared by PCR amplification using a pair of primers. The upstream primer (SP65′Oct-1; CTATTTAG GTGACACTATAGAAACAGACACCATGGACAATCCG TCAGAAACC) provided the SP6 promoter for transcription (positions 3 to 20), β-globin 5′ untranslated sequences, initiation codon, and 3′ G residue for translation (positions 21 to 35) and 17 nucleotides that hybridize immediately downstream of the Oct-1 initiation codon described previously (49) for amplification (positions 36 to 52). The downstream primer (OCT1.10; GGCTTCTGGCAGCCCAGC) hybridizes immediately downstream of the stop codon of the *oct-1* coding sequences. For synthesis of the parental proteins Pit-1 and Oct-3, the T7 promoter of pBS and T3 promoter of pBlueScriptOct-3.32 (37), respectively, were used to direct RNA synthesis. Oct-2 mRNA was synthesized as described previously (52). The RNA from each template was used to program reticulocyte lysates (Promega Inc.) according to the manufacturer's instructions. The extent and quality of synthesis were monitored by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (23).

The proteins corresponding to full-length product were excised from dried gels, and the yield was measured by scintillation counting. The molar concentration of the proteins was adjusted by dilution with an identically treated but unprogrammed reticulocyte lysate. The number of methionines in each protein was taken into account as follows: Oct-1 (1.1.1), 12 methionines; Oct-2 (2.2.2), 11 methionines; Oct-1.2.1, 12 methionines; Oct-3 (3.3.3), 5 methionines; Oct-1.3.1, 11 methionines; Pit-1 (P.P.P), 7 methionines; and Oct-1.P.1, 11 methionines.

The amino-terminal Oct-1 truncation product used for Fig. 1 was generated by using an oligonucleotide primer (SP65'OcPOU; CTATTTAGGTGACACTATAGAAACAG ACACCATGGAGGAGCCCAG) that contains promoter sequences for SP6 RNA polymerase, the β-globin 5′ untranslated region, and a sequence that hybridizes to the 5′ end of the POU domain (junction 1 in Fig. 6) and the primer OCT1.10 to make a template for SP6 RNA polymerase by PCR. The carboxyl-terminal truncation used the same 5′ primer as was used to make full-length Oct-1 and a second primer that provides a stop codon and hybridizes to the *Pfl*MI site at the end of sequences corresponding to the Oct-1 POU domain (junction 5 in Fig. 6) (OcHBam; CAGGATC CTATGGGTTGATTCTTTTTC). The resulting constructs were transcribed by SP6 RNA polymerase, and the resulting RNAs were used to program rabbit reticulocyte lysates as described above. A HeLa cell nuclear extract was prepared according to method of Wildeman et al. (58).

Electrophoretic mobility retardation assay probes. All probes used result from synthetic oligonucleotides that were cloned into sites within the pUC119 polylinker. The sequences are shown in Fig. 2. Some of the sites (see Results for definitions) have been described previously as follows: ch. H2B, SV40 site I, SV40 site II, SV40 site II/Perf. Octa., and ICP4 TAAT-2 (2); and ICP0 (46). The sphII/II site was prepared by digestion of pβ6XsphII/sphII (30) with *Xho*I, end repair of the resulting 28-bp *Xho*I fragments, and ligation into the *Sma*I site of pUC119. The hu. H2B, mu. hu. H2B, ICP4 TAAT-1, mu. ICP4 TAAT-1, and prolactin sites were similarly prepared from pβ6X constructs that will be described elsewhere (6, 53) except that the first four sites listed were cloned into the pUC119 *Hin*cII site. The ICP4 TAAT-1 site is from −367 to −350 upstream of the ICP4 initiation site (34).

The DNA probes for gel retardation were prepared in parallel by PCR amplification, using the same pair of radiolabelled forward and reverse sequencing primers prepared by phosphorylation with polynucleotide kinase in the presence of [γ-^{32}P]ATP (7,000 Ci/mmol; ICN Radiochemicals). The products were gel purified by electrophoresis through a 6% polyacrylamide gel and eluted overnight at 4°C in 10 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES; pH 7.9)–60 mM KCl.

Electrophoretic mobility retardation assay. In vitro-translated proteins (3 μl; equal to 10^{-14} mol of protein, assuming a 5 μM endogenous concentration of methionine in the translation extracts prepared by Promega) were preincubated with the competitor DNAs and buffer for 15 min on ice prior to the addition of 5×10^4 cpm (2.5 μl) of probe and further incubation on ice for 15 min. The final reaction volume was 10 μl in 8 mM HEPES (pH 7.9)–60 mM KCl–2 mM EDTA–0.2 mM dithiothreitol–4 mM spermidine–0.1 mg of bovine serum albumin per ml–0.03% Nonidet P-40–2 to 3 μg of poly(dI-dC) · (dI-dC)–0.5% Ficoll–10% glycerol–500 ng of single-stranded pUC119. At the end of the incubation, all reactions were transferred to room temperature and loaded onto a Tris-glycine 5% polyacrylamide gel (acrylamide/bisacrylamide, 29:1) (48) which had been preelectrophoresed for 1.5 h. DNA-binding activity was found to be sensitive to the binding conditions. The conditions described above were used to maximize the difference between the DNA-binding activities of Oct-1 and Pit-1.

POU domain-DNA UV cross-linking. Wild-type POU domain- and POU-X_A–glutathione-*S*-transferase (GST) fusion proteins were expressed in *Escherichia coli* by using the T7 expression system (pET11c) (47). The factor X_A site was introduced into the Oct-1 linker segment by site-directed mutagenesis between the Pro-Ser/Ala-Leu sequence in the Oct-1 linker (see Fig. 6), using single-stranded templates and the primer TCCAGCCCAAGTATCCAAGGGCGGGCCCT GAATTCTCCA (the sequence insertion is underlined) as described above. The GST cassette was obtained by PCR amplification using primers that contained *Nde*I and *Xba*I sites from pGEX2-T. The sequences containing wild-type POU-domain and POU-X_A fragments were also obtained by PCR amplification of the appropriate constructs, using primers that contained *Xba*I and *Bam*HI sites. The GST cassette and sequences containing the POU domain were digested with the appropriate enzymes and cloned into pET11c digested by *Nde*I and *Bam*HI, in a three-fragment ligation. The constructs were verified by dideoxy sequencing. The constructs were transformed into *E. coli* BL21(DE3), kindly provided by F. W. Studier. Soluble proteins were expressed as described previously (47) except that cells were grown at 23°C. After 3 to 6 h of isopropyl-β-D-thiogalactopyranoside (IPTG) induction, the cells were harvested by centrifugation and resuspended in phosphate-buffered saline. Following lysozyme treatment at 30°C, Nonidet P-40 (Sigma) was added to a final concentration of 1%. The cells were lysed by three cycles of freeze-thawing. All subsequent steps were performed at 4°C. The lysate was sonicated five times with 10-s pulses (60% duty cycle), using a Branson tip sonicator, and subsequently cleared by centrifugation at 14,000 × *g* for 15 min. The proteins were batchwise affinity purified from the supernatant on glutathione-agarose (Sigma) and eluted with 5 mM reduced glutathione. Approximately 500 μg of 90 to 95% pure protein (as judged by Coomassie staining after SDS-PAGE) were obtained from 50 ml of induced culture.

The bromodeoxyuridine (BrdU)-substituted DNA probes were prepared by primer extension using the M13 forward sequencing primer. The annealed single-stranded template containing the human H2B octamer motif was extended, in the presence of 62.5 μM α-^{32}P-labelled deoxynucleoside triphosphates (800 Ci/mmol) with unlabelled BrdU (Sigma) substituted for TTP, by the Klenow fragment of DNA polymerase I (U.S. Biochemical). This protocol labelled the strand shown in Fig. 2. For cross-linking, the purified proteins (200 nM, final concentration) were preincubated in 43.5 μl of 8 mM HEPES (pH 7.9)–60 mM KCl–1 mM EDTA–0.2 mM dithiothreitol–4% Ficoll–0.1% Nonidet P-40–130 ng of poly(dI-dC) (final concentration) in a 96-well microtiter dish on ice for 15 min. The probe (50 nM) in 6.5 μl was added, and the mixture was incubated on ice for 5 min and then exposed to UV light in a Stratalinker (Stratagene Inc.) for 10 min on ice. The reaction mixture was then treated with DNase I (5 U) and micrococcal nuclease (1 U) for 30 min and repurified on glutathione-agarose, and half of the reaction mixture was treated with factor X_A (Boehringer Mannheim) at 23°C for 2.5 h. Laemmli buffer was added, and the mixture was boiled and loaded on a 12.5% polyacrylamide gel (21). After electrophoresis, the gels were fixed, dried on 3MM paper, and autoradiographed. ^{14}C-labeled

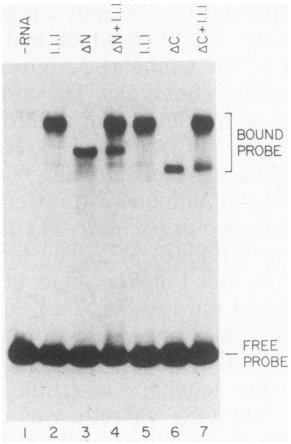

FIG. 1. Evidence that Oct-1 binds an octamer motif as a monomer. Oct-1 (1.1.1) and N-terminally (ΔN) and C-terminally (ΔC) truncated forms of Oct-1 were translated in vitro, either alone or in combination, and assayed by electrophoretic mobility retardation of a human histone H2B octamer motif probe as indicated above the lanes. The unprogrammed reticulocyte lysate exhibits low levels of endogenous Oct-1 activity (lane 1).

molecular weight standards were obtained from Bethesda Research Laboratories.

RESULTS

Experimental strategy. To define the influence of regions within the Oct-1 POU domain on DNA sequence recognition, we exchanged regions of the Oct-1 POU domain for homologous regions of other POU-domain proteins and assayed the DNA-binding specificity of the chimeric proteins. To assess a true change in DNA-binding specificity, as opposed to a general change in DNA-binding affinity, the DNA-binding activities of the parent and the chimeras were compared on numerous DNA-binding sites. A difference in DNA-binding specificity is best indicated if the binding-site preference is actually reversed between two different DNA-binding proteins. To examine POU domain-DNA interactions in the absence of the protein-protein interactions involved in dimer formation, we analyzed the DNA-binding activities of the chimeric proteins on sites to which they bind as monomers. We first compared the DNA-binding specificity of Oct-1 with those of Oct-1 chimeras carrying the entire POU domain from three related POU proteins, Oct-2 (7, 26, 40), Oct-3 (31, 37; also referred to as Oct-4 [42]), and Pit-1 (16; also referred to as GHF-1 [4]), and with the DNA-binding specificities of the three related POU proteins themselves. Then we tested the DNA-binding specificity of a series of Oct-1/Pit-1 chimeras in which only individual segments of the POU domain were exchanged. To assay the DNA-binding properties of the various wild-type and hybrid proteins, an equimolar concentration of each protein, synthesized in vitro, was assayed on probes of identical specific activity by the electrophoretic mobility retardation assay (see Materials and Methods for details).

Lack of heterodimer formation suggests that Oct-1 binds an octamer motif as a monomer. Figure 1 shows an electrophoretic mobility retardation assay for heterodimer formation between differently sized Oct-1 proteins, performed by the method of Hope and Struhl (15). Oct-1 (1.1.1) and the N-terminally (ΔN) and C-terminally (ΔC) truncated Oct-1 proteins were synthesized in vitro, either alone or together by cotranslation, and assayed with a perfect octamer motif probe derived from the human histone H2B promoter. The unprogrammed reticulocyte lysate (lane 1) contains a small amount of endogenous 95- to 100-kDa Oct-1 that generates an Oct-1–DNA complex that migrates somewhat more slowly than the complex generated by the in vitro-translated 90-kDa Oct-1 protein (lane 2), which lacks some N-terminal sequences (8). As expected, the Oct-1 ΔN (lane 3) and ΔC (lane 6) truncated proteins generate more rapidly migrating Oct-1–DNA complexes. Cotranslation of Oct-1 with the N-terminally (lane 4) or C-terminally (lane 7) truncated Oct-1 proteins did not yield any new complexes of intermediate mobility that could indicate formation of heterodimers. If heterodimers did arise freely, the intermediate-mobility complex resulting from heterodimer formation should be a prevalent complex in lanes 4 and 7. This assay cannot absolutely establish monomer binding because each truncation may remove a dimerization domain that is not essential for DNA binding but results in the preferential formation of full-length homodimers over heterodimers. We specifically used the N- and C-terminal truncations, which alternately retain the C or N terminus, to uncover such a dimerization domain. The lack of heterodimer formation by either of these two truncations strongly suggests that Oct-1 binds to a single octamer motif as a monomer. Therefore, in all likelihood, in the following experiments only complex formation between a single POU domain and DNA is being assayed.

The POU domain is the major determinant of DNA-binding specificity. The DNA-binding activity of chimeric POU-domain proteins was assayed on a series of 11 octamer- and TAATGARAT-related binding sites, as well as a Pit-1 binding site from the rat prolactin promoter (28). The sequence of each site is shown in Fig. 2. The sites have been aligned such that there is maximal similarity between adjacent sequences (2). HSV carries two types of TAATGARAT motifs: class I (OCTA$^+$)TAATGARAT motifs carry an overlapping imperfect octamer motif (ATGCTAATGARAT), whereas class II (OCTA$^-$)TAATGARAT motifs lack the overlapping octamer motif (GCGGTAATGARAT). In Fig. 2, a class I (OCTA$^+$)TAATGARAT motif from the ICP0 (IE110, α0) promoter is shown at the top, followed by sites that are more octamer related and then sites that are more class II (OCTA$^-$)TAATGARAT related. The octamer-related sequences include (i) natural octamer sites in the human (hu. H2B) and chicken (ch. H2B) histone H2B promoters and (ii) the two divergent SV40 enhancer sites I (seven of eight match) and II (five of eight match). Three different point mutants of these sites were also included: a single point mutation of SV40 site I that does not prevent Oct-1 binding (sphII/II), a double point mutation of the human H2B octamer motif that debilitates Oct-1 binding (mu. hu. H2B), and a triple point mutation of SV40 site II that converts the divergent octamer motif to a perfect octamer motif (site II/Perf. Octa.). The two HSV-derived class II (OCTA$^-$) TAATGARAT motifs (ICP4 TAAT-1 and TAAT-2) are from the ICP4 (IE175, α4) promoter and are shown in the inverted orientation, because this orientation gives the best alignment between an octamer motif, the divergent SV40 site II octamer-related sequence, and a TAATGARAT motif (2). The best alignment of SV40 site II and the (OCTA$^-$)TAATGARAT motif is achieved by deletion of a single central base pair, as indicated by the boldface type in Fig. 2. In these studies, a double point mutant of the ICP4 TAAT-1 site was also included.

Figure 3 shows the results of assaying the Oct-1 chimeras carrying the Oct-2, Oct-3, and Pit-1 POU domains on six

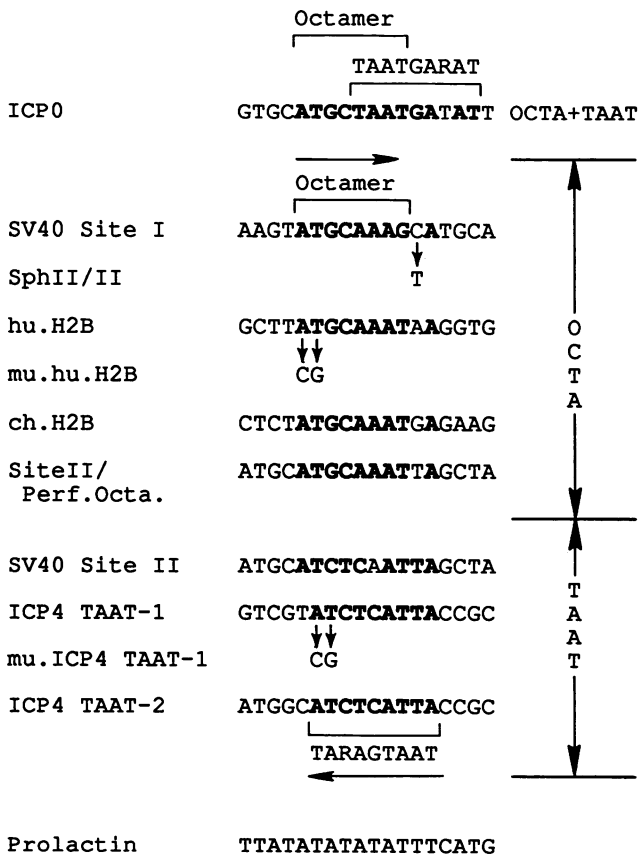

FIG. 2. Nucleotide sequence of binding-site probes. The sequences of the 12 binding sites used in the electrophoretic mobility retardation assays are shown. Sites were cloned into the polylinker of pUC119. The sequences are aligned to give maximum homology between adjacent sequences. Single or double point mutations that were introduced into a site are indicated by the arrows below the nucleotide sequence. Bold letters indicate identity to either the octamer/decamer sequence or the TAATGARAT sequence. Note that the TAATGARAT element can be best aligned to SV40 site II by deletion of the single A residue shown in plain type in the site II sequence and by inverting the orientation of the TAATGARAT motif.

different sites: three octamer-related sites (Fig. 3A to C), the divergent SV40 site II (Fig. 3D), an (OCTA$^-$)TAATGARAT site (Fig. 3E), and the prolactin Pit-1 binding site (Fig. 3F). The various proteins are referred to by the origin of the N terminus, POU domain, and C terminus. Thus, for example, Oct-1 carrying the Pit-1 POU domain is named Oct-1.P.1. The heterologous POU domains in these chimeras are 87% (Oct-2), 56% (Oct-3), and 66% (Pit-1) identical to the Oct-1 POU domain. In each panel of Fig. 3, the DNA-binding activities of Oct-1 (lanes 2 and 6) and the three Oct-1 POU-domain exchanges (lanes 3 to 5) are compared to those of the wild-type Oct-2, Oct-3, and Pit-1 proteins (lanes 7 to 9). Lane 10 shows the complex formed by HeLa cell-derived Oct-1 as a mobility marker; comparison with the in vitro-translated Oct-1 in lanes 2 and 6 shows that the in vitro-translated and HeLa cell Oct-1 proteins display the same relative preference for each binding site but the HeLa cell Oct-1 displays a sharper gradient of relative binding affinity for the different sites. We do not know the reason for this difference, but it could be due to either posttranslational modification of one or the other protein or a nonspecific effect of the different types of extracts.

As expected, under conditions selected to accentuate the differences in DNA binding by Oct-1 and Pit-1 (see Materials and Methods), Oct-1 and Oct-2 (2.2.2) manifest similar relative binding affinities for each probe, although Oct-2 generally displays a weaker affinity. In contrast, the more distantly related Oct-3 (3.3.3) protein and the Oct-1.3.1 chimera bind with low affinity to all of the probes, whereas Pit-1 (P.P.P) and Oct-1.P.1 bind the perfect octamer motif with an affinity between that displayed by the Oct-2 and Oct-3 proteins (Fig. 3A; compare lane 9 with lanes 7 and 8, and lane 5 with lanes 3 and 4). Thus, under our conditions, Pit-1 is a better octamer motif-binding protein than Oct-3. Pit-1 (lane 9), however, displays a different DNA-binding specificity to Oct-1, because although it binds less effectively than Oct-1 (lane 6) to the octamer sites (Fig. 3A to C), it binds more avidly than Oct-1 to SV40 site II (Fig. 3D) and the (OCTA$^-$)TAATGARAT site (Fig. 3E). Comparison of the wild-type Oct-2, Oct-3, and Pit-1 proteins to the POU-domain exchanges Oct-1.2.1, Oct-1.3.1, and Oct-1.P.1 on the different probes shows that the relative affinities of the chimeras and their corresponding wild-type protein are similar on each probe. Thus, here the POU domain is the major determinant of DNA-binding specificity.

The comigration of the Oct-1.P.1 complex with the Oct-1.1.1 complex in Fig. 3A to E indicates that the Pit-1 POU domain binds to octamer- and TAATGARAT-related sites as a monomer. On the natural prolactin promoter Pit-1 binding site, however, Oct-1.P.1 (Fig. 3F, lane 5) creates a prominent complex that migrates more slowly than the Oct-1 or Oct-1.2.1 complexes (lanes 2 and 3), probably as a result of dimer formation induced by the Pit-1 POU domain as described previously (17). Because of the difficulty in discriminating between the effects of POU-domain sequences on DNA contacts versus protein-protein contacts between monomers in the dimer complex, the prolactin site was not included in the following studies.

The POU-specific and homeodomain regions contact DNA. The results of several studies have suggested that when the POU domain is bound to DNA, the POU-specific region might contact DNA (17, 22, 50, 57), but there has been no direct evidence of such contacts. To assay for such contacts directly, we have tested whether the Oct-1 POU-specific region can be cross-linked to BrdU-substituted DNA. UV-induced cross-linking of protein to BrdU-substituted DNA identifies zero-range contacts between protein and DNA (29). Because, unlike the POU homeodomain, the POU-specific region has not been shown to bind DNA on its own (17, 22, 57), we devised a protease cleavage strategy to detect POU-specific region contacts with DNA in the context of the entire POU domain. As illustrated in Fig. 4A, we inserted the recognition site (IEGR) for the blood coagulation restriction protease factor X_A within the region of the Oct-1 linker segment that was shown previously to tolerate an insertion without any obvious effect on Oct-1 DNA-binding specificity (50). Indeed, as shown in Fig. 4B and by data not shown, the insertion of the factor X_A recognition site also did not obviously affect the DNA-binding specificity of the Oct-1 POU domain. To assay for POU-specific region contacts with DNA, we (i) cross-linked the entire POU domain to ^{32}P-labeled BrdU-substituted DNA, (ii) treated the cross-linked complex with DNase, (iii) cleaved the cross-linked POU domain with factor X_A, and (iv) fractionated the resulting protein-DNA adducts on an SDS-poly-

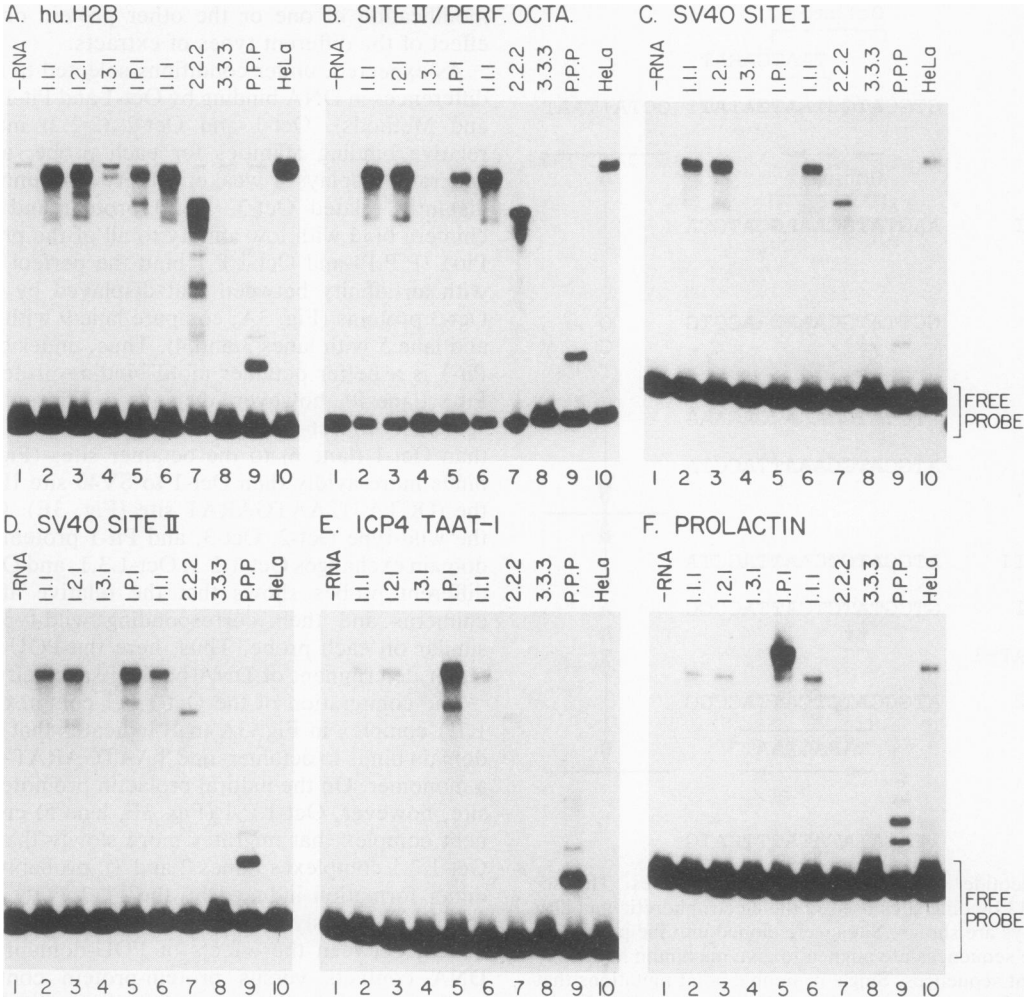

FIG. 3. Comparison of the DNA-binding affinities and specificities of Oct-1 hybrids with the Oct-2, Oct-3, or Pit-1 POU domains and the wild-type proteins. Each panel represents an electrophoretic mobility retardation assay on a different binding-site probe. Equal moles of in vitro-translated Oct-1 (1.1.1; lanes 2 and 6), Oct-1 chimeras carrying either the Oct-2 (1.2.1; lane 3), Oct-3 (1.3.1; lane 4), or Pit-1 (1.P.1; lane 5) POU domain, or wild-type Oct-2 (2.2.2; lane 7), Oct-3 (3.3.3; lane 8), or Pit-1 (P.P.P; lane 9) proteins were assayed in each panel. Lanes 10 contained HeLa cell nuclear extract. Background from the unprogrammed reticulocyte lysate is shown in lane 1 of each panel.

acrylamide gel to test whether label from the DNA had been transferred to the POU-specific region.

Figure 4C shows the results of such an experiment using Oct-1 POU-domain fusions to GST expressed and purified from *E. coli* either with or without the factor X_A recognition site (Fig. 4A). Both proteins become labeled with ^{32}P after exposure of the DNA-bound proteins to UV light (lanes 1, 3, and 5), whereas GST alone is not labeled (lane 7), indicating that the cross-linking conditions are specific for the DNA-binding POU domain. When the cross-linked wild-type Oct-1 POU-domain fusion protein is cleaved with thrombin between the GST and POU-domain sequences (Fig. 4A), only the POU-domain fragment is labeled with ^{32}P (lane 2), indicating that the entire protein is not nonspecifically cross-linked to DNA by the cross-linking procedure. As expected, the cross-linked wild-type Oct-1 POU-domain fusion protein is not affected by treatment with factor X_A (lane 4). However, identical treatment of the Oct-1 POU domain with the engineered factor X_A recognition site results in quantitative cleavage of the GST-POU domain fusion protein (lane 6). Two labeled fragments result from the factor X_A digestion: one that migrates as expected for the homeodomain alone and the other which has the expected mobility of the POU-specific region fusion to GST. These results indicate that both the POU homeodomain and POU-specific region, but not the GST sequences, can make contacts with the major groove of DNA where the photoreactive bromine atoms are positioned. Together with the effects of the POU-specific region on DNA sequence recognition described below, these cross-linking results suggest that the POU-specific region makes sequence-specific contacts when the POU domain is bound to DNA.

Sequence specificity of the Oct-1 and Pit-1 POU domains. To study the contribution of different segments of the Oct-1 POU domain to binding to the diverse Oct-1 regulatory targets, we compared the DNA-binding specificities of a series of Oct-1 and Pit-1 POU-domain segment swaps. Figure 5 shows a comparison of the DNA-binding specificities and relative affinities of the parental proteins Oct-1 and Oct-1.P.1 for the 11 different octamer- and TAATGARAT-related sites shown in Fig. 2. The DNA-binding assays shown in Fig. 5 are from the same experiment shown in Fig.

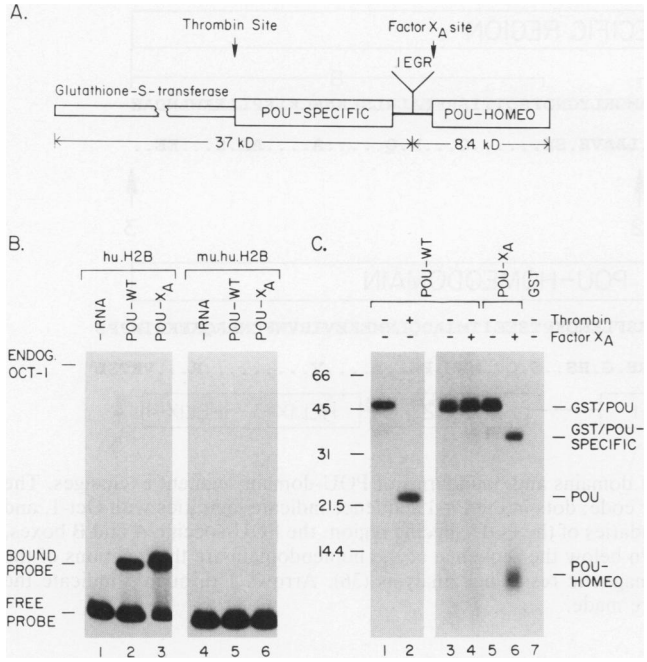

FIG. 4. Evidence that the POU-specific region and POU homeodomain contact DNA. (A) Schematic of the GST-POU domain fusion protein. The positions of the recognition sites for proteases factor X_A and thrombin are shown above the diagram. The expected sizes of protein fragments after digestion with factor X_A are indicated below. (B) Electrophoretic mobility retardation assays comparing the binding activities of the wild-type POU domain and the POU-X_A POU domain, containing the factor X_A site in the linker, shown on the wild-type (hu. H2B) and mutant (mu. hu. H2B) human H2B octamer motifs. (C) SDS-PAGE of the wild-type GST-POU domain fusion (lanes 1 to 4), the GST–POU-X_A fusion (lanes 5 and 6), and GST (lane 7) proteins after cross-linking to ^{32}P-labeled BraU-substituted DNA. The sample in lane 2 was treated with thrombin, samples in lanes 4 and 6 were treated with factor X_A, and the samples in lanes 1, 3, and 5 were mock treated. The migration of molecular weight standards is indicated at the left in kilodaltons.

7, but here, because equivalent exposures of the entire lanes are shown, the relative affinities of Oct-1 for the octamer and TAATGARAT sites as well as Oct-1.P.1 for these sites can be observed. Examination of their relative affinities for the 11 sites shows that canonical octamer-related sites (lanes 3 to 8 and 11 to 14) bind Oct-1 better than Oct-1.P.1, whereas the TAATGARAT sites (lanes 17, 18, 21, and 22) bind Oct-1.P.1 better than Oct-1. Consistent with this result, in this comparison the best Oct-1 binding sites are canonical octamer motifs (lanes 7, 11, and 13), whereas the best Oct-1.P.1 binding site is a TAATGARAT motif (lane 22). Point mutations within the octamer (lanes 9 and 10) and TAATGARAT (lanes 19 and 20) motifs adversely affect binding of both Oct-1 and Oct-1.P.1. The comparison of the two TAATGARAT sites shows that changes flanking the core sequence can have a marked effect on the relative affinity of both Oct-1 and Oct-1.P.1 for the binding site (see TAAT-1 and TAAT-2 in Fig. 2 and compare lanes 17 and 18 with lanes 21 and 22 in Fig. 5).

As shown in Fig. 2, sequence comparison of octamer and TAATGARAT sequences suggested that the divergent SV40 site II octamer motif is more closely related to the TAATGARAT motif than to the octamer motif (2). Consistent with this hypothesis, SV40 site II binds to the Oct-1.P.1 chimera better than to Oct-1 (compare lanes 15 and 16); this result represents the first biochemical test of the similarity between SV40 site II and a TAATGARAT motif. As expected, the 3-bp mutation that converts SV40 site II into a perfect octamer motif (site II/Perf. Octa.; Fig. 2) switches the relative binding preference in favor of the Oct-1 protein (compare lanes 13 and 14). Together, these results show that the Oct-1 and Pit-1 POU domains have different but overlapping DNA-binding specificities.

Structures of Oct-1 and Oct-1.P.1 POU-domain segment chimeras. Figure 6 shows a comparison of the Oct-1 and Pit-1 POU-domain sequences, with conserved motifs indicated. The relative positions of the four α helices identified in the Antennapedia homeodomain (36) are shown. The numbered arrows identify the positions at which POU-domain segments were precisely exchanged. As illustrated at the top of Fig. 7, two reciprocal series of exchanges between Oct-1 and Oct-1.P.1 were constructed in which either the A, B, linker,

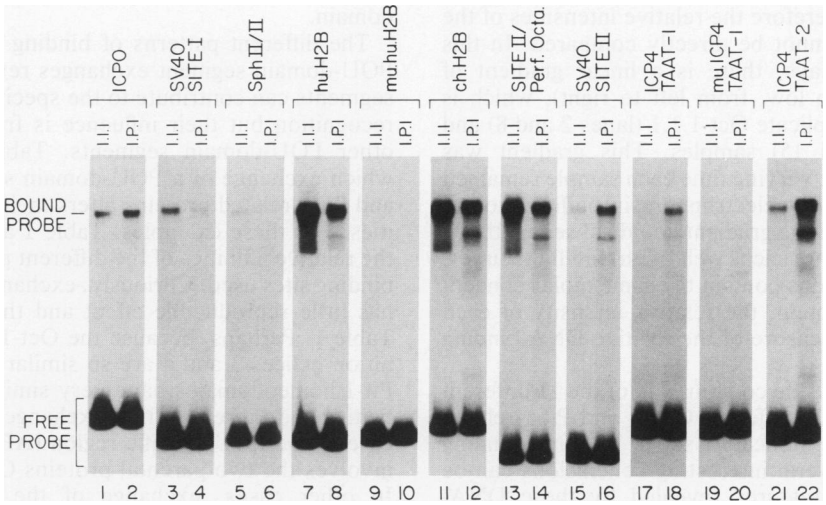

FIG. 5. Direct comparison of Oct-1.1.1 and Oct-1.P.1 DNA-binding activities. The probes indicated above the lanes were of identical specific activity, and similar exposures of the gels are shown for the binding of the Oct-1.1.1 (1.1.1) and Oct-1.P.1 (1.P.1) proteins. The sequences of the different sites are shown in Fig. 2.

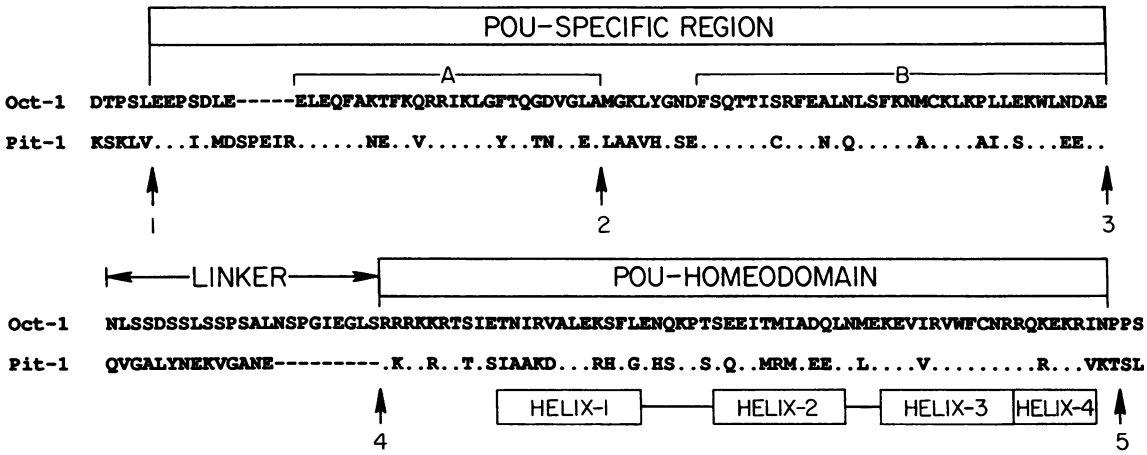

FIG. 6. Amino acid sequence comparison of the Oct-1 and Pit-1 POU domains and boundaries of POU-domain segment exchanges. The sequences of the Oct-1 and Pit-1 POU domains are shown in single-letter code; dots in the Pit-1 sequence indicate identities with Oct-1, and dashes indicate gaps used to maximize the sequence alignment. The boundaries of the POU-specific region, the POU-specific A and B boxes, the linker, and the homeodomain are shown above the sequences. Shown below the sequence of the homeodomain are the locations of the four α helices of the Antennapedia homeodomain defined by nuclear magnetic resonance analysis (36). Arrows 1 through 5 indicate the junctions at which Oct-1 and Pit-1 POU-domain segment exchanges were made.

or homeodomain segment was exchanged individually or the entire POU-specific region (labeled AB) was exchanged. The chimeras are named according to the structure of the parent construct and the origin of the foreign POU-domain segment(s). For example, when the linker of Oct-1 is exchanged for that of Pit-1, the resulting chimera is named Oct-1.1.1[L_p]. For clarity, we refer to the chimeric POU domains by using O and P to represent, in order, the origin of the POU-specific A and B boxes, linker, and homeodomain. In this case, the POU domain of Oct-1.1.1[L_p] is designated OOPO.

DNA-binding activity of POU-domain segment chimeras. Figure 7 shows the protein-DNA complexes generated by the POU-domain segment chimeras on different octamer- and TAATGARAT-related sites. Each row (A through K) represents a different binding site as listed to the right, and each column (1 through 15) represents binding to the protein indicated at the top of the figure. Here, different-length exposures are shown for different binding sites (see the legend to Fig. 7), and therefore the relative intensities of the bands between rows cannot be directly compared. In this particular series of assays, there is a linear gradient of binding activity (high to low, from left to right), which is controlled for by the duplicate Oct-1.1.1 (lanes 2 and 8) and Oct-1.P.1 (lanes 9 and 15) samples. This gradient was apparently caused by the varying time each sample remained in the well surrounded by electrophoresis buffer prior to electrophoresis, because the gradient could subsequently be eliminated by preloading each well with binding buffer. Because all of the reactions contain the same molar concentration of DNA and protein, the relative intensity of each complex in a row is a measure of the relative DNA-binding affinity of each protein.

As illustrated in Fig. 7, the combination of the 11 different DNA-binding sites and 12 different Oct-1- and Pit-1-related proteins results in a complicated series of DNA-binding patterns from which several interesting conclusions can be made. Among general features revealed by these DNA-binding patterns are that each of the 10 POU-domain chimeras can bind DNA, as illustrated by the ability of all the proteins to bind to the human histone H2B octamer motif (row D), and retains sequence specificity because they still do not bind effectively to the mutated H2B octamer motif (row E) or the mutated TAAT-1 motif (row J). These results indicate that all four regions of Oct-1 and Pit-1 POU domains are compatible with one another. Surprisingly, in many instances, the chimeras bind particular sites with higher affinity than either parent. For example, Oct-1.P.1[AB_o] binds SV40 site I (row B, column 10 [i.e., B-10]) with higher affinity than do the parents Oct-1.1.1 (B-8) or Oct-1.P.1 (B-9). Indeed, the highest DNA-binding activity exhibited in Fig. 7 is with Oct-1.1.1[H_p] on the human histone H2B octamer motif (D-7). Thus, under these in vitro assay conditions, the natural parents are not necessarily the highest-affinity DNA-binding proteins. If these results reflect in vivo DNA-binding activities, then an optimal Oct-1 octamer motif-binding activity may not have been favored during evolution. Instead, other activities, such as protein-protein interactions like that between the Oct-1 homeodomain and VP16 (46), may constrain the evolution of the DNA-binding domain.

The different patterns of binding by the Oct-1 and Pit-1 POU-domain segment exchanges reveal that different POU segments can contribute to the specificity of DNA sequence recognition but their influence is frequently dependent on other POU-domain segments. Table 1 lists examples in which exchange of a POU-domain segment between Oct-1- and Pit-1-related proteins alters their DNA-binding specificities. For these examples, Table 1 also lists quantitation of the relative affinities of the different proteins for the different binding sites used. Curiously, exchange of the homeodomain has little reproducible effect and therefore is not listed in Table 1. Perhaps, because the Oct-1 and Pit-1 DNA recognition helices 3 and 4 are so similar (Fig. 6), the Oct-1 and Pit-1 homeodomains have very similar DNA-binding specificities and therefore their exchange has little effect. In the case of the POU-specific region (AB), the exchange shown involves the two parental proteins Oct-1.1.1 and Oct-1.P.1. In other cases, exchange of the POU-domain segment between Oct-1.1.1 and Oct-1.P.1 does not have a strong effect on binding to the DNA sequences that we have tested, in the case of the A-box exchange in Oct-1.P.1

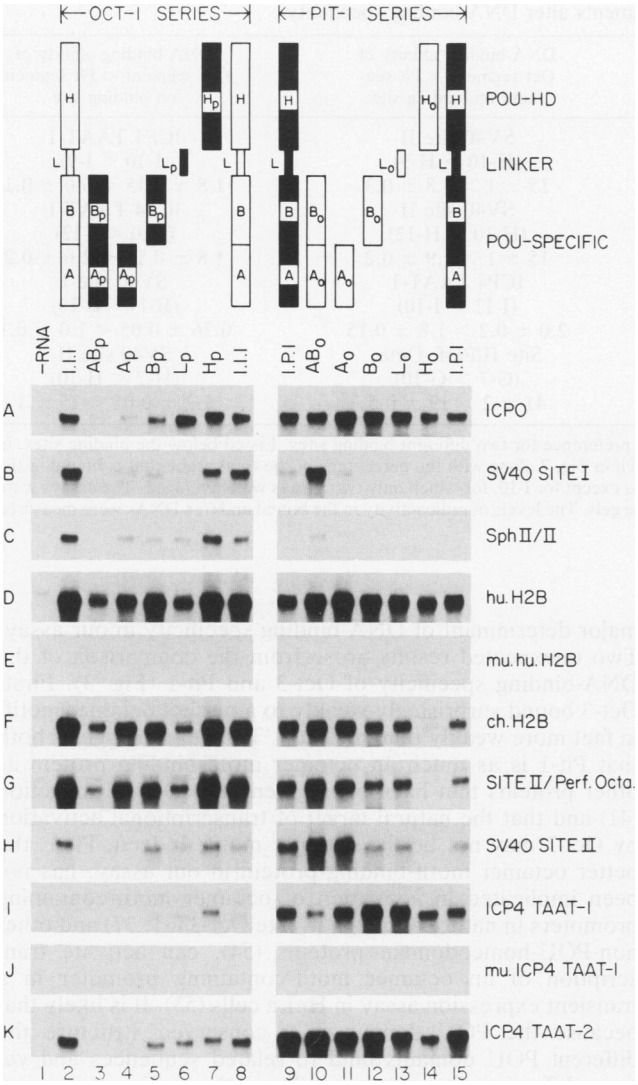

FIG. 7. DNA-binding activity of Oct-1/Pit-1 POU-domain segment chimeras. Each row (A through K) represents a different DNA-binding site, and each column (1 through 15) represents a different in vitro-translated protein whose structure is shown schematically at the top. For each chimera, only the region that has been exchanged is drawn; the remainder of the POU domain is derived from the parent construct. Only the protein-DNA complexes generated by the electrophoretic mobility retardation assays are shown. All probes had identical specific activities, and each binding reaction contains an equimolar concentration of protein. To allow comparison of the DNA-binding patterns for the weaker binding sites, exposures of different length are shown as follows: ICP0, 12 h; SV40 site I, 28 h; *Sph*II/II, 28 h; hu. H2B, 6 h; mu. hu. H2B, 28 h; ch. H2B, 6 h; site II/Perf. Octa., 12 h; SV40 site II, 12 h; ICP4 TAAT-1, 12 h; mu. ICP4 TAAT-1, 28 h; ICP4 TAAT-2, 12 h.

(PPPP) and Oct-1.P.1[A_o] (OPPP) (compare columns 9 and 11), but exchange of the A box between POU-domain chimeras, such as Oct-1.P.1[B_o] (POPP) and Oct-1.P.1 [AB_o] (OOPP), can have a large effect on DNA-binding specificity (compare columns 10 and 12). For example, Oct-1.P.1[B_o] (POPP), which carries the Pit-1 A box, binds better to the ICP4 TAAT-1 site than does Oct-1.P.1[AB_o] (OOPP), which contains the Oct-1 A box (compare I-12 with I-10), whereas on SV40 site II the opposite preference is exhibited (compare H-12 with H-10). Thus, the contribution of the A box to sequence-specific DNA recognition is dependent on its exact POU-domain context. In this case, the Oct-1 B box combined with the Pit-1 linker and homeodomain reveals the influence that the A box can have on DNA-binding specificity.

Unexpectedly, the hypervariable linker region influences the DNA-binding specificity of the POU domain but only in particular POU-domain contexts, not when the linkers are exchanged between Oct-1.1.1 and Oct-1.P.1. For example, Oct-1.1.1 [H_p] (OOOP), which carries the Oct-1 linker in the context of the Oct-1 POU-specific region and Pit-1 homeodomain, binds to site II/Perf. Octa. better than does the corresponding chimera with the Pit-1 linker, Oct-1.P.1 [AB_o] (OOPP) (compare G-7 and G-10), but the reverse preference is exhibited on the original SV40 site II sequence (compare H-7 and H-10). Perhaps the linker influences the DNA-binding specificity by altering how the POU-specific and homeodomains contact DNA. This kind of positioning effect may also explain the fact that the determination of the DNA-binding preference of a POU domain for an octamer-related versus TAATGARAT-related sequence cannot be ascribed to the origin of a particular POU-domain segment. For example, when the Pit-1 homeodomain is replaced by the Oct-1 homeodomain in Oct-1.P.1[H_o] (PPPO), binding to two different octamer motifs (chicken H2B [F-14] and site II/Perf. Octa. [G-14]) is not improved; on the contrary, it is diminished. Perhaps the Pit-1 POU-specific and linker segments do not allow the Oct-1 homeodomain to be positioned appropriately to recognize certain perfect octamer motifs.

Subtle changes in the binding site can also influence how the segments bind DNA. SV40 site I and sphII/II differ only at the nonconserved position between positions 8 and 10 of the decamer motif (Fig. 2). This region of the octamer element probably contacts the homeodomain (57). Yet this single base pair difference affects how the POU-specific region influences DNA binding, because when the Oct-1 POU-specific region is exchanged for the Pit-1 POU-specific region in Oct-1.P.1, binding to the SV40 site I sequence is increased (compare B-9 and B-10) considerably more than is binding to the sphII/II site (compare C-9 and C-10). If indeed a change in the DNA sequence where the homeodomain contacts DNA can influence how the POU-specific region recognizes DNA, then this result may explain how Oct-1 can bind to divergent and yet specific sequences because a change in one portion of the sequence results in favoring a different sequence elsewhere in the binding site.

DISCUSSION

We have shown previously that Oct-1 can bind to a diverse set of regulatory sequences because it makes flexible contacts with DNA (2). In the present study, we have examined how Oct-1 is capable of such flexibility. The results suggest that the POU domain is a dynamic DNA-binding domain in which multiple segments interdependently confer flexible sequence-specific interactions. Previously, we showed that both the POU-specific region and POU homeodomain contribute to DNA binding but could not distinguish whether only one or both regions contributed to the specificity of DNA binding (50). The studies described here, together with those of Ingraham et al. (17) and Verrijzer et al. (57), show that indeed both segments, and surprisingly the linker region, contribute to sequence-specific interactions; the ability to cross-link the POU-specific region to the major groove of DNA suggests that this region confers DNA-binding speci-

TABLE 1. Evidence that POU-domain segments alter DNA-binding specificity[a]

POU-domain segment swap	Oct-1 segment	Pit-1 segment	DNA-binding affinity of Oct segment > Pit segment on binding site	DNA-binding affinity of Oct segment < Pit segment on binding site
POU specific	Oct-1.P.1[AB_o] (OOPP)	Oct-1.P.1 (PPPP)	SV40 site II (H-10 > H-9) 15 ± 1 > 2.8 ± 0.3	ICP4 TAAT-1 (I-10 < I-9) 1.8 ± 0.15 < 6.0 ± 0.1
A box	Oct-1.P.1[AB_o] (OOPP)	Oct-1.P.1[B_o] (POPP)	SV40 site II (H-10 > H-12) 15 ± 1 > 1.9 ± 0.2	ICP4 TAAT-1 (I-10 < I-12) 1.8 ± 0.15 < 2.6 ± 0.2
B box	Oct-1.P.1[AB_o] (OOPP)	Oct-1.P.1[A_o] (OPPP)	ICP4 TAAT-1 (I-11 > I-10) 2.0 ± 0.2 > 1.8 ± 0.15	SV40 site I (B-11 <B-10) 0.36 ± 0.05 < 1.0 ± 0.1
Linker	Oct-1.1.1[H_p] (OOOP)	Oct-1.P.1[AB_o] (OOPP)	Site II/Perf. Octa. (G-7 > G-10) 41 ± 2 > 19 ± 0.5	SV40 site II (H-7 < H-10) 1.2 ± 0.05 < 15 ± 1

[a] Examples are shown in which exchanging the POU-segment listed changes the preference for two different binding sites. Listed below the binding sites, in parentheses, are the coordinates (row-column) of the protein-DNA complexes shown in Fig. 7 along with the percentage of the total probe that is bound in the complex (bound/free plus bound). The results shown are the averages of three assays except for I-10, for which only two assays were averaged. The assays were performed as for Fig. 7 except that there was no gradient of binding activity across the gels. The levels of radioactivity in the bound and free DNAs were measured with a Phosphor Imager (Molecular Dynamics).

ficity by making sequence-specific contacts with DNA. These results argue against the model proposed by Garcia-Blanco et al. (11), which suggested that the POU homeodomain is entirely responsible for sequence-specific DNA contacts by the closely related Oct-2 protein.

In a detailed comparison of the binding specificities of the Oct-1 POU domain and the Oct-1 homeodomain alone, Verrijzer et al. (57) showed that in the absence of the POU-specific region, the homeodomain makes contacts with a subset of the residues contacted by the entire POU domain. These results suggested that the POU-specific region contacts the 5′ portion of the octamer/decamer motif (ATGCAAATNA) and the POU homeodomain contacts the A/T-rich 3′ half. Homeodomains recognize the A/T-rich sequence ATTA or its complement TAAT, and the X-ray crystallographic structure of the engrailed homeodomain-DNA complex indicates homeodomain interactions with three of the four positions of the ATTA or TAAT sequence (21). These interactions are consistent with the homeodomain contacting the 3′ end of the decamer motif (ATNA) or the 5′ end of the TAATGARAT motif (TAAT), which further justifies the sequence alignment shown in Fig. 2, in which the TAATGARAT sequence is inverted.

In some studies, it is evident that the POU homeodomain alone displays different DNA-binding properties than in the context of the POU domain. Kristie and Sharp (22) showed that the Oct-1 homeodomain fused to protein A bound an octamer-related element as a dimer whereas a POU-domain counterpart bound only as a monomer. Furthermore, Ingraham et al. (17) showed that the entire Pit-1 POU domain can bind to a very different binding site than does the Pit-1 homeodomain alone on the same DNA fragment. The activities of the Oct-1/Pit-1 POU-domain chimeras described here show that indeed segments of the POU domain can influence one another's DNA-binding specificity. In our studies, we have used Oct-1- and Pit-1-related proteins synthesized in vitro. In vivo, Oct-1 and Pit-1 may be modified in ways that also affect DNA binding such as by phosphorylation.

The POU domain is responsible for the DNA-binding specificity of POU proteins. The comparison of the DNA-binding specificity of the POU domains of Oct-1 through Oct-3 and Pit-1 in their natural contexts or in the context of Oct-1 N- and C-terminal sequences shows that the POU domain is the major determinant of DNA-binding specificity in our assay. Two unexpected results arose from the comparison of the DNA-binding specificity of Oct-3 and Pit-1 (Fig. 3). First, Oct-3 bound surprisingly weakly to a perfect octamer motif, in fact more weakly than did Pit-1. This result suggests both that Pit-1 is as much an octamer motif-binding protein as other proteins that have been given the "Oct" designation (41) and that the natural target of transcriptional activation by Oct-3 may not be the octamer motif. Indeed, Pit-1, the better octamer motif-binding protein in our assay, has not been implicated in activation of octamer motif-containing promoters in nature although it, like Oct-3 (31, 37) and other non-POU-homeodomain proteins (54), can activate transcription of an octamer motif-containing promoter in a transient expression assay in HeLa cells (53). It is likely that because the POU domain is a conserved structure the different POU domains bind to related sequences and yet have different natural targets. Thus, the finding that many proteins bind to the octamer motif (41) may be more indicative of the existence of many POU-domain proteins than bona fide octamer motif regulatory proteins. The only protein to display the same binding specificity as Oct-1 in our assay is Oct-2. This result is consistent with those of a previous study (45) and the finding that the octamer motif is present in promoters that are apparently regulated in vivo by either Oct-1 or Oct-2.

The second unexpected finding is that Pit-1 can bind better than Oct-1 to the class II (OCTA$^-$)TAATGARAT motif; a site on which Oct-1 binds weakly but is able, by association with VP16, to activate transcription of HSV IE promoters (34, 55). Unlike Oct-1, Pit-1 is unable to associate effectively with VP16 (46). Thus, if Pit-1 or a POU protein with a similar DNA-binding specificity were expressed in an HSV-infected cell, it could interfere with activation of HSV IE transcription by blocking the association of Oct-1 with VP16 on these VP16-responsive targets. This overlapping and yet opposite preference of Oct-1 and Pit-1 for regulatory targets is reminiscent of λ repressor and λ Cro, which also bind to related sites, the λ operators O_R1, O_R2, and O_R3, with the opposite binding-site preference, resulting in differential activation and repression of transcription (35).

The POU domain: a dynamic DNA-binding structure. Several aspects of the studies described here suggest that the

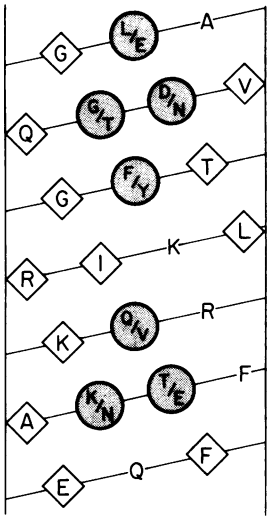

FIG. 8. Comparison of Oct-1 and Pit-1 POU-specific A-box sequences displayed as an α helix. The Oct-1 and Pit-1 POU-specific A-box sequences are superimposed on an α-helical projection. Invariant amino acids among all reported POU-domain sequences are enclosed in diamonds, and the amino acid positions that differ between Oct-1 and Pit-1 are circled and shaded. The Oct-1 amino acid at these positions is indicated in the upper left of the circled positions, and the Pit-1 sequence is shown at the lower right.

POU domain is a dynamic DNA-binding structure. This conclusion is illustrated best by the influence that segments of the POU domain have on one another. An unexpected result is that the hypervariable linker region can influence the DNA-binding specificity of a chimeric POU domain carrying the Oct-1 POU-specific region and Pit-1 homeodomain. It is not evident at present whether this influence on DNA binding is a result of sequence-specific DNA contacts by the linker or rather the different sizes of the Pit-1 (15 amino acids) and Oct-1 (24 amino acids) linkers, which may affect how the POU-specific region and POU homeodomain are aligned on the DNA. Originally, the variability in size and sequence of the linker in different POU proteins suggested that it was not critical for determining DNA-binding specificity (14, 50), but the remarkable 12-of-16-amino-acid similarity between the linkers in the *Drosophila* Cf-1a and mammalian SCIP POU proteins ($S^{TT}_{SS}GSPT^{SI}_{NL}DKIAAQG$ [13, 19, 25]) suggests that the linker could have a significant function such as in DNA-binding specificity or protein-protein interactions.

In contrast to the linker region, the POU-specific region is the most highly conserved region among POU domains. The structure of the POU-specific region is unknown, but secondary structure predictions suggest a large amount of α-helical structure (17). Figure 8 shows an interesting outcome of superimposing the POU-specific A-box sequences of Oct-1 and Pit-1 as an α helix. Positions that differ between Oct-1 and Pit-1 are circled, and positions that are conserved among all POU proteins described are indicated by the diamonds. Although the A box may not exist as a single uninterrupted α helix as shown in Fig. 8, it is interesting that the differences between the Oct-1 and Pit-1 A boxes cluster on one face of the hypothetical α-helical structure. Furthermore, the neighboring surface is completely conserved among the known POU proteins. The context-dependent influence of the A box on DNA-binding specificity suggests that the regions of difference between the Oct-1 and Pit-1 A-box sequences serve to specify DNA sequence recognition by contacting DNA directly or affecting contacts with other segments of the POU domain (e.g., the B box).

The POU domain is an unusual DNA-binding structure. Most other DNA-binding proteins bind DNA either with a single contiguous DNA-binding domain, as in the case of other homeodomain proteins, or as homomeric or heteromeric dimers. In the POU domain, a heterodimeric-type structure has been retained in a single polypeptide. The two halves of this structure are very different in sequence, suggesting different structures, and yet through evolution the POU-specific region and POU homeodomain have remained linked as a complete POU domain. The interdependence of the different POU-domain segments for binding to DNA may explain why these two regions have evolved together.

The surprising influence that different segments of the POU domain can have on one another and the apparent effect a base pair change in the homeodomain-binding site (as in SV40 site I and sphII/II) can have on DNA binding by the POU-specific region lead to a picture of the POU domain as a dynamic structure that an adopt different overall conformations on different binding sites. This picture can explain how the Oct-1 POU domain is able to recognize such a divergent and yet specific set of regulatory targets because an alteration in contacts with one region of the binding site can influence the contacts made in the other region of the binding site. Such an apparently flexible protein-DNA complex may display binding-site-specific structures that influence the ability of the complex to interact with other proteins involved in transcriptional activation as in the case of the adaptor protein VP16, which interacts with the Oct-1 POU domain on particular binding sites.

ACKNOWLEDGMENTS

We thank Nouria Hernandez, Bruce Stillman, and Angus Wilson for helpful comments on the manuscript and Louis Staudt for providing the Oct-3 clone. R.A. gratefully acknowledges Masafumi Tanaka for his help in many phases of this work and Mario Cerritelli for his help in establishing the T7 expression system. We thank Jane Reader for typing the manuscript, and we thank Jim Duffy and Phil Renna for artwork.

This work was supported by Public Health Service grant CA-13106 from the National Cancer Institute.

REFERENCES

1. **apRhys, C., D. Ciufo, E. O'Neil, T. Kelly, and G. Hayward.** 1989. Overlapping octamer and TAATGARAT motifs in the VF65-response elements in herpes simplex virus immediate-early promoters represent independent binding sites for the cellular Nuclear Factor III. J. Virol. **63**:2798–2812.
2. **Baumruker, T., R. Sturm, and W. Herr.** 1988. OBP100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. Genes Dev. **2**:1400–1413.
3. **Bebenek, K., and T. A. Kunkel.** 1989. The use of native T7 DNA polymerase for site directed mutagenesis. Nucleic Acids Res. **17**:5408.
4. **Bodner, M., J.-L. Castrillo, L. E. Theill, T. Deerinck, M. Ellisman, and M. Karin.** 1988. The pituitary-specific transcription factor GHF-1 is a homeobox-containing protein. Cell **55**: 505–518.
5. **Bürglin, T., M. Finney, A. Coulson, and G. Ruvkun.** 1989. *Caenorhabditis elegans* has scores of homeobox-containing genes. Nature (London) **341**:239–243.
6. **Cleary, M., and W. Herr.** Unpublished results.
7. **Clerc, R., L. Corcoran, H. LeBowitz, D. Baltimore, and P. Sharp.** 1988. The B-cell specific Oct-2 protein contains a POU box and homeo-box-type domains. Genes Dev. **2**:1570–1581.

8. **Das, G., and W. Herr.** Unpublished results.
9. **Eghtedarzadeh, M., and S. Henikoff.** 1986. Use of oligonucleotides to generate large deletions. Nucleic Acids Res. **14**:5115.
10. **Falkner, F. G., and H. G. Zachau.** 1984. Correct transcription of an immunoglobulin κ gene requires an upstream fragment containing conserved sequence elements. Nature (London) **310**:71–74.
11. **Garcia-Blanco, M., R. Clerc, and P. Sharp.** 1989. The DNA-binding homeo domain of the Oct-2 protein. Genes Dev. **3**:739–745.
12. **Goding, C., and P. O'Hare.** 1989. Herpes simplex virus Vmw65-octamer-binding protein interaction: a paradigm for combinatorial control of transcription. Virology **173**:363–367.
13. **He, X., M. Treacy, D. Simmons, H. Ingraham, L. Swanson, and M. Rosenfeld.** 1989. Expression of a large family of POU-domain regulatory genes in mammalian brain development. Nature (London) **340**:35–42.
14. **Herr, W., R. A. Sturm, R. G. Clerc, L. M. Corcoran, D. Baltimore, P. A. Sharp, H. A. Ingraham, M. G. Rosenfeld, M. Finney, G. Ruvkun, and H. R. Horvitz.** 1988. The POU-domain: a large conserved region in the mammalian *pit*-1, *oct*-1, *oct*-2, and *Caenorhabditis elegans unc-86* gene products. Genes Dev. **2**:1513–1516.
15. **Hope, I. A., and K. Struhl.** 1987. GCN4, a eucaryotic transcriptional activator protein, binds as a dimer to target DNA. EMBO J. **6**:2781–2784.
16. **Ingraham, H. A., R. Chen, H. J. Mangalam, H. P. Elsholtz, S. E. Flynn, C. R. Lin, D. M. Simmons, L. Swanson, and M. G. Rosenfeld.** 1988. A tissue-specific transcription factor containing a homeodomain specifies a pituitary phenotype. Cell **55**:519–529.
17. **Ingraham, H., S. Flynn, J. Voss, V. Albert, M. Kapiloff, L. Wilson, and M. Rosenfeld.** 1990. The POU-specific domain of Pit-1 is essential for sequence-specific, high affinity DNA binding and DNA-dependent Pit-1-Pit-1 interactions. Cell **61**:1021–1033.
18. **Johnson, P., and S. McKnight.** 1989. Eukaryotic transcriptional regulatory proteins. Annu. Rev. Biochem. **58**:799–839.
19. **Johnson, W., and J. Hirsch.** 1990. Binding of a *Drosophila* POU-domain protein to a sequence element regulating gene expression in specific dopaminergic neurons. Nature (London) **343**:467–470.
20. **Kemler, I., E. Schreiber, M. Müller, P. Matthias, and W. Schaffner.** 1989. Octamer transcription factors bind to two different sequence motifs of the immunoglobulin heavy chain promoter. EMBO J. **8**:2001–2008.
21. **Kissinger, C. R., B. Liu, E. Martin-Blanco, T. Kornberg, and C. Pabo.** 1990. Crystal structure of an *engrailed* homeodomain-DNA complex at 2.8 Å resolution: a framework for understanding homeodomain-DNA interactions. Cell **63**:579–590.
22. **Kristie, T., and P. Sharp.** 1990. Interactions of the Oct-1 POU subdomains with specific DNA sequences and the HSV α-*trans*-activator protein. Genes Dev. **4**:2383–2396.
23. **Laemmli, U.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
24. **Laughon, A., and M. Scott.** 1984. Sequence of *Drosophila* segmentation gene: protein structure homology with DNA-binding proteins. Nature (London) **310**:25–31.
25. **Monuki, E., G. Weinmaster, R. Kuhn, and G. Lemke.** 1989. SCIP: a glial POU domain gene regulated by cyclic AMP. Neuron **3**:783–793.
26. **Müller, M., S. Ruppert, W. Schaffner, and P. Matthias.** 1988. A cloned octamer transcription factor stimulates transcription from lymphoid-specific promoters in non-B cells. Nature (London) **336**:544–551.
27. **Mullis, K. B., and F. A. Faloona.** 1987. Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. Methods Enzymol. **155**:335–350.
28. **Nelson, C., V. R. Albert, H. P. Elsholtz, L. I.-W. Lu, and M. G. Rosenfeld.** 1988. Activation of cell-specific expression of rat growth hormone and prolactin genes by a common transcription factor. Science **239**:1400–1405.
29. **Ogata, R., and W. Gilbert.** 1976. Contacts between the *lac* repressor and thymines in the *lac* operator. Proc. Natl. Acad. Sci. USA **74**:4973–4976.
30. **Ondek, B., L. Gloss, and W. Herr.** 1988. The SV40 enhancer contains two distinct levels of organization. Nature (London) **333**:40–45.
31. **Okamoto, K., H. Okazawa, A. Okuda, M. Sakai, M. Murmatsu, and H. Hamada.** 1990. A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells. Cell **60**:461–472.
32. **Otting, G., Y. Q. Qian, M. Billeter, M. Müller, M. Affolter, W. Gehring, and K. Wüthrich.** 1990. Protein-DNA contacts in the structure of a homeodomain-DNA complex determined by nuclear magnetic resonance spectroscopy in solution. EMBO J. **9**:3085–3092.
33. **Parslow, T. G., D. L. Blair, W. J. Murphy, and D. K. Granner.** 1984. Structure of the 5′ ends of immunoglobulin genes: a novel conserved sequence. Proc. Natl. Acad. Sci. USA **81**:2650–2654.
34. **Preston, C. M., M. C. Frame, and M. Campbell.** 1988. A complex formed between cell components and a herpes simplex virus structural polypeptide binds to a viral immediate early gene regulatory DNA sequence. Cell **52**:425–434.
35. **Ptashne, M.** 1986. A genetic switch. Cell Press, Cambridge, and Blackwell Scientific Publications, Palo Alto.
36. **Qian, Y. Q., M. Billeter, G. Otting, M. Müller, W. J. Gehring, and K. Wüthrich.** 1989. The structure of the *Antennapedia* homeodomain determined by NMR spectroscopy in solution: comparison with prokaryotic repressors. Cell **59**:573–580.
37. **Rosner, M., A. Vigano, K. Ozato, P. Timmons, F. Poirier, P. Rigby, and L. Staudt.** 1990. A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature (London) **345**:686–692.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Schaffner, W.** 1989. How do different transcription factors binding the same DNA sequence sort out their jobs? Trends Genet. **5**:37–39.
40. **Scheidereit, C., J. Cromlish, T. Gerster, K. Kawakami, C.-G. Balmaceda, R. Curie, and R. Roeder.** 1988. A human lymphoid-specific transcription factor that activates immunoglobulin genes is a homeobox protein. Nature (London) **336**:551–557.
41. **Schöler, H. R., A. Hatzopolous, R. Balling, N. Suzuki, and P. Gruss.** 1989. A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor. EMBO J. **8**:2543–2550.
42. **Schöler, H., S. Ruppert, N. Suzuki, K. Chowdhury, and P. Gruss.** 1990. New type of POU-domain in germ line-specific protein Oct-4. Nature (London) **344**:435–439.
43. **Shepard, J., W. McGinnis, A. Carrasco, E. DeRobertis, and W. Gehring.** 1984. Fly and frog homeodomains show homologies with yeast mating type regulatory proteins. Nature (London) **310**:70–71.
44. **Singh, H., R. Sen, D. Baltimore, and P. Sharp.** 1986. A nuclear factor that binds to a conserved sequence motif in transcriptional control elements of immunoglobin genes. Nature (London) **319**:154–158.
45. **Staudt, L., H. Singh, R. Sen, T. Wirth, P. Sharp, and D. Baltimore.** 1986. A lymphoid-specific protein binding to the octamer motif of immunoglobulin genes. Nature (London) **323**:640–643.
46. **Stern, S., M. Tanaka, and W. Herr.** 1989. The Oct-1 homeodomain directs formation of a multiprotein-DNA complex with the HSV transactivator VP16. Nature (London) **341**:624–630.
47. **Studier, F. W., A. Rosenberg, J. Dunn, and J. Dubendorf.** 1990. Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol. **185**:60–89.
48. **Sturm, R., T. Baumruker, B. Franza, Jr., and W. Herr.** 1987. A 100 kD HeLa cell octamer binding protein (OBP100) interacts differently with two separate octamer-related sequences within the SV40 enhancer. Genes Dev. **1**:1147–1160.
49. **Sturm, R. A., G. Das, and W. Herr.** 1988. The ubiquitous octamer-binding protein Oct-1 contains a POU-domain with a

homeo box subdomain. Genes Dev. **2**:1582–1599.
50. **Sturm, R. A., and W. Herr.** 1988. The POU-domain is a bipartite DNA-binding structure. Nature (London) **336**:601–604.
51. **Suzuki, N., H. Rohdenwohld, T. Neuman, P. Gruss, and H. Schöler.** 1990. Oct-6: a POU transcription factor expressed in embryonal carcinoma and in the developing brain. EMBO J. **9**:3723–3732.
52. **Tanaka, M., and W. Herr.** 1990. Differential transcriptional activation by Oct-1 and Oct-2: interdependent activation domains induce Oct-2 phosphorylation. Cell **60**:375–386.
53. **Tanaka, M., and W. Herr.** Unpublished results.
54. **Thali, M., M. Müller, M. DeLorenzi, P. Mathias, and M. Bienz.** 1988. *Drosophila* homeotic genes encode transcriptional activators similar to mammalian OTF-2. Nature (London) **336**:598–601.
55. **Triezenberg, S., K. LaMarco, and S. McKnight.** 1988. Evidence of DNA:protein interactions that mediate HSV-1 immediate early gene activation by VP16. Genes Dev. **2**:730–742.
56. **Verrijzer, C., A. Kal, and P. van der Vliet.** 1990. The DNA binding domain (POU domain) of transcription factor Oct-1 suffices for stimulation of DNA replication. EMBO J. **9**:1883–1888.
57. **Verrijzer, C., A. Kal, and P. van der Vliet.** 1990. The oct-1 homeo domain contacts only part of the octamer sequence and full oct-1 DNA-binding activity requires the POU-specific domain. Genes Dev. **4**:1964–1974.
58. **Wildeman, A., P. Sassone-Corsi, T. Grundström, M. Zenke, and P. Chambon.** 1984. Stimulation of in vitro transcription from the SV40 early promoter involves specific a trans-acting factor. EMBO J. **3**:3129–3133.