Structure of Pit-1 POU domain bound to DNA as a dimer: unexpected arrangement and flexibility

Eric M. Jacobson,1 Peng Li,2,3 Alfonso Leon-del-Rio,2,3 Michael G. Rosenfeld,2 and Aneel K. Aggarwal1,4

1Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032 USA; 2Howard Hughes Medical Institute, University of California at San Diego, La Jolla, California 92093 USA

Pit-1, a member of the POU domain family of transcription factors, characterized by a bipartite DNA-binding domain, serves critical developmental functions based on binding to diverse DNA elements in its target genes. Here we report a high resolution X-ray analysis of the Pit-1 POU domain bound to a DNA element as a homodimer. This analysis reveals that Pit-1 subdomains bind to perpendicular faces of the DNA, rather than opposite faces of the DNA as in Oct-1. This is accomplished by different spacing and orientation of the POU-specific domain. Contrary to previous predictions, the dimerization interface involves the carboxyl terminus of the DNA recognition helix of the homeodomain, which in an extended conformation interacts with specific residues at the amino terminus of helix α1 and in the loop between helices α3 and α4 of the POU-specific domain of the symmetry related monomer. These features suggest the molecular basis of disease-causing mutations in Pit-1 and provide potential basis for the flexible allostery between protein domains and DNA sites in the activation of target genes.

[Key Words: Crystal structure; Pit-1; POU-domain factors; dimerization; spacing; orientation]

Received October 25, 1996; revised version accepted December 3, 1996.

The expression of specific genes in cells is often governed by families of transcription factors harboring DNA-binding motifs. In eukaryotes, these factors often exhibit extraordinary versatility. For instance, the bZip and helix-loop-helix families of transcription factors bind to DNA as both homo- and heterodimers (Bexevanis and Vinson 1993; Glover and Harrison 1995). The nuclear receptors go a step further and recognize DNA elements with different spacings and polarities between the half-sites (Naar et al. 1991; Umesono et al. 1991; Kurokawa et al. 1993). The POU domain family of transcription factors are perhaps the most remarkable in their ability to bind DNA as both monomers and dimers and in their ability to adopt diverse configurations (Wegner et al. 1993; Herr and Cleary 1995). Part of this versatility derives from their unusual bipartite structure, consisting of a highly conserved ~75 amino acid POU-specific domain [POU1], tethered by a linker of variable length and composition, to a 60-amino-acid homeodomain [POU2]. The flexibility of the linker, in particular, allows the possibility of different relative spacings and orientations between the subdomains. This helps to expand the repertoire of DNA sites recognized by POU-domain proteins and has important implications for understanding the mechanisms of combinatorial control of gene expression in eukaryotes.

The POU domain was first identified as a novel DNA-binding motif from the cloning of mammalian transcriptional regulators Pit-1, Oct-1, and Oct-2, and the Caenorhabditis elegans developmental modulator Unc-86 (Herr et al. 1988). Numerous other members of the POU family have now been identified in metazoan organisms, ranging from flies to humans. In almost all cases they have important roles in cellular differentiation and exhibit a characteristic ability of recognizing diverse DNA elements. Pit-1, for instance, is expressed exclusively in the anterior pituitary gland and is responsible for both the development of three of the five endocrine cell types and the trans-activation of several genes, including prolactin (Prl) and growth hormone (GH) genes (Bodner et al. 1988; Ingraham et al. 1988; Li et al. 1990; Lin et al. 1992). Pit-1 binds to a complex pattern of sites within the promoters and enhancers of the genes it regulates. The sequences of these sites vary widely around a weak consensus sequence (A/T)(A/T)TATNCAT, raising the intriguing possibility that Pit-1 might be configured differently on the different sites. Most of these sites bind Pit-1 as a dimer, although the prolactin Prl-1D site accommodates Pit-1 as a monomer (Ingraham et al. 1990).
Holloway et al. 1995). This differential binding can lead to alternative uses of the synergistic activation domains (Holloway et al. 1995). Oct-1 was identified originally by its recognition of the conserved octamer motif TAAAG-CAT found in the promoters of a number of tissue-specific and ubiquitously expressed genes (Wegner et al. 1993; Herr and Cleary 1995). Oct-1 can also bind to the distinct TAATGARAT (R-purine) motif in regulating the expression of the immediate-early genes of herpes simplex virus [Baumruker et al. 1988; apRhys et al. 1989]. Oct-1 functions primarily as a monomer, but can also form cooperative homo- and heterodimers with Oct-2 on a site in the immunoglobulin heavy chain promoter [Kemler et al. 1989; Poellinger and Roeder 1989]. Pit-1 and Oct-1 can also cooperate to form a heterodimer on the prolactin Prl-1P site [Voss et al. 1991]. A more dramatic example of POU domain flexibility is provided by the neuronal protein Brn-2. Brn-2 accommodates different spacings between its subdomain-binding sites, namely 0, 2, or 3 bp [Li et al. 1993]. Moreover, its POU-specific domain appears capable of switching orientations on different DNA elements. Together, these studies allude to the tremendous flexibility of POU-domain proteins that is crucial to their functional specificities. Our knowledge of POU domain structure, however, is limited to nuclear magnetic resonance (NMR) studies of proteins that is crucial to their functional specificities. Our knowledge of POU domain structure, however, is limited to nuclear magnetic resonance (NMR) studies of the individual subdomains [Assa-Munt et al. 1993; Dekker et al. 1993; Sivaraia et al. 1994], and a crystal structure of Oct-1 POU domain bound to the octamer motif as a monomer [Klemm et al. 1994].

To gain a better molecular understanding of POU domain flexibility in terms of their abilities to multimerize and adopt alternative configurations, we undertook a high-resolution X-ray analysis of Pit-1 POU domain bound to DNA as a homodimer. The structure, determined at 2.3Å resolution, reveals a remarkable alteration in the configuration of the individual POU domains. Compared with Oct-1, the Pit-1 POU-specific domain has not only completely flipped its orientation relative to the homeodomain but has also altered its spacing. As a result, the POU-specific domain and the homeodomain no longer bind to opposite sides of the DNA as in Oct-1, but to perpendicular faces of the DNA. The structure also reveals the dimerization interface for the first time, which, surprisingly, turns out to be the site of mutational hot spot in Pit-1 associated with combined pituitary hormone deficiency (CPHD) in humans [Cohen et al. 1995]. The structure provides a framework for understanding the functional specificity of Pit-1. Above all, the structure provides a dramatic illustration of the flexibility of POU domain proteins and their versatility as transcription factors.

### Results

#### Structure determination

The rat Pit-1 POU domain was overexpressed in Escherichia coli and purified from inclusion bodies [Jacobson et al. 1996]. The DNA fragment is very similar to the prolactin response element Prl-1P with 10 of the 12 central base pairs, over which the POU domains contact DNA, being identical [Fig. 1]. The crystallographic asymmetric unit contains two Pit-1 POU domains bound to the 28-bp DNA fragment as a dimer. The structure was solved by the multiwavelength anomalous diffraction (MAD) method using brominated DNA oligomers [Hendrickson 1991]. The method worked exceptionally well despite a PI space group and a fairly large crystallographic asymmetric unit. The final model consists of two homeodomains (POU$_{H1}$ and POU$_{H2}$), two POU-specific domains (POU$_{S1}$ and POU$_{S2}$), and 24 DNA base pairs.

As in the Oct-1/DNA complex [Klemm et al. 1994], the linker[s] connecting the POU-specific domain to the homeodomain is disordered in our complex, indicative of its flexibility. Because the linker is disordered, a question arises as to which POU-specific domain is connected to which homeodomain. The distance between the end of POU$_{S1}$ and the beginning of POU$_{H1}$ is ~20Å, which is easily bridgeable by the short 15-amino-acid Pit-1 linker. [In Oct-1, the 24-amino-acid linker bridges a distance of ~28Å]. The alternative, that POU$_{S1}$ is connected to POU$_{H2}$, a distance of ~42Å, would require the linker to adopt a much more extended and rigid structure. The connection between POU$_{S1}$ and POU$_{H1}$ was confirmed by a series of biochemical experiments that are described below.

#### Overall arrangement

The complex reveals an arrangement of POU domains that was completely unexpected. We expected each Pit-1 monomer to recognize DNA in a similar manner as Oct-1, with the homeodomain and the POU-specific domain binding to opposite faces of the DNA [Klemm et al. 1994]. Therefore, it was a surprise when we discovered that the POU-specific domain has not only flipped its orientation relative to the homeodomain but has also altered its spacing [Fig. 2]. As a result, the POU-specific domain and homeodomain no longer bind to opposite sides of DNA but to perpendicular faces of the DNA [Fig. 2]. We also expected the dimerization contacts to be mediated solely between the POU-specific domains [Ingram et al. 1990]. Instead, the dimerization interface is formed between the homeodomain of one monomer and the POU-specific domain of the symmetry related monomer. The last few residues of the recognition helix of the homeodomain are unraveled to form the dimer contacts. Intriguingly, the dimerization interface also turns out to be the site of a mutation (R271W) commonly associated with CPHD in humans [Radovick et al. 1992; Cohen et al. 1995].

The Pit-1 dimer interacts with DNA over a remarkably short stretch of sequence [Fig. 2]. The majority of contacts to DNA base pairs are localized to just the central 12 bp of the site. The four subdomains insert their recognition helices into adjacent major grooves, leading to an appearance of almost surrounding the DNA [Fig. 2B]. Because of the compact binding, it is not surprising
Figure 1. (A) Comparison of Pit-1 and Oct-1 POU domain sequences. The Pit-1 POU domain begins at amino acid 124 and the Oct-1 POU domain begins at amino acid 303. Numbering in the context of the whole proteins is used for the description of mutations. To ease description of the structure, we have also numbered the POU-specific domain and the homeodomain independently. Shown below the sequences are the relative locations of the α-helices. The dashed box at the end of helix a3 indicates residues that are unraveled in Pit-1.

(B) Sequence of the 28-bp DNA fragment used in cocrystallization studies and its match to the prolactin response element Prl-1P. In bold are shown the central 12 bp over which the majority of the base-pair contacts occur.

that the protein-DNA interactions are highly overlapping, with some of the base pairs recognized simultaneously in both the major groove and the minor groove. A number of the protein-DNA interactions we observe are distinct from the Oct-1 complex [Klemm et al. 1994], reflecting flexibility at the level of individual amino acids.

POU-specific domains and DNA contacts

The POU-specific domain consists of a cluster of four α-helices, with helices a2 and a3 forming the helix-turn-helix [HTH] motif of the domain [Fig. 2B]. The substructure is strikingly similar to the DNA-binding domains of phage 434 and λ repressors and 434 Cro, as noted previously for Oct-1 [Assa-Munt et al. 1993; Dekker et al. 1993]. This similarity even extends to many of the protein–DNA contacts (see below). The most notable difference in structure in comparing the two Pit-1 POU-specific domains with each other and with Oct-1, is in a segment beginning in the middle of helix a2 and continuing into the turn preceding helix a3 [amino acids 33-41].

In the Oct-1/DNA complex, the POU-specific domain recognized a GCAT DNA subsite [Fig. 2A] [Klemm et al. 1994]. The corresponding subsite in our complex is GTAT; but in an opposite alignment, ATAC, to match the flipped orientation of the POU-specific domain [Fig. 2C]. The DNA base pairs are contacted exclusively in the major groove. A number of these interactions are different from the Oct-1 structure, reflecting the single-base pair difference between the two subsites (GTAT versus GCAT). The common interactions are the bidentate hydrogen bonds joining Gln-44, the first residue of the recognition helix a3, to the adenine residue of the fourth base pair of the subsite, and a hydrogen bond linking Thr-45 to the thymine residue of the third base pair [Fig. 3A]. The most striking difference occurs in the interactions involving Arg-49. In the Oct-1 structure, Arg-49 lies in between the guanines of the first and second base pairs of the subsite, and donates hydrogen bonds to the O6 atoms of both guanines [Klemm et al. 1994]. In our structure, Arg-49 adopts a different configuration and donates both hydrogen bonds to the guanine of the first base pair [Fig. 3]. The hydrogen bonds are donated in a bidentate fashion to the O6 and N7 atoms of the receiving guanine. Therefore, the same residue adopts different configurations in the two complexes in order to make alternative protein–DNA contacts. The second difference relates to the role of Thr-45. In our structure, Thr-45 appears to be donating a hydrogen bond (via a shared proton) to the thymine of the second base pair, whereas in the Oct-1/DNA complex it received a hydrogen bond from a cytosine at that position [Fig. 3]. In our complex, specificity for a thymine is further conferred by van der Waals contacts between its methyl group and the Cβ methylene groups of Thr-45 and Ser-43. It appears, from a comparison of the two crystal structures, that the POU-specific domain is remarkably flexible in adapting to variations in sequence within the subsite.

The HTH motif is anchored onto DNA via extensive contacts to the sugar-phosphate backbone. Residues Arg-
Structure of Pit-1 POU domain homodimer

Figure 2. [A] A schematic overview of Pit-1 and Oct-1 DNA complexes. The Pit-1 POU domains are bound to DNA as a homodimer, whereas the Oct-1 POU domain is bound to DNA as a monomer [Klemm et al. 1994]. The two Pit-1 POU domains are shown in salmon and blue colors, while the single Oct-1 POU domain is shown in salmon color. Arrows indicate the relative orientations of the individual POU-specific and homeo subdomains. [B] The Pit-1-DNA complex viewed approximately down the noncrystallographic twofold axis. The broken lines show the putative flexible linkers connecting the POU-specific domains (POU_{S1} and POU_{S2}) to the homeo domains (POU_{H1} and POU_{H2}). The dimer interface is formed between the homeo domain of one monomer and the POU-specific domain of the other monomer. The figure was drawn with the aid of program SETOR [Evans 1993]. [C] A comparison of Pit-1 and Oct-1 POU domains. Compared with Oct-1 [yellow color], the Pit-1 POU-specific domain [salmon color] is both inverted in orientation and shifted to a different face of the DNA.

20, Thr-26, Cys-48, and Asn-52 form hydrogen bonds with the phosphate groups through their side chains, whereas Gln-27 forms hydrogen bonds through both its main chain NH group and its side chain NH\textsubscript{2} group. Contacts to the other side of the major groove are provided by residues Ser-43 and Thr-46 that precede and lie...
Figure 3. Protein–DNA interactions. [A] A view along the recognition helix of a Pit-1 POU-specific domain [POU\textsubscript{s1}], showing residues Ser-43, Gln-44, Thr-45, and Arg-49 interacting with DNA base pairs in the major groove. Hydrogen bonds are shown as dotted lines. [B] A view along the recognition helix of a Pit-1 homeo domain [POU\textsubscript{H1}] in the major groove, showing residues Asn-51 and Gln-54 making direct hydrogen bonds to DNA base pairs, and Arg-46 making a water-mediated hydrogen bond. Val-47 and Cys-50 make van der Waals contacts with the base pairs. [A] and [B] were drawn with the aid of program GRASP [Nicholls 1993]. [C] A sketch comparing base pair contacts in Pit-1 and Oct-1 DNA complexes. Contacts occur in both the major and minor grooves; hydrogen bonds are shown by solid arrows and nonpolar interactions by broken arrows. The two Pit-1 monomers are shown in salmon and blue colors, respectively. The DNA bases are colored according to the color of the subdomain making the major groove contacts. Note the highly overlapping nature of these contacts in Pit-1.

within the recognition helix, respectively, and residues Ser-56, Asn-59, and Lys-62 from helix a4. Overall, the DNA backbone contacts are similar to those observed in the Oct-1/DNA complex [Klemm et al. 1994], suggesting that the two POU-specific domains dock onto DNA in a similar manner—despite their opposite orientations.
As noted previously, a number of the residues and protein-DNA contacts are strikingly similar to the phage 434 and λ repressors and 434 Cro protein (Assa-Munt et al. 1993; Klemm et al. 1994). Residues Arg-20, Gln-27, Gln-44, and Glu-51 are conserved among all POU-specific domains and have their counterparts in the 434 proteins. Surprisingly, our structure reveals additional similarities to the prokaryotic proteins. Pit-1 contains an asparagine residue (Asn-52 vs. Ala-52 in Oct-1) that overlaps with an asparagine in 434 and λ repressors and 434 Cro and makes a geometrically equivalent hydrogen bond to a phosphate group (Pabo et al. 1990). The DNA backbone contacting residues Ser-43 and Thr-46 also bond to a phosphate group (Pabo et al. 1990). The DNA backbone contacts points to a similar mode of docking in the two classes. It has been argued that the domain may represent a particularly successful evolutionary design for presenting the HTH motif (Assa-Munt et al. 1993). The conservation of many of the DNA backbone contacts points to a similar mode of docking in which the first, second, and sixth residues along the recognition helix are in positions to recognize a 4-bp subsite in the major groove. The two classes of proteins appear, however, to have evolved different mechanisms for achieving high-affinity DNA binding. The prokaryotic proteins achieve high affinity through dimerization (Harrison and Aggarwal 1990), whereas the POU-specific domain achieves high affinity via a linker attachment to a second DNA-binding domain.

**Homeodomains and DNA contacts**

As expected, the homeodomain contains three α-helices and an amino-terminal arm that fits into the DNA minor groove [Fig. 2B]. Helices α2 and α3 form the HTH motif of the homeodomain, with helix α3 [also known as the recognition helix] interacting with base pairs in the major groove of the DNA. The Pit-1 homeodomain resembles other homeodomain–DNA crystal structures [Kissingler et al. 1990; Wolberger et al. 1991; Hirsch and Aggarwal 1995; Li et al. 1995; Wilson et al. 1995], except for a striking deviation at the carboxy-terminal end of the recognition helix. The last four residues of the recognition helix are in an unraveled or extended conformation that is closer to the NMR structures of some homeodomains. This unraveling allows for the formation of a dimerization interface in our complex. Flexibility for this portion of the recognition helix has been indicated by NMR studies of Antp and Ftz homeodomains [Qian et al. 1989, 1994], as well as isolated POU homeodomains [Sivaraja et al. 1994; Cox et al. 1995; Morita et al. 1995]. Strikingly, the conformation we observe is similar to the conformation at the end of helix α3 of the Oct-2 POU homeodomain [Sivaraja et al. 1994]. The role of this flexibility has been somewhat ambiguous. Our results suggest that the inherent flexibility of this region may be important in allowing protein–protein interactions with partner proteins.

The traditional way to view homeodomain–DNA interactions is with respect to an ATTA [or TAAT on the complementary strand] core motif, in which the first two base pairs are contacted in the major groove and the last two base pairs are contacted in the minor groove (Gehring et al. 1994). The matching sequence in our oligonucleotide is ATGT [Figs. 2 and 3]. Some of the major groove contacts such as the bidentate hydrogen bonding between Asn-51 and an adenine of the second base pair of the motif and hydrophobic interactions between Val-47 and the methyl group of a thymine of the first base pair of the motif, are characteristic of almost all homeodomains [Fig. 3]. Asn-51 is invariant among all homeodomains and Ile/Val-47 is strongly conserved [Burglin 1994]. However, we also observe a new interaction that appears to be unique to POU homeodomains. We find Gln-54 reaching back towards the major groove in order to make bidentate hydrogen bonds with an adenine that is the Watson–Crick mate of the thymine contacted by Val-47 [Fig. 3B]. Gln-54 is invariant among homeodomains of the POU family [Burglin 1994]. This interaction with an adenine provides an extra level of specificity for an AT sequence in the major groove. Specificity for base pairs preceding the AT sequence is conferred by Cys-50, which makes van der Waals contacts with the methyl group of a thymine at position –2, and Arg-46 which forms a water mediated hydrogen bond [Fig. 3].

Contacts in the minor groove arise primarily from a single residue, Arg-5, from the amino-terminal arm. In POU111, Arg-5 appears to be in a position to donate hydrogen bonds to four acceptor atoms on three consecutive base pairs [Fig. 3C]. In POU122, the configuration of Arg-5 differs somewhat from POU111, but it is still in a position to donate hydrogen bonds to two consecutive AT base pairs [Fig. 3C]. It is best to consider the minor groove contacts as tentative at present because the electron density for Arg-5 is rather diffuse in both homeodomains, perhaps indicating more than one configuration.

As with the POU-specific domain, contacts to the DNA backbone bracket the major groove containing the recognition helix α3. The side chains of residues Ser-28, Arg-46, and Arg-53 make direct hydrogen bonds with phosphate groups to one side of the groove, whereas both the side chain and main chain atoms of residue Thr-6 form hydrogen bonds to a phosphate group on the other side. Many of the DNA backbone contacts are conserved between Pit-1 and Oct-1, suggesting that the homeodomains dock onto DNA in a similar manner, despite variations in DNA core sequence (TCATGT versus TTATTT) and the presence and absence of dimer contacts in the two structures, respectively [Fig. 3].

**Dimer interface**

The dimer interface is formed between the homeodomain of one monomer and the POU-specific domain of the symmetry related monomer. This was unanticipated because previous studies had hypothesized the POU-spe-
sific domain as the sole determinant of dimerization [In- 
graham et al. 1990]. The interface is formed by the in- 
teraction of residues at the carboxy-terminal end of the 
recognition helix in the homeodomain with residues at 
the amino-terminal end of helix α1 and the loop between 
helices α3 and α4 in the POU-specific domain (Fig. 4). The 
interacting surfaces are remarkably complementary 
in terms of both hydrogen bonding and hydrophobic con- 
tacts. Two of the intermolecular hydrogen bonds are 
formed between the main chain NH groups of residues 
Val-59 and Lys-60 in the homeodomain and the main 
chain CO groups of residues Leu-53 and Leu-55 in the 
POU-specific domain. A third hydrogen bond is estab- 
lished between Nη1 of Arg-58 of the homeodomain and 
Oε1 of Gln-54 of the POU-specific domain (Fig. 4). The 
hydrogen bonds are augmented by a set of hydrophobic 
interactions attributable to Val-59 from the homeodo- 
mamin inserting into a small hydrophobic pocket on the 
surface of the POU-specific domain. The pocket is lined 
by nonpolar residues, including Met-6 and Leu-9 from 
the amino-terminal portion of helix α1 and Leu-53 and 
Phe-57 from the loop between helices α3 and α4 (Fig. 4). 
Lys-60, the last residue of the homeodomain, appears to 
be well positioned to reach over and form a salt link with 
Glu-10 at the amino-terminal end of helix α1 in the 
POU-specific domain, but its electron density is pointing 
avay from Glu-10 in one of the homeodomains and is 
defined poorly in the other. The dimerization contacts 
are only feasible because of the unraveling of the recog- 
nition helix in the homeodomain. Interestingly, the 
complementary segment on the POU-specific domain, 
the loop between helices α3 and α4, undergoes relatively 
little conformational change on dimerization.

Pit-1 exists as a monomer in solution and only forms 
cooperative dimers on specific DNA elements [Ingraham 
et al. 1990]. There are two equivalent dimerization in-
terfaces in our complex because of the twofold symme-
try of the complex. Each interface loses ~600 Å² of sol-
vent accessible surface area. The dimerization contacts 
can be compared with those observed in the DNA com-
plexes of Drosophila Paired homeodomain homodimer 
and the yeast Mata2/Mata1 homeodomain heterodimer 
[Li et al. 1995, Wilson et al. 1995]. In the Paired complex, 
the amino-terminal arm and the amino terminus of helix 
α3 of one homeodomain contact the first few residues of 
helices α2 and α3 of the twofold related homeodomain. 
The interface loses ~1000 Å² of solvent accessible sur-
face area, giving rise to an estimated cooperativity of 
~180-fold [Wilson et al. 1995]. Analogous to the Pit-1 
complex, the interacting surfaces are complementary in 
terms of both polar and nonpolar interactions. In con-
trast, the interface in the Mata2/Mata1 heterodimer is 
predominately hydrophobic, resulting in a total buried 
surface area of 754Å² [Li et al. 1995]. Mata2 contains a 
carboxy-terminal tail, immediately following the ho-
meodomain, that becomes ordered in the presence of 
Mata1 to assume a helical conformation. Interestingly, 
in the Mata2/Mata1 complex it is the ordering of the 
disordered carboxy-terminal arm that is crucial to het-
 erodimer formation, whereas in our structure it is the 
unraveling of a helical segment. Taken together, the 
structures illustrate the different mechanisms by which 
monomeric homeodomain containing proteins can 
dimerize on DNA.

**DNA conformation**

The DNA retains an overall B-DNA like conformation, 
lacking any major bends or kinks [Fig. 5]. The average 
helical parameters are 33.7° for twist and 3.38 Å for rise 
[Lavery and Sklenar 1988]. The DNA is markedly 
straighter than in the Oct-1/DNA complex, possessing a 
net bend of −7° as opposed to −30° [Klemm et al. 1994]. 
This difference in DNA bending provides an elegant 
mechanism for understanding the differential effects of 
phosphorylation in Pit-1 and Oct-1 [Fig. 5] (see Discus-
sion).

---

**Figure 4.** A stereo view of the dimer inter-
face, showing the unraveled recognition heli-
x of the homeodomain of one monomer (salmon color) coming in close contact with 
the POU-specific domain of the other monomer (blue color). Val-59 penetrates a 
small hydrophobic cavity on the surface of the POU-specific domain. In addition, three 
intermolecular hydrogen bonds are established (yellow broken lines); two of which are 
between main chain atoms and the third is between Arg-58 and Gln-54. Arg-58 has been identified as mutational hot spot in combined pituitary hormone deficiency in humans [Ohta et al. 1992; Radovick et al. 1992; Cohen et al. 1995].
Structure of Pit-1 POU domain homodimer

Correlation with biochemical data

To test our structural model we undertook a series of biochemical experiments. The first set of experiments were designed to test the inverse orientation model in which POU_{H1} is connected to POU_{S1} and POU_{H2} is connected to POU_{S2}. Because the flexible linker is invisible in our electron density we could not rule out the alternative that POU_{H1} is connected to POU_{S2} [Fig. 2B]. However, from structural arguments alone, this appeared to be an unlikely possibility because it would have required the linker to span a much longer distance of ~42Å. (The short Pit-1 linker can barely traverse this distance by adopting a highly extended and rigid structure.) To test which model was correct, we deleted 3 and 5 amino acids from the Pit-1 linker. Our reasoning was that in the inverse orientation model, the linker would be of sufficient length to accommodate the deletions, whereas in the alternative model, the linker could not accommodate the loss of residues, and binding would be highly sensitive to the deletions. Consistent with the inverse orientation model, the deletions did not affect the ability of Pit-1 to bind to the 28-bp DNA fragment, or to the prolactin Prl-1P and the growth hormone GH-1 DNA response elements [Fig. 6B]. As an additional test, we made two mutant versions of the 28-bp DNA fragment, the first of which contained mutations in the putative POU_{H2} and POU_{S2} subsites, whereas the second contained mutations in the POU_{H2} and POU_{S1} subsites. If POU_{H1} is connected to POU_{S1}, then Pit-1 would still be able to bind as a monomer to the first mutant oligomer but fail to bind the second oligomer. Opposite results were expected if the connection was otherwise. We obtained results that were again consistent with the inverse orientation model [Fig. 6C].

One of the surprising consequences of the inverse orientation model is the dimerization interface in which the homeodomain of one monomer interacts with the POU-specific domain of the twofold related monomer. Intriguingly, the R271W mutation causing combined pituitary hormone deficiency in humans also maps to this same interface. To test this interface in solution, we mutated the last three amino acids at the carboxy-terminal end of the homeodomain recognition helix and a glutamine residue in the loop between helices 3 and 4 of the POU-specific domain to alanines, and measured the binding of the mutated protein to the 28-bp DNA element as well as the natural Prl-1P and GH-1 response elements. In a striking correlation with the crystal structure, the mutant protein lost its ability to dimerize on all three DNA elements [Fig. 6D]. The experiment lends further support to the connectivity between POU_{H1} and POU_{S1}, and establishes a viable dimer interface in a POU domain protein. We also undertook a cotransfection assay to address whether the 28-bp fragment in our crystals can function as a response element in vivo. The transcription activation levels directed by the 28-bp DNA fragment were found to be comparable with those achieved with the Prl-1P site [Fig. 6A].

Together, our experiments have suggested that Pit-1 may adopt a similar dimer configuration when bound to the 28-bp DNA fragment, the Prl-1P element, or the GH-1 response element. However, there must also be subtle differences, because a single substitution at the dimerization interface, corresponding to the R271W CPHD mutation, was found to have a substantial effect on dimerization on the Prl-1P response element, while having little effect on dimerization on the 28-bp and the GH-1 elements. The exploration of these subtle differences will require a knowledge of the structure of Pit-1 bound to different elements.

Discussion

POU domain proteins constitute a growing family of transcription factors that regulate the expression of developmental and tissue-specific genes in a wide range of eukaryotic organisms (Wegner et al. 1993; Herr and Cleary 1995). A characteristic of POU domain proteins is their extraordinary versatility in recognizing diverse DNA elements. In this article, we have described a high-
resolution structure of Pit-1 POU domain bound to DNA as homodimer, revealing an arrangement of POU domains that was completely unexpected. Compared with Oct-1 (Klemm et al. 1994), the Pit-1 POU-specific domain has not only completely flipped its orientation relative to the homeodomain but has also altered its spacing. As a result, the POU-specific domain and the homeodomain no longer bind to opposite sides of the DNA as in Oct-1, but rather to perpendicular faces of the DNA. The structure also provided an initial visualization of a POU domain dimer interface, surprisingly formed between the homeodomain of one monomer and the POU-specific domain of the twofold related monomer. Mutating some of the key residues at this interface to alanines is sufficient to abolish the dimerization of Pit-1 on natural response elements Prl-1P and GH-1. The interface is only possible because of the unraveling of the last few residues of the recognition helix of the homeodomain, fully consistent with long-standing NMR measurements that indicated flexibility at the carboxy-terminal end of the recognition helix of homeodomains. The role of this flexibility, however, has been uncertain. We suggest that this flexibility is important in allowing homeodomains to interact with each other as well as with partner proteins. Visually, one of the most striking features of our complex is the compact binding, in which four DNA-binding domains are accommodated over a remarkably short stretch of DNA sequence. Not surprisingly, many of the contacts to the DNA are highly overlapping, with some of the base pairs recognized simultaneously in both the major and minor grooves.

**POU domain flexibility and transcription**

The Pit-1 complex brings into focus several key mechanisms for understanding POU domain flexibility. First, the alteration in both spacing and orientation of the POU-specific domain relative to the homeodomain appears to be a function of both dimerization and the fact that Pit-1 has the shortest linker (15 amino acids) among known POU domain proteins. Because up to 5 amino acids can be removed from the Pit-1 linker without any adverse effects on binding to the prolactin Prl-1P or growth hormone GH1 response elements, the full 15-amino-acid linker of Pit-1 may only become necessary when it samples other configurations on other response elements. Interestingly, the neuronal POU domain pro-
tein Brn-2 can accommodate different spacings (n=0, 2, and 3 nt) between its subdomain binding sites. For these spacings, we find a striking correlation between DNA-binding affinities and the distances between the end of POU$_5$ and the beginning of POU$_H$. The distances of 29 Å, 20 Å, and 24 Å for n=0, 2, and 3, respectively, are related inversely to the DNA-binding affinities of Brn-2, suggesting that the relative strain imposed on the 17-amino-acid linker may be an important source of differences in affinities for different spacings (Li et al. 1993).

The second mechanism highlighted by our work is the diversity of side chain–base pair interactions. The DNA subsite [ATAC] contacted by POU$_5$ in our structure differs by a single base-pair from the subsite [ATGC] contacted by the Oct-1 POU$_5$ (Klemm et al. 1994). This causes residue Arg-49 to adopt different configurations in the two structures. In Oct-1, Arg-49 lies in between base pairs 3 and 4 of the Oct-1 subsite and hydrogen bonds to adjacent guanines (Klemm et al. 1994). In contrast, Arg-49 adopts a different arrangement in our complex, enabling it to donate both hydrogen bonds to the guanine of the fourth base pair. Therefore, the same residue appears capable of adopting different configurations, allowing it to recognize more than one set of base pairs, consistent with the mutation data of Cleary and Herr (1995). Alternate side chain interactions have also been observed in the structures of the DNA-binding domains of phage 434 repressor and the estrogen receptor (Rodgers et al. 1993; Schwabe et al. 1995). Therefore, the recognition of alternate DNA sites through alternate side chain rearrangements is a common theme extending through several of classes of transcription factors.

The third key mechanism for understanding POU domain flexibility relates to their ability to bind DNA as both monomers and homo- and heterodimers (Kemler et al. 1989; LeBowitz et al. 1989; Poellinger and Roeder 1989; Ingraham et al. 1990; Voss et al. 1991; Holloway et al. 1995). Pit-1 loses ~1200 Å$^2$ of solvent accessible surface area attributable to dimerization on DNA. This value is well below the values observed with most oligomeric proteins (Janin et al. 1988), consistent with Pit-1 existing as a monomer in solution. On DNA binding, however, the dimer contacts will mark a major contribution to the cooperativity and the complex (Ingraham et al. 1990). Dimerization will be favored on DNA elements that contain appropriately positioned half-sites. One “appropriate” arrangement, based on our structure, is that allowing the carboxy-terminal end of the recognition helix of homodomains to interact with a complementary surface on the POU-specific domain of the two-fold related monomer. However, it is likely that POU domains have evolved more than one type of dimerization interface. Overall, the dimerization interface in our complex appears to be fairly promiscuous in allowing pairings between different POU domain partners. Some of the key residues such as Arg-58 of POU$_5$ and Gln/Asn-54 of POU$_5$ are well conserved, and residue 59 of the POU$_H$ tends to be hydrophobic. Moreover, two of the hydrogen bonds at the interface are formed between main chain rather than side chain atoms.

The Pit-1 complex challenges the more traditional notion of a transcription factor binding DNA with a unique configuration. The various mechanisms of flexibility may expand the repertoire of DNA sequences recognized and the number of genes regulated by POU domain proteins. The various Pit-1 response elements on a variety of target genes (e.g., Bodner et al. 1988; Ingraham et al. 1988; Li et al. 1990; Lin et al. 1992; Rhodes et al. 1993) show substantial variation in sequence, raising the possibility that Pit-1 may be configured differently on different DNA sites, as may be the case for Oct-1 (Baurucker et al. 1988; apRhys et al. 1989). This flexibility, achieved through alterations in orientation, spacing, and multimerization, may enhance the range of cooperative interactions between POU-domain proteins and cofactors. In the case of Oct-1, its ability form a complex with the acidic activator VP16 has been shown to vary depending on the nature of the DNA site (Cleary and Herr 1995). Similarly, Oct-1 and Oct-2 have been shown to interact with the activator OCA-B but only on a subset of octamer containing sites (Luo and Roeder 1995; Gstaiger et al. 1996). Likewise, interactions with distinct putative coactivators may occur when Pit-1 is bound as a monomer or as a dimer to its specific DNA response elements (Holloway et al. 1995). Therefore, the same POU-domain protein may be used in different contexts on different DNA elements to achieve gene-specific action.

**CPHD**

Many of the mutations causing cPHD in humans have been mapped to the Pit-1 POU domain, causing combined deficiencies in prolactin, growth hormone, and thyroid stimulating hormone, with the patients often being characterized by mental and growth retardation. Figure 7 shows the mapping of these mutations onto the Pit-1 POU domain. Some of these mutations can be understood in terms of their negative effects on DNA binding. For instance, the nonsense mutations R172X and E250X (using a numbering scheme based on the whole protein, where X stands for a stop codon) cause deletions of protein segments important for DNA binding (Tatsumi et al. 1992; Irie et al. 1995). The missense mutation R143Q substitutes a critical arginine residue with a glutamine (Ohta et al. 1992). Arg-143 has a key dual role in our structure in anchoring Pit-1 to the negatively charged DNA backbone and in the folding of POU-specific domain through hydrophobic and electrostatic interactions.

A more puzzling mutation is A158P, which maps to helix $\alpha$2 of the POU-specific domain (Pfaffle et al. 1992). The mutation disrupts the ability of Pit-1 to activate transcription from prolactin and growth hormone DNA response elements, but has only a minimal effect on DNA binding (Pfaffle et al. 1992). Intriguingly, the mutation is located in a region (amino acids 33–41) showing the most variation in structure, when comparing the Pit-1 POU-specific domains with each other and with Oct-1. Although the exact function of this region is un-
known, the inherent flexibility may be important in mediating protein-protein interactions. Interestingly, mutating single amino acids in this region generates mutants of Pit-1 with dramatically different affinities for the TATA-binding protein (TBP) associated factor TAF40 (T. Heinzel and M.G. Rosenfeld, unpubl.). This potential linkage between region 33-41 and TAF40 is intriguing because the analogous region in A. repressor is the site of several positive control mutants (Ptashne 1992). Hence, like its prokaryotic counterpart, region 33-41, encompassing the A158P CPHD mutation, may also have a role in the recruitment of cofactors and in trans-activation.

The most frequent CPHD mutation is R271W (Ohta et al. 1992; Radovick et al. 1992; Cohen et al. 1995). The mutation maps near the carboxy-terminal end of the homeodomain DNA recognition helix and again displays only a minimal effect on DNA binding (Radovick et al. 1992). Unexpectedly, our structure reveals Arg-271 to be involved in the formation of the Pit-1 dimer interface, making an intermolecular hydrogen bond with a glutamine residue on the POU-specific domain of the twofold related monomer. Therefore, it is possible to view the R271W mutation in terms of its affect on the ability of Pit-1 to dimerize on various DNA response elements. We find that this mutation has a substantial effect on dimerization on the Prl-1P element, implying consequences on a subset of target genes.

DNA conformation and phosphorylation

A comparison of Pit-1 and Oct-1 DNA conformations suggests a mechanism for understanding their differential responses toward phosphorylation (Fig. 5). Both Pit-1 and Oct-1 are phosphorylated at residue 7 of the homeodomain. Pit-1 can be phosphorylated in vitro in response to cAMP-dependent protein kinase; whereas in vivo phosphorylation may occur during cell-cycle and/or in response to homeostatic regulators (Kapiloff et al. 1991; Caelles et al. 1995). Whereas binding to most sites is decreased by phosphorylation (Kapiloff et al. 1991; Caelles et al. 1995), detailed kinetic analysis revealed that binding was markedly decreased on some sites, but only minimally affected on others (Kapiloff et al. 1991). In our complex, Thr-7 is pointing away from the negatively charged DNA backbone, a distance of 6.6-7Å from the nearest phosphate group. Therefore, on those sites with the configuration observed in our complex, phosphorylation of Thr-7 would be expected to have little effect on Pit-1 DNA-binding activity. In contrast, Ser-7 in the Oct-1/DNA complex is only 3.6Å away from the DNA backbone (Klemm et al. 1994), consistent with ability of phosphorylation of this residue to disrupt binding (15- to 20-fold) to the H2B octamer motif, as part of the cell cycle regulation of the H2B gene (Segil et al. 1991). The variation in distance of residue 7 from the DNA backbone is primarily attributable to a difference in DNA bending in the two complexes (Fig. 5). The Pit-1 DNA is markedly straighter than the Oct-1 DNA, resulting in the splaying of DNA backbone away from residue 7. This suggests a model in which certain conformational and multimerization states of the POU domains may be more sensitive to effects of phosphorylation than others, through their effects on DNA conformation.

Materials and methods

Protein and DNA preparation

A fragment of the Pit-1 gene encoding the POU domain (amino acids 128-273) was subcloned into the poly-His vector pET14b and expressed in E. coli. Because of cloning constraints, amino acids 128 and 129 were changed to glycine and methionine from the naturally occurring residues glutamate and isoleucine (Fig. 1A). Purification of the protein has been described (Jacobson et al. 1996). The DNA fragments were synthesized on an Applied Biosystems 391 PCRmate automated synthesizer using phosphoramidite chemistry and purified through the use of PAGE.

Data collection

The crystallization and preliminary X-ray analysis of the complex has been described (Jacobson et al. 1996). The cocryystals belong to space group P1 with unit cell dimensions of a=42.5 Å, b=50.1 Å, c=55.8 Å, α=76.7°, β=79.3°, and γ=67.2°. For MAD phasing, we grew cocryystals with derivatized DNA oligomers containing bromouracils. Two such oligomers (Brl and
Structure of Pit-1 POU domain homodimer

Br2), each containing six bromouracils, yielded data quality crystals (Table 1).

All data measurements were performed at Cornell High Energy Synchrotron Source (CHESS) on frozen crystals (−160°C). The native data were recorded from two cocrystals, using the charge coupled device (CCD) detector at beamline A1 (λ=0.908Å). Diffraction data for MAD analysis were collected from a Br1 cocrystal, using imaging plates at beamline F2. X-ray fluorescence spectrums for the Br K-edge were found to be too noisy when measured directly from the crystal. Consequently, we used spectra measured from a bromine (NaBr) foil to select the monochromator settings for the different wavelengths. The data were measured at three wavelengths corresponding to the edge (0.9224Å) and peak (0.9223Å) of the Br K absorption profile, plus a remote point (0.9074Å). Because of the low symmetry of the P1 space group the Friedel pairs could only be recorded by the inverse beam method (θ and θ + 180°). The data were collected by the oscillation method (α=5°), with no overlap between successive frames. Inverse data were recorded after every 30° of rotation. The imaging plates were digitized using a Fuji BAS2000 scanner. By the end of the data collection (a total of 36 hr), the cocrystal had been rotated a complete 360°. In addition to the MAD data, we measured two more data sets near the peak wavelength (0.920Å). The first corresponded to another Br1 cocrystal, and the second was from a Br2 cocrystal that has four out of six bromines at different positions from Br1 (Table 1). All data sets were processed using the HKL software package [Molecular Structure Corp.]. The data collection statistics are shown in Table 1.

Bromine MAD phasing

The bromine positions were determined by inspection of anomalous difference Patterson maps. The maps were exceptionally clean and allowed easy identification of the six bromine sites for Br1 and Br2, despite the absence of Harker sections in space group P1. For phasing we used the program package PHASES (Furey 1993), and a total of five data sets: Br1-edge, Br1-peak, Br1-remote, Br1, and Br2 (Table 1). The Br1-edge data set was treated as the “native” in the phasing process. To incorporate Br2 data, which has bromines at different positions from the native, the scattering from the six Br1 sites was first subtracted, by attaching negative values to the scattering factors. All of the data sets contributed to the phasing through their anomalous differences (Table 1). The only isomorphous contribution came from Br1-remote, which had the largest f’ difference with respect to the native. Attempts were made to include isomorphous contributions from the other data sets but they led to an unstable refinement. The final figure of merit was 0.720 at 3Å resolution. The resulting electron density map showed clear positions for the DNA and the POU-specific domains and homeodomains. The map was further improved by cycles of averaging, using the noncrystallographic twofold axis of symmetry determined from the refined bromine positions.

Table 1. Data collection and phasing statistics

| Item                              | Native | Br1 edge | Br1 peak | Br1 remote | Br1  | Br2  |
|-----------------------------------|--------|----------|----------|------------|------|------|
| Wavelength (Å)                    |        | 0.91995  | 0.91962  | 0.9049     | 0.9193| 0.9193|
| Resolution (Å)                    | 2.3    | 2.7      | 2.7      | 2.7        | 2.7  | 2.7  |
| Unique reflections                | 14,575 | 9,914    | 10,021   | 9,926      | 6,810| 10,759|
| Data coverage (%)                 | 84.7   | 86.8     | 87.2     | 86.8       | 58.1 | 95.1 |
| Rmerge (%)                        | 6.8    | 5.2      | 5.8      | 6.0        | 5.1  | 5.0  |
| MAD phasing statistics:           |        |          |          |            |      |      |
| phasing power (isomorphous to 3Å) | 1.16   | 1.65     | 1.67     | 1.80       | 2.04 |    |
| phasing power (anomalous to 3Å)   |        |          |          |            |      |      |
| R-factor (6-2.3Å, %)              | 23.0   |          |          |            |      |      |
| R-free (6-2.3Å, %)                | 30.2   |          |          |            |      |      |
| nonhydrogen atoms                 | 3099   |          |          |            |      |      |
| Water molecules                   | 176    |          |          |            |      |      |
| rms bond lengths (Å)              | 0.013  |          |          |            |      |      |
| rms bond angles (°)               | 1.71   |          |          |            |      |      |

aMerging R factor (Rmerge = Σ|Iobs − ⟨I⟩|/|⟨I⟩|) calculated for all data.

bPhasing power for isomorphous differences is defined as ⟨|Fh+⟩/rms(e)⟩ where ⟨|Fh+⟩⟩ is the mean calculated amplitude for the heavy-atom model and rms(e) is the weighted root mean square lack of closure error.

Phasing power for anomalous differences is defined as ⟨|AFh+⟩/rms(e')⟩ where ⟨|AFh+⟩⟩ is the mean calculated Bijvoet difference from the heavy atom model and rms(e') is the weighted root mean square lack of closure for anomalous differences.

Table 2. DNA and POU-specific domain structure

| Item                              | Br1  | Br2  |
|-----------------------------------|------|------|
| Br1                               |      |      |
| 3'-G G A A G A G G U A C A U T A U A C A U G T A C U C T T C C-5' |      |      |
| Br2                               |      |      |
| 3'-G G A A G A G G U A C A T A U T A G U A C T T C C-5' |      |      |

GENES & DEVELOPMENT 209
The R-factor between observed and calculated structure factor amplitudes following map inversion converged to 0.187 from an initial value of 0.375. The averaged map showed clear densities for the DNA base pairs and the side chains of most amino acids. The map was further improved by the inclusion of solvent flattening, leading to a rise in the figure of merit to 0.898 at 3 Å resolution.

**Building and refinement**

Using the program O [Jones et al. 1991], copies of the Oct-1 POU-specific domain and the homeodomain were extracted from the Oct-1/DNA complex and manually docked into the experimental electron density map (3 Å). An idealized B-DNA oligomer was docked into the DNA densities, along the body diagonal of the unit cell. Even at this early stage, it was clear that the Pit-1 POU domains were bound to DNA in a radically different manner from the Oct-1/DNA complex. The POU-specific domains, the homeodomains, and the DNA were manually rebuilt to fit the densities better. Certain regions, such as the linkers connecting the POU-specific domains to homeodomains and base pairs at the ends of the DNA oligomer, were excluded from building because of the lack of experimental densities. The rebuilt model was first subjected to rigid body and positional refinement against the native data, at 3 Å resolution, using the program XPLOR (Brunger 1992). The refinement was then extended to 2.3 Å resolution, using all data within 6–2.3 Å resolution and $\frac{F_{o}}{F_{c}}-2\sigma(F_{c})$. The progress of refinement was monitored by $R_{free}$ calculations based on 5% of the data. An anisotropy correction was applied in later stages of the refinement, leading to a drop in the $R$-factor and $R_{free}$ values by ~1.3%. In addition, an extensive number of simulated annealing omit maps were calculated and used to confirm the model. A total of 176 water molecules were eventually added to the model, with the aid of an automated procedure. All water positions were verified by inspecting simulated annealing omit maps. Near the end of refinement when the $R$-factor was 23.8% and the $R_{free}$ was 30.2%, all reflections set aside previously for the $R_{free}$ calculation were used in the refinement and the R-factor further dropped to 23.0%.

**Gel shift**

Recombinant protein [1–100 ng] was incubated in reaction buffer (2 mM HEPES at pH 7.8, 60 mM KCl, 1 mM EDTA, 0.1% NP-40, 15% glycerol, 0.1 μg/μl Polh[dG][dC]), 25 ng/μl BSA, 5 mM DTT, 5 ng/μl salmon sperm DNA, and 0.1% milk) in the presence of $^{32}$P-labeled DNA probe [100,000 cpm] for 30 min at room temperature. After incubation, the samples were loaded onto a 6% acrylamide gel.

**Effect of linker length on Pit-1 dimerization**

Two mutant Pit-1 fragments [128–273] were made in which 5 or 3 amino acids were removed from the linker region (amino acids 203–207 and 204–206). The two mutant proteins did not show any significant change in their ability to bind 28-bp, GH-1, or Prr-1P elements supporting the inverse model.

**Dimerization experiments**

A recombinant Pit-1 fragment [128–273] in which amino acids Gln-177, Arg-271, Val-272, and Lys-273 were substituted by alanines was tested in its ability to dimerize on three DNA binding sites: 28-mer, GH-1, and Prr-1P. The mutant Pit-1 fragment (POU mutant) showed a decrease in its ability to dimerize on any of the DNA-binding sites tested. Because the mutant monomer band was found to run in an intermediate position to the dimer and monomer bands produced by the wild-type protein, the following experiment was done. The 4-amino-acid substitutions were introduced into full-length Pit-1 (Pit-1 mutant), and the protein mixed POU mutant. The gel shift band pattern produced by the mixture was the sum of the bands obtained with Pit-1 mutant or POU mutant alone, no new bands were produced.

**Cotransfection assays**

The cotransfection assays were performed essentially as described by Li et al. [1993]. Three copies of the 28-mer were inserted into the BamHI site of the p36-luc vector [Ingraham et al. 1988]. CMV–Pit-1 and Prr-1P-p36-luc were gifts from K. Scully (University of California at San Diego, La Jolla). One microgram each of CMV–Pit-1 expression plasmid and luciferase reporter plasmid were cotransfected into HeLa cells.

**Acknowledgments**

We thank the staff at CHESS (particularly B. Miller, J. Navie, and O. Shen) for assistance with the bromine MAD experiment. We thank C. Escalante for advice on protein expression and purification; J. Hirsch and D. Wahl for help with data collection; and later stages of the project, P. Kwong, V. Ramakrishnan and L. Shapiro for advice on computer programs; H. Vidiu for helpful suggestions, F. Hedjan for assistance with the biochemical assays. The work was supported by National Institutes of Health grants GM-49327 (A.K.A) and NIDDK-18477 (M.G.R.). M.G.R. is an investigator with Howard Hughes Medical Institute. Co-ordinates are being deposited in the Brookhaven Protein Database. They can also be obtained by email (Aggarwal convex. hhmi.columbia.edu).

**References**

Aggarwal, A.K., D.W. Rodgers, M. Drottter, M. Ptashne, and S.C. Harrison. 1988. Recognition of a DNA operator by the repressor of phage 434: A view at high resolution. Science 242: 899–907.

Aprikakis, C.M., D.M. Ciufio, E.A. O’Neill, T.J. Kelly, and G.S. Hayward. 1989. Overlapping octamer and TAATGA-RAT motifs in the FV65-response elements in herpes simplex virus immediate-early promoters represent independent binding sites for cellular Nuclear Factor III. /: Virology 63: 2798–2812.

Assa-Munt, N., R.I. Mortshire-Smith, R. Aurora, W. Herr, and P.E. Wright. 1993. The solution structure of the Oct-1 POU-specific domain reveals a striking similarity to the bacteriophage λ repressor DNA binding domain. Cell 73: 193–205.

Baumrucker, T., R. Sturm, and W. Herr. 1988. OBP100 binds remarkably degenerate octamer motifs through specific interactions with flanking sequences. Genes Dev: 2: 1400–1413.

Beexvanis, A.D. and C.R. Vinson. 1993. Interactions of coiled coils in transcription factors: Where is the specificity? Curr. Opin. Genet. Dev. 3: 278–285.

Bodner, M., J. Castrillo, L.E. Theill, T. Deerinck, M. Ellisman, and M. Karin. 1988. The pituitary-specific transcription factor GHFI is a homeobox-containing protein. Cell 55: 505–518.
Structural Research Institute, Yale University, New Haven, CT.

Burglin, T. 1994. A comprehensive classification of homeobox genes. In Guidebook to the homeobox genes (ed. D. Duboule), p. 27–71. Oxford University Press, Oxford, UK.

Caelles, C., H. Hennemann, and M. Karin. 1995. M-phase-specific phosphorylation of the POU transcription factor GHF-1 by a cell cycle-regulated protein kinase inhibits DNA binding. Mol. Cell. Biol. 15: 6694–6701.

Cleary, M.A. and W. Herr. 1995. Mechanisms for flexibility in X-PLOR, version 3.1, a system for X-ray crystallography and NMR. Yale University Press, New Haven, CT.

Brunger, A.T. 1992. X-PLOR, version 3.1, a system for X-ray crystallography and NMR. Yale University Press, New Haven, CT.

Gehring, W.J., M. Aifolter, and T. Burglin. 1994. Homeodomain proteins with the helix-turn-helix motif. Annu. Rev. Biochem. 63: 487–526.

Glover, J.N.M. and S.C. Harrison. 1995. Crystal structure of the heterodimeric bZIP transcription factor cFos-cJun bound to DNA. Nature 373: 257–261.

Hendrickson, W.A. 1991. Variable effects of phosphorylation of Pit-1 dictated by the DNA response elements. Science 253: 786–789.

Hess, 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.

Hess, 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 Caeorhabditis elegans unc-86 gene products. Genes & Dev. 2: 1513–1516.

Hirsh, J.A. and A.K. Aggarwal. 1995. Structure of the Even-skipped homeodomain complexed to AT-rich DNA: new perspectives on homeodomain specificity. EMBO J. 14: 6280–6291.

Ingraham, F.A., S.E. Flynn, J.W. Voss, V.R. Albert, M.S. Kaploff, L. Wilson, and M.G. Rosenfeld. 1990. The POU-specific domain of Pit-1 is essential for sequence specific, high affinity DNA binding and DNA dependent Pit-1-Pit1 interactions. Cell 61: 1021–1033.

Irie, Y., K. Tatsumi, M. Ogawa, T. Kamiio, C. Preeyasombat, C. Suprasongis, and N. Amino. 1995. A novel E250X mutation of the Pit1 gene in a patient with combined pituitary hormone deficiency. Endocrine J. 42: 351–354.

Jacobson, E.M., P. Li, M.G. Rosenfeld, and A.K. Aggarwal. 1996. Crystalization and preliminary X-ray analysis of Pit-1 POU domain complexed to a 28 base pair DNA element. Proteins 24: 263–265.

Janin, J., S. Miller, and C. Chothia. 1988. Surface, subunit interfaces and interior of oligomeric proteins. J. Mol. Biol. 204: 155–164.

Jones, A.T., J.Y. Zou, S.W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A47: 110–119.

Kaploff, M.S., Y. Farkash, M. Wegner, and M.G. Rosenfeld. 1991. Variable effects of phosphorylation of Pit-1 dictated by the DNA response elements. Science 253: 786–789.

Kemler, I., E. Schreiber, M.M. Muller, 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.

Kissinger, C.R., B. Liu, E. Martin-Blanco, T.B. Kornberg, and C.O. 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.

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

Kurokawa, R., Y. Yu, A. Naar, S. Kyakumoto, Z. Han, S. Silverman, M.G. Rosenfeld, and C.K. Glass. 1993. Different orientations of the carboxy-terminal dimerization interface determine binding site selection by nuclear receptor heterodimers. Genes & Dev. 7: 1423–1435.

Lavery, R. and H. Sklenar. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. J. Biomol. Struct. Dyn. 6: 63–91.

LeBowitz, J.H., R.G. Clerc, M. Benowitz, and P.A. Sharp. 1989. The Oct-2 protein binds cooperatively to adjacent octamer sites. Genes & Dev. 3: 1625–1638.

Li, P., X. He, M.R. Gerrerio, 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.

Lin, S., E.B. Crenshaw, E.J. Rawson, D.M. Simmons, L.W. Swanson, and M.G. Rosenfeld. 1990. Dwarf locus mutants lacking three pituitary cell types result from mutations in the POU domain gene pit-1. Nature 347: 528–533.

Lin, C.R., S.-C. Lin, C.P. Chang, and M.G. Rosenfeld. 1992. Pit-1 dependent expression of the receptor for growth hormone releasing factor mediates pituitary cell growth. Nature 356: 262–269.
Jacobson et al.

360: 765–768.
Luo, Y. and R.G. Roeder. 1995. Cloning, functional characterization, and mechanism of action of the B-cell specific transcriptional coactivator OCA-B. Mol. Cell. Biol. 15: 4115–4124.

Morita, E.H., M. Shirakawa, F. Hayashi, M. Imagawa, and Y. Kyogoku. 1995. Structure of the Oct-3 POU homeodomain in solution, as determined by triple resonance heteronuclear multidimensional NMR spectroscopy. Protein Sci. 4: 729–739.

Naar, A.M., J.M. Boutin, S.M. Lipkin, V.C. Yu, J.M. Holloway, C.K. Glass, and M.G. Rosenfeld. 1991. The orientation and spacing of core-DNA binding motifs dictate selective transcriptional responses to three nuclear receptors. Cell 65: 1267–1279.

Nicholls, A.J. 1993. GRASP: A graphical representation and analysis of surface properties. Columbia University, New York, NY.

Ohta, K., Y. Nobukuni, H. Mitsubuchi, S. Fujimoto, N. Matsuo, H. Ingaki, F. Endo, and I. Matsuda. 1992. Mutations in the Pit-1 gene in children with combined pituitary hormone deficiency. Biochem. Biophys. Res. Comm. 189: 851–855.

Pabo, C.O., A.K. Aggarwal, S.R. Jordan, L.J. Beamer, U.R. Obeyesekere, and S.C. Harrison. 1990. Conserved residues make similar contacts in two repressor-operator complexes. Science 247: 1210–1213.

Poellinger, L. and R.G. Roeder. 1989. Octamer transcription factors 1 and 2 each bind to two different functional elements in the immunoglobulin heavy-chain promoter. Mol. Cell. Biol. 9: 747–756.

Ptashne, M. 1992. A genetic switch. Cell Press & Blackwell Scientific Publications, Cambridge, UK.

Qian, Y.Q., M. Billeter, G. Otting, M. Muller, A.K. Aggarwal, S.R. Jordan, L.J. Beamer, U.R. Obeyesekere, and S.C. Harrison. 1990. Conserved residues make similar contacts in two repressor-operator complexes. Science 247: 1210–1213.

Radovick, S., M. Nations, Y. Du, L.A. Berg, B.D. Weinstein, and F.E. Wondisford. 1992. A mutation in the POU-homeodomain of Pit-1 is responsible for combined pituitary hormone deficiency. Science 257: 115–1118.

Rhodes, S.J., R. Chen, G.E. DiMattia, K.M. Scully, K.A. Kalla, S.C. Lin, V.C. Yu, and M.G. Rosenfeld. 1993. A tissue-specific enhancer contains Pit-1-dependent morphogen inducibility and autoregulation on the Pit-1 gene. Genes & Dev. 7: 913–931.

Rodgers, D.W. and S.C. Harrison. 1993. The complex between phase 434 repressor DNA-binding domain and operator site OR3: Structural differences between consensus and non-consensus half-sites. Structure 1: 227–240.

Schwabe, J.W., L. Chapman, and D. Rhodes. 1995. The oestrogen receptor recognizes an imperfectly palindromic response element through an alternative side-chain conformation. Structure 3: 201–213.

Sivaraja, M., M.C. Botfield, M. Mueller, A. Jancso, and M.A. Weiss. 1994. Solution structure of a POU-specific homeodomain: 3D NMR studies of human B-cell transcription factor Oct-2. Biochemistry 33: 9845–9855.

Tatsumi, K., K. Miyai, T. Notomi, K. Kaibe, N. Amino, Y. Mizuno, and H. Kohno. 1992. Cretinism with combined hormone deficiency caused by a mutation in the Pit-1 gene. Nature Genet. 1: 56–58.

Umesono, K., K.K. Murakami, C.C. Thompson, and R.M. Evans. 1991. Direct repeats as selective response elements for the thyroid-hormone, retinoic acid, and vitamin D3 receptors. Cell 65: 1255–1266.

Voss, J.W., L. Wilson, and M.G. Rosenfeld. 1991. POU domain proteins Pit-1 and Oct-1 interact to form a heteromeric complex and can cooperate to induce expression of the prolactin promoter. Genes & Dev. 5: 1309–1320.

Wegner, M., D.W. Drolet, and M.G. Rosenfeld. 1993. POU-domain proteins: Structure and function of developmental regulators. Curr. Opin. Cell Biol. 5: 488–498.

Wilson, D.S., B. Guenther, C. Desplan, and J. Kuriyan. 1995. High resolution crystal structure of a Paired (PAX) class co-operative homeodomain dimer on DNA. Cell 82: 709–719.

Wolberger, C., A.K. Veshon, B. Liu, A.D. Johnson, and C.O. Pabo. 1991. Crystal structure of a MATa2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. Cell 67: 517–528.
Structure of Pit-1 POU domain bound to DNA as a dimer: unexpected arrangement and flexibility.

E M Jacobson, P Li, A Leon-del-Rio, et al.

Genes Dev. 1997, 11:
Access the most recent version at doi:10.1101/gad.11.2.198