Structure, Vol. 13, 601–607, April, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.01.023

Structural Basis of Prospero-DNA Interaction: Implications for Transcription Regulation in Developing Cells

Mohammad S. Yousef[1,2] and Brian W. Matthews[1,*]
[1]Institute of Molecular Biology
Howard Hughes Medical Institute and
Department of Physics
1229 University of Oregon
Eugene, Oregon 97403
[2]The Biophysics Department
Faculty of Science
Cairo University
Giza
Egypt

Summary

The crystal structure of a complex between the novel homeodomain of the neural transcription factor Prospero and DNA shows that the invariant residues Lys1290, Asn1294, and Asp1297 make specific contacts with the noncanonical DNA binding site. The overall structure includes the homeodomain and the adjacent Prospero domain and confirms that they act as a single structural unit, a Homeo-Prospero domain. The Prospero domain facilitates the proper alignment of the protein on the DNA. Knowledge of the structure reconciles two different DNA sequences that have been proposed as transcriptional targets for Prospero. As in the apo structure, the C terminus of the Prospero domain shields a short helix within the homeodomain that includes a nuclear export signal (NES). The structural results suggest that exposure of the NES is not coupled directly to DNA binding. We propose a DNA recognition mechanism specific to Prospero-type homeodomains in developing cells.

Introduction

Prospero (Pros), a multidomain 153.5 kDa protein, is a divergent homeodomain transcription factor (Chu-Lagraff et al., 1991; Hirata et al., 1995; Matsuzaki et al., 1992). Pros is expressed in all of the neural lineages in *Drosophila* embryos. It is required for the proper expression of several nervous system proteins and in determining the cell fate of the stem cell (neuroblast progeny) in the central nervous system (Doe et al., 1991; Hassan et al., 1997). Pros is asymmetrically localized to the cytoplasmic membrane of neuroblasts and consequently segregated into one daughter cell, the ganglion mother cell (GMC), upon cell division. Pros then translocates from the cytoplasm of the GMC to the nucleus to initiate the differentiation toward either neurons or glial cells (Hirata et al., 1995; Spana and Doe, 1995). In fully differentiated neurons, Pros is degraded to undetectable levels (Vaessin et al., 1991).

The Pros family is widely distributed in developing cells from different organisms. Within the family, there is extensive sequence homology between *Drosophila* Pros and its *C. elegans*, chicken, mouse, and human homologs extending over the 160 C-terminal amino acids of the five proteins (Bürglin, 1994a, 1994b; Dudas et al., 2004; Oliver et al., 1993; Tomarev et al., 1996). Part of this well-conserved C-terminal region has a sequence that corresponds to an atypical homeodomain (HD). The remainder is identified as a Prospero domain (PD).

The homeodomain is one of the most important DNA binding motifs in eukaryotes and has provided a model system for studying protein-DNA interactions. The structure of the *Drosophila* engrailed homeodomain-DNA complex (Kissinger et al., 1990) was the first crystal structure to reveal how this motif recognizes DNA. Since then, structures of a number of homeodomain-DNA complexes have been determined by crystallography and NMR (reviewed in Gehring et al., 1994; Kornberg, 1993; Laughon, 1991; Wolberger, 1996). These studies have demonstrated that the overall fold and DNA-docking arrangements of the homeodomain are well conserved. Most homeodomains bind to very similar DNA sites that contain the sequence TAAT and typically have differential DNA binding specificity for the two base pairs following the conserved core binding sites (Fraenkel et al., 1998).

Atypical homeodomain proteins (like members of the Pros family) are structurally homologous to typical homeodomain proteins but display significant variation in sequence (Bürglin, 1994a). This may explain why Pros binds to a noncanonical DNA sequence (Cook et al., 2003; Hassan et al., 1997) and raises the possibility that it may use a somewhat different DNA recognition mode.

Atypical homeodomain structures have been solved previously (Ceska et al., 1993; Wolberger et al., 1991) and threading analysis predicted that the putative Prospero homeodomain was capable of assuming an atypical homeodomain structure (Banerjee-Basu et al., 1999). Crystal structure determination of the C-terminal region of Prospero (residues 1245–1403; Figure 1A) revealed that this region does indeed include a novel homeodomain (HD) structure (Ryter et al., 2002) but, unexpectedly this HD forms an integral structural unit with the adjacent Prospero domain (PD), a domain that is unique to Pros and its family. The combined structural unit is called Homeo-Prospero (HPD) domain.

At the extreme C terminus of the HPD domain, a short helix is masked by the PD. When this helix is unmasked, Prospero is exported from the nucleus (Bi et al., 2003). Therefore, it was suggested that the structure assumes a closed form that is essential for nuclear localization and is prearranged for DNA binding.

The C-terminal 236 amino acids of Pros, which includes the HPD and an additional amino-terminal 70 residues, was shown to bind a novel DNA sequence and activate the transcription of reporter genes in transiently transfected tissue culture cells (Demidenko et al., 2001; Hassan et al., 1997). Here we report the crystal structure of the Pros HPD (the most conserved 160 C-terminal amino acids) in complex with this DNA sequence. The structure helps understand the molecular

*Correspondence: brian@uoxray.uoregon.edu

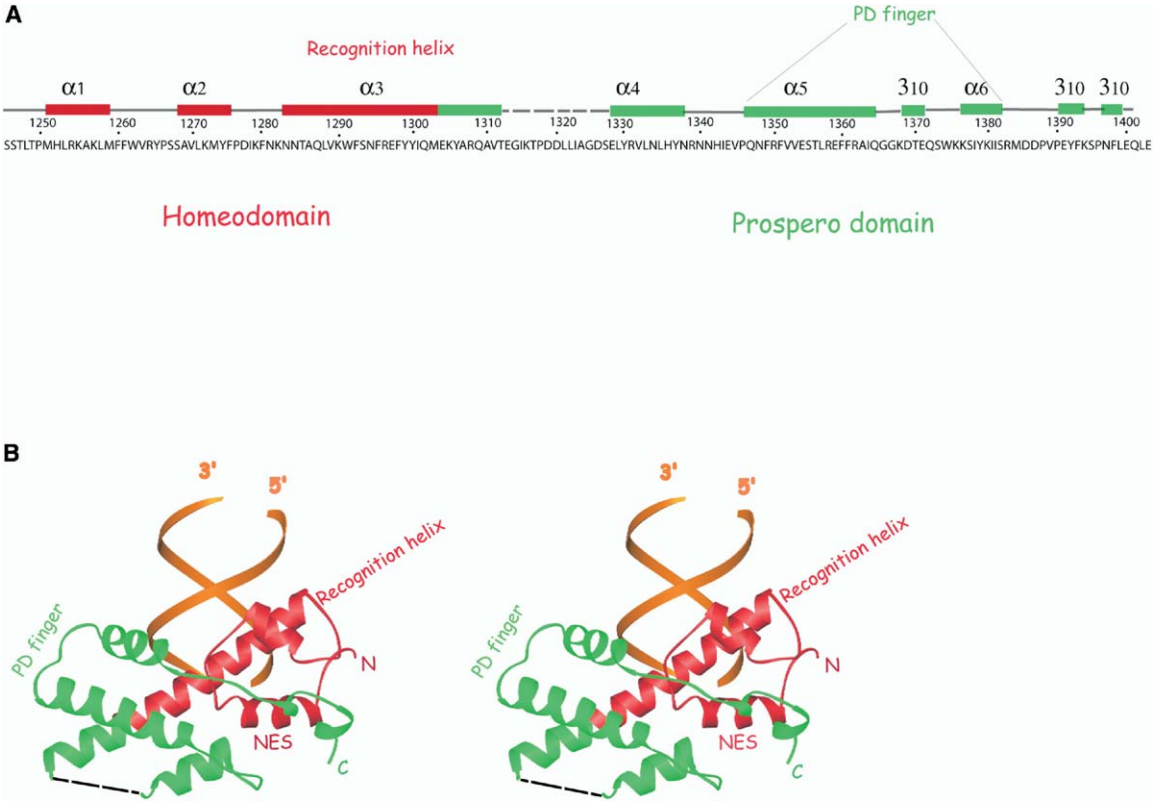

Figure 1. Overall Structure of the Homeo-Prospero Domain (HPD) and its Complex with DNA

(A) The amino acid sequence and secondary structure of the Prospero HPD domain. The homeodomain region (HD) is shown in red, the Prospero domain (PD) in green.

(B) Stereo view showing the complex between the protein and the DNA. The color-coding is as in (A) with the DNA shown in orange. The recognition helix connects the HD to the PD. The PD finger restricts the DNA binding site to the N-terminal part of the recognition helix. The C terminus of the PD also shields a helix that contains the nuclear export signal (NES) that lies within the HD. Dashed lines indicate disordered residues.

basis for recognition of the noncanonical DNA binding site. We have identified the invariant residues that specifically contact the DNA and propose a recognition mechanism that might be specific to Pros-type atypical homeodomains.

Results and Discussion

Overall Arrangement of the Prospero HPD/DNA Complex

The asymmetric unit of the Prospero HPD/DNA complex consists of a single copy of the Pros HPD bound to a 10 bp DNA duplex that includes the consensus sequence of Hassan et al. (1997) (see Experimental Procedures). In summary, the HD region contains three helices that are very similar to the overall fold of canonical homeodomains while the PD region is composed of a four-helical bundle with up-down-up-down topology (Figure 1B). The DNA-recognition helix bridges between the HD and the PD domains. The structure of HPD in the complex is essentially identical (rmsd 0.5 Å) to the apo form (Ryter et al., 2002). This close structural correspondence confirms that the HD and the PD regions act as a single HPD structural unit.

In the structure of the apo protein the extreme C-terminal residues of the PD were seen to mask the defined nuclear export signal (NES) that is within the HD region. Because the structure of the complex shows the conformation of five additional residues at the C terminus (see Experimental Procedures) it allows this arrangement to be seen more fully. The C terminus of the PD includes two consecutive 3_{10} helices that "wrap around" the surface of the α helix that includes the NES (Figures 1A and 1B). The amino acid sequence between the two 3_{10} helices includes a proline (Pro1396) that is invariant in Prospero family members and presumably helps define the sharp bend that occurs at this point (Figure 1B).

The DNA decamer in the crystal is a relatively straight segment of right-handed helical B DNA. The average helical twist is 34.4°, which corresponds to 10.5 bp per turn. Successive decamers are stacked end-to-end and create a pseudo-continuous duplex running through the crystal (Figure 2). Each decamer is related to the next by a crystallographic 2-fold axis that is normal to

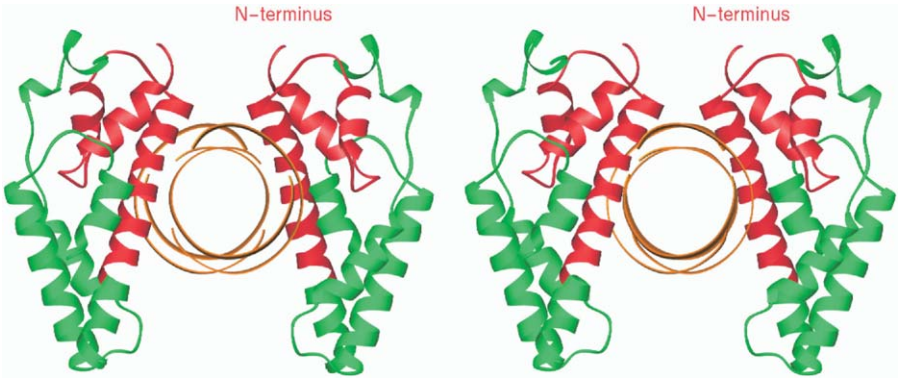

Figure 2. Stereo View along the Pseudocontinuous Helix in the Crystal

Two symmetry mates are related by a 2-fold axis with their N termini about 20 Å apart.

the long axis of the DNA and passes between successive duplexes.

Comparison with the Engrailed Homeodomain

In the HPD complex, the recognition helix spans the DNA duplex parallel to the sugar-phosphate backbone. The HD portion of this "hybrid" helix fits into the major groove with several of its side chains making specific contacts with the DNA. The hydrophobic face of this helix packs against the other helices in the structure to form the core of the protein (Figure 1B). To compare the mode of binding of the HPD to DNA with that of a "standard" homeodomain, we superposed the structure of the HPD-DNA complex on that of the Engrailed homeodomain (Kissinger et al., 1990) (Figure 3). The two structures were aligned by superimposing the backbone of the Engrailed homeodomain on that of the Prospero homeodomain. As can be seen, the two homeodomains superimpose remarkably well, except at their N termini. The respective DNA targets, however, have very different alignments with the axis of one DNA duplex tilted by about 36° relative to the other.

If one attempts to align the DNA in the HPD:DNA complex so that it corresponds to the Engrailed-DNA complex, there is an obvious steric clash with part of the PD (Figure 3). In particular, there is a protruding finger of the PD, made up of helices α_5 and α_6 plus a connecting 3_{10} helix, that contacts the DNA and appears to act as a structural guide that limits the possible orientations of the protein as it binds to the DNA (Figures 1B and 3).

Contacts with the Phosphate Backbone

Most of the contacts made by the protein with the DNA are in the vicinity of base pairs 2–4 (Figure 4B). These include sequence specific and nonspecific contacts. The invariant residues Trp1291 and Gln1287 contact the phosphate moiety of the critical adenine at base pair 2.

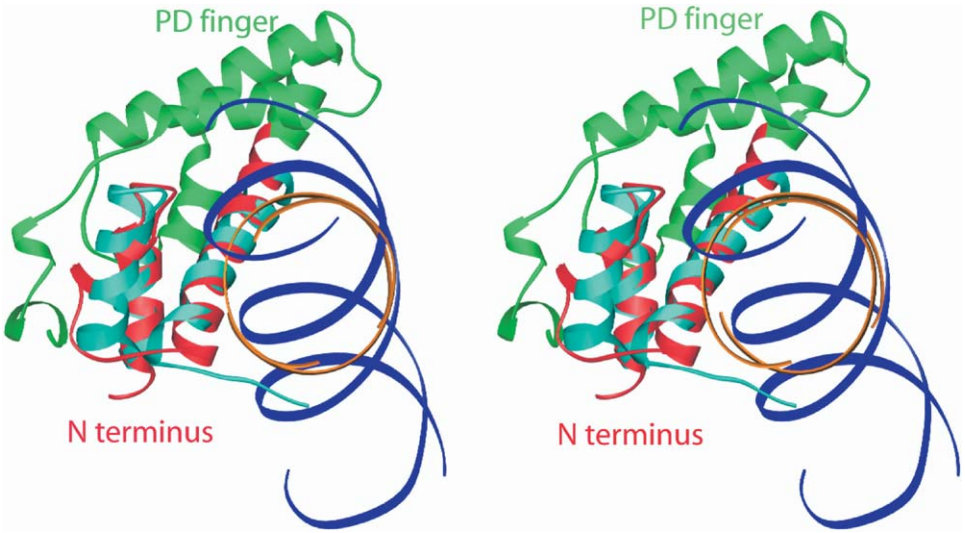

Figure 3. Comparison of DNA Binding by Homeo- and Homeo-Prospero Domains

Superposition of the HPD-DNA complex, colored as in Figure 1B, on that for the Engrailed homeodomain (Kissinger et al., 1990), colored in cyan with DNA in dark blue. The two complexes are superimposed based on the respective homeodomains (Prospero in red, Engrailed in cyan), which have very similar structures. As can be seen, the respective DNA targets have very different orientations. The PD finger of the HPD structure appears to restrict the alignment of the DNA and prevent it assuming an orientation similar to that in the Engrailed complex.

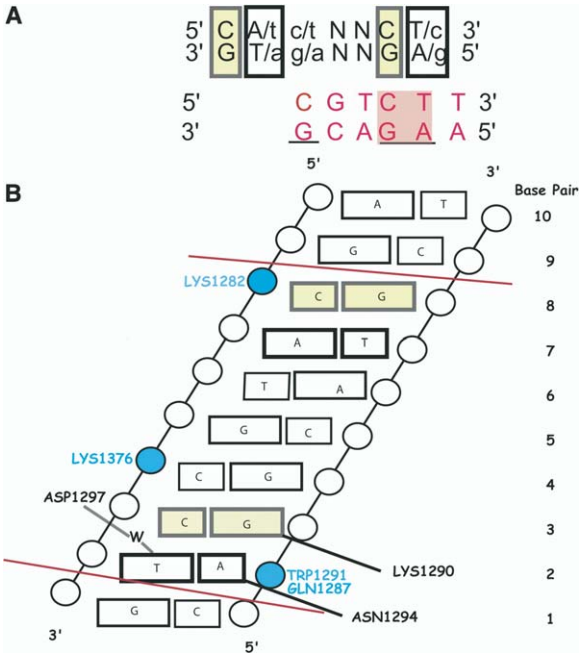

Figure 4. Recognition of DNA by the Homeo-Prospero Domain

(A) The upper DNA sequence, in black, is the consensus reported by Hassan et al. (1997) for preferential binding by Prospero. The yellow boxes indicate the base pairs that are most strongly favored. Noncolored boxes indicate sites that can be occupied by more than one base with uppercase letters indicating a stronger preference. N indicates sites at which there is no apparent preference. The sequence shown in red is a proposed transcriptional target for Pros in the *Drosophila* genome (Cook et al., 2003). The consensus sequence is underlined and the two bases that were found to be most critical for Pros binding are highlighted. The black and the red sequences are aligned based on the structure of the HPD-DNA complex (see text).

(B) Sketch summarizing all the contacts made by the Pros HPD. The red lines define the consensus sequence. Phosphates are represented with circles; blue circles show phosphates that are contacted by HPD (residue names are in blue). The thick black lines show major groove contacts made by the HPD (residue names are in black). The gray line indicates a water-mediated contact.

The invariant Lys1376 contacts the phosphate group at base pair 4. This interaction is particularly important since it involves the PD finger and the DNA backbone. Lys1282 (either Lys or Arg in the Pros family) contacts the phosphate group of the cytosine at base pair 8. In addition, Arg1339 (lysine in some family members) contacts the guanine backbone at what would correspond to base pair –4 in Figure 4B). In the crystal this contact is to a separate piece of DNA, but since this DNA extends the helix in a pseudocontinuous manner we assume that this contact would also occur when HPD binds to a sufficiently long segment of continuous DNA.

Invariant Residues and Critical Contacts with the DNA

Lys1290, Asn1294, and Asp1297 are conserved in every member of the Prospero family (Ryter et al., 2002) and lie within the exposed hydrophilic face of the HD recognition helix. In the crystal structure, these invariant residues occur in the section of the recognition helix closest to the major groove and make sequence-specific contacts with the DNA (Figures 5A and 5B). This region of the DNA also corresponds to part of the consensus sequence that was identified in a screen for preferred binding sites (Hassan et al., 1997; Figure 4A). Lys1290 makes a hydrogen bond with guanine at base pair 3 by donating a proton to the O6 position. A modest rotation of this side chain would allow it to hydrogen bond also with N7, but the distance between the ϵ-amino group of side chain and N7 of the guanine is 4.3 Å in the current structure. An interaction that Prospero HPD shares with all homeodomain structures is the bidentate hydrogen bonding between Asn1294 (conserved in all HD proteins) and the adenine at base pair 2. Asn1294 donates a hydrogen bond to N7 and accepts a hydrogen bond from N6. This explains the strong preference for the adenine base at this position compared to the guanine. Asp1297 makes a water-mediated interaction with the thymine at base 2 (Figure 5B).

These protein residues that contact the DNA have more or less the same side chain conformation in the free (Ryter et al., 2002) and the DNA bound structures. This suggests that once Pros is in the nucleus (i.e., adopting the closed conformation to mask the nuclear export signal), these amino acids are preorganized to bind to the DNA. Unexpectedly, the base pairs at positions 7 and 8 (Figure 4B), which were identified as being strongly preferred in the binding assay (Figure 4A), do not make a base-specific contact with the protein. This raises the possibility that some additional residues at the N terminus of the Pros HPD may be involved in specific contacts at this region. This assumption seems reasonable in that the consensus sequence was deduced using a protein construct that is 70 residues longer at the N terminus end of the HPD (Hassan et al., 1997).

An Alternative Transcriptional Target for Prospero

Cook et al. (2003) have shown that Prospero is involved in photoreceptor development in *Drosophila*. They also found that Prospero preferentially recognizes the sequence element AAGACG but could not detect significant similarity between this element and that previously found by Hassan et al. (1997).

In the crystal structure of the HPD-DNA complex reported here there are two adjacent base pairs that make sequence-specific contacts with the protein (Figure 4B). The sequence of this dinucleotide (5′A-G3′) is present in the sequence element of Cook et al. (2003) and leads to the alignment shown in Figure 4A. When aligned in this way there are no contradictions between the sequence of Cook et al. (2003) and the consensus proposed by Hassan et al. (1997). In particular, the base pair at position 6 (GC) (Cook et al., 2003) is consistent with the preference for GC or AT at this site (Hassan et al., 1997). More important, however, Cook et al. (2003) showed that mutation of either the A or G at sites 2 or 3 drastically decreased binding to Prospero. Thus it

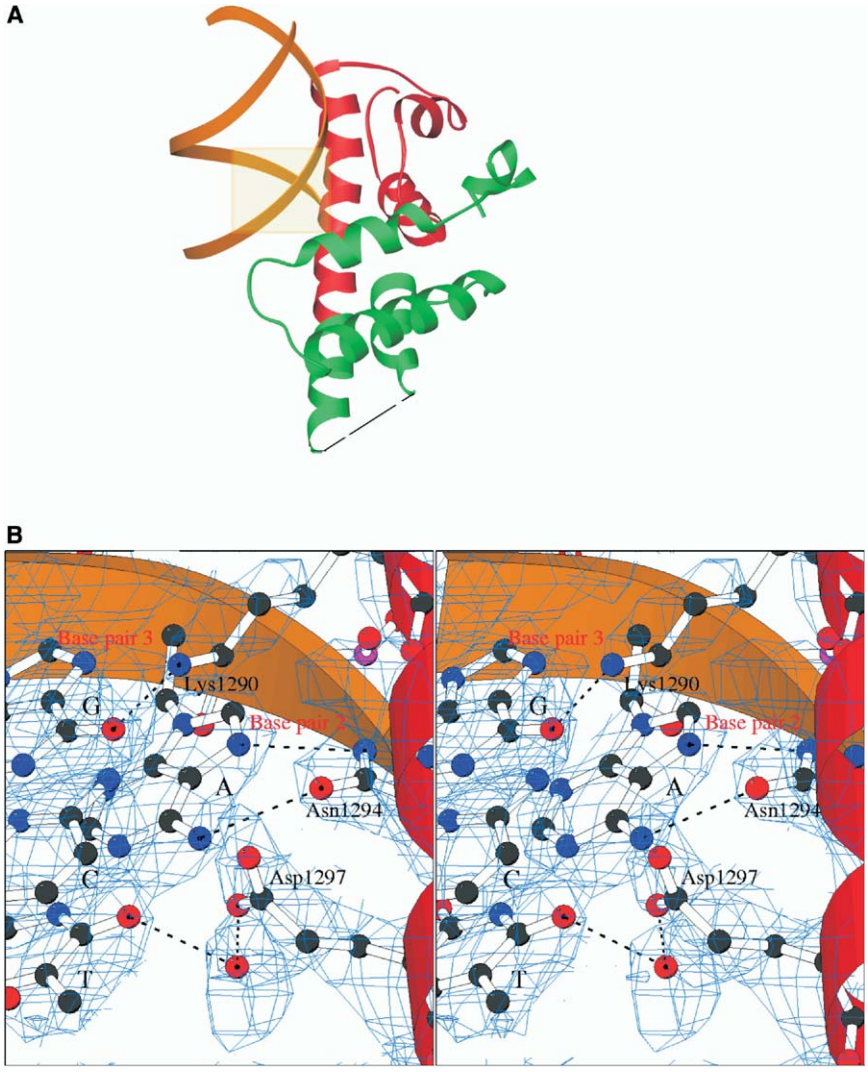

Figure 5. The Sequence-Specific Base Contacts Made by Invariant Residues of the Pros HPD

(A) The region of contacts (highlighted) relative to the overall arrangement of the complex.

(B) A stereo view showing details of the base contacts. Putative hydrogen bonds are indicated in dotted lines. The figure includes a sigmaa-weighted $2F_o - F_c$ map (Read, 1986) contoured at 1σ.

appears that the data of Hassan et al. (1997) and Cook et al. (2003) can be reconciled by aligning their respective target sequences as in Figure 4A.

Relationship between DNA Binding and Unmasking of the Nuclear Export Signal

The Pros HPD-DNA structure described here reveals that the NES is masked, similar to the apo structure. It is an open question whether masking the NES is a prerequisite for DNA binding (or vice versa). However, the C-terminal segment of Pros that masks the NES is on the opposite side of the protein to the DNA binding region. This suggests that DNA binding is not directly coupled with masking or unmasking of the NES. It is also consistent with the observation that the structure of the HPD is essentially identical in the apo and DNA bound forms and, in particular, that the NES remains masked in both structures. The present structural results therefore suggest that DNA binding and exposure of the NES are not directly coupled.

The N Terminus of the HPD and a Proposed DNA Recognition Mechanism

In most HD structures there is a flexible region, called the N-terminal arm, which extends from the N terminus of the first helix and wraps around the DNA to make base-specific and phosphate backbone contacts in the minor groove, supplementing the contacts made by the recognition helix (Fraenkel et al., 1998; Grant et al., 2000; Mathias et al., 2001). In the Prospero HPD-DNA complex structure, the N terminus actually points away from the minor groove (Figures 2 and 3). The N termini

Table 1. Data Processing and Structural Refinement Statistics

Data Processing	
Space group	$P3_22_12$
Unit cell dimensions (Å)	a = b = 44.7, c = 226.6
Monomers per asymmetric unit	1
Resolution range (outer shell) (Å)	30–2.8 (2.85–2.8)
Number of observations	20,738
Number of unique reflections	6,354
Completeness (outer shell) (%)	90 (85.8)
R_{sym} (outer shell) (%)	3.8 (10.0)
$I/\sigma(I)$ (outer shell)	33 (7.5)
Structural Refinement	
Resolution range (Å)	30–2.8
R_{work}/R_{free} (%)	21.7/28.2
Rms deviations	
Bond lengths (Å)	0.006
Bond angles (°)	1.1
Distribution of (φ,ϕ) for most favored and additionally allowed regions (%)	100
Total number of non-hydrogen atoms	1,659
Number of water molecules	15

of two symmetry mates in the crystal are relatively close to each other (~20 Å) along the pseudocontinuous DNA double helix (Figure 2). Similarly, if two monomers were to bind to a palindromic DNA sequence, which has 2-fold symmetry, their N termini would also be positioned close in space. This suggests that in the full Prospero protein the extension at this N terminus might provide a dimer or other interface that would allow the HPD and its partner to bind to longer DNA sequences than used in the present study. In this context it has been shown that the extreme N terminus of the HPD along with 70 additional residues (residues 1171–1288) can bind and modulate homeodomain activities of other proteins (Hassan et al., 1997).

An increase of DNA binding affinity and sequence selectivity due to the formation of cooperative complexes (heterodimers) has been observed for homeodomain proteins in developing cells both in *Drosophila* (Passner et al., 1999) and in human (Mann and Chan, 1996). Such complex formation can reconcile the highly specific in vivo function of many DNA transcription factors with their low specificity in vitro.

It is possible that Pros could bind to a palindromic sequence as a homodimer, or to a nonpalindromic sequence as a heterodimer. Prokaryotic homeodomains bind to palindromic sites as homodimers while many eukaryotic homeodomains bind to DNA either as homo- or heterodimers (Luscombe et al., 2000). The fact that HPD alone makes relatively few contacts with the DNA (Figure 4B) also suggests that these interactions might be supplemented by a partner protein. In this context, the role played by the PD to limit the orientation of the recognition helix could be particularly important since it not only orients the key residues optimally for DNA binding but also exposes the N termini of the HPD, bringing them close together.

In summary, we conclude that the following steps may be involved in Pros function as a transcription factor and gene regulator in developing neural cells. (1) In the neuroblast, Pros is in an open state that allows the NES to direct nuclear export of Pros. (2) In the ganglion mother cell (GMC), a physiological signal induces Pros to adopt a closed conformation so that the PD shields the nuclear export signal and Pros enters the nucleus. (3) Once in the closed conformation, the Pros HPD is preorganized for DNA binding. The Pros PD sterically restricts alignment of the protein on the DNA and may help position the N-terminal region for interaction with a suitable partner. (4) Sequence-specific contacts of Pros at one end of the consensus site help confer specific recognition. (5) Putative sequence-specific contacts with a different region of the protein at the other end of the optimal site might enhance the specificity and stability of the complex. (6) Putative homo- or hetero-dimerization events could further stabilize the complex and allow Pros to function as either a transcription factor or a cofactor. (7) In the neuron, the NES masking signal is turned off or blocked so that cytoplasmic Pros adopts a more open structure, reexposes the nuclear export signal, and exits the nucleus.

Experimental Procedures

Sample Preparation

Protein Expression and Purification

The C-terminal of Pros (residues 1241–1403) was overexpressed and purified essentially as described previously (Ryter et al., 2002). The only difference is that a pET15b-based vector that adds a cleavable N terminus His tag (instead of C terminus noncleavable His tag) was used during cloning. The N terminus His tag was digested by incubation with TEV-protease, overnight, at 16°C. The cleaved tag was then removed by passing the digestion reaction mixture through a Ni-NTA agarose (Qiagen) column. The removal of the tag and the purity of the sample were confirmed by matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry and SDS-PAGE electrophoresis.

DNA Preparation

The sequence of the DNA used for cocrystallization was 5′-AGC ATGCCTG-3′, together with the complementary strand. It contains the consensus sequence 5′-CATGCCT-3′ that preferentially binds to Prospero (Hassan et al., 1997). The single-stranded oligonucleotides (cartridge purified, Invitrogen) were dissolved in 100 mM Na-cacodylate buffer (pH 7). To form duplex DNA, equal molar ratios of the single strands were mixed and heated to 90°C for 5 min and then cooled slowly to room temperature over a period of 1 hr.

Complex Formation and Crystallization

The complex was formed with 1:1.3 molar ratio of Prospero to duplex DNA (0.5 M protein [~20 mg/ml] was mixed with 0.65 M DNA). Rather small crystals (0.1 × 0.15 × 0.15 mm) were grown in 3 weeks at 4°C using the hanging drop vapor diffusion method against a well solution containing 100 mM citrate (pH 5.0) and 24% PEG4000.

Data Collection and Structure Determination

Diffraction data were collected at the Advanced Light Source (beamline 8.3.1, λ = 1.0 Å) at 100 K, with 20% glycerol as a cryoprotectant. Although the crystal diffracted strongly to 2.5 Å, the alignment of the crystal axes relative to the detector made it difficult to use data beyond 2.8 Å. Data were integrated and scaled using the HKL suite of programs DENZO, XDISPLAYF, and SCALEPACK (Otwinowski and Minor, 1997). The structure was solved by molecular replacement (Rossmann, 1972), using the Prospero apo structure (Ryter et al., 2002) as a search model, and refined using CNS (Brunger et al., 1998). R_{free} was 44% after a few cycles of rigid-body refinement. The DNA duplex was unambiguously identified in $2F_o - F_c$ and $F_o - F_c$ electron density maps and a DNA model with the correct sequence was constructed. The program *O* (Jones et al., 1991) was employed for the manual building of the model and 15 water molecules were added. Later, electron density was observed for five residues at the C terminus, which were then added

(1397–1401). These residues were disordered in the apo form (Ryter et al., 2002), possibly due to the existence of a highly mobile C-terminal His tag. As in the apo structure, however, residues 1314–1326 were disordered and were not modeled. A summary of the data processing and refinement statistics is given in Table 1.

Acknowledgments

We thank Jodi Ryter and Blaine Mooers for helpful technical advice, Richard Kingston for the expression vector, Anthony Addlagatta for help with data collection, Chris Doe for background and insightful comments on the manuscript, and the staff of the Advanced Light Source for help with data collection. This work was supported in part by National Institutes of Health grant GM20066 to B.W.M.

Received: October 26, 2004
Accepted: January 8, 2005
Published: April 11, 2005

References

Banerjee-Basu, S., Landsman, D., and Baxevanis, A.D. (1999). Threading analysis of prospero-type homeodomains. In Silico Biol. *1*, 163–173.

Bi, X., Kajava, A.V., Jones, T., Demidenko, Z.N., and Mortin, M.A. (2003). The carboxy terminus of Prospero regulates its subcellular localization. Mol. Cell. Biol. *23*, 1014–1024.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Bürglin, T.R. (1994a). A comprehensive classification of homeobox genes. In Guidebook to the Homeobox Genes, D. Duboule, ed. (Oxford: Oxford University Press)), pp. 25–71.

Bürglin, T.R. (1994b). A *Caenorhabditis elegans* prospero homologue defines a novel domain. Trends Biochem. Sci. *19*, 70–71.

Ceska, T.A., Lamers, M., Monaci, P., Nicosia, A., Cortese, R., and Suck, D. (1993). The X-ray structure of an atypical homeodomain present in the rat liver transcription factor LFB1/HNF1 and implications for DNA binding. EMBO J. *12*, 1805–1810.

Chu-Lagraff, Q., Wright, D.M., McNeil, L.K., and Doe, C.Q. (1991). The prospero gene encodes a divergent homeodomain protein that controls neural identity in Drosophila. Development *2* (Suppl.), 79–85.

Cook, T., Pichaud, F., Sonneville, R., Papatsenko, D., and Desplan, C. (2003). Distinction between color photoreceptor cell fates is controlled by Prospero in *Drosophila*. Dev. Cell *4*, 853–864.

Demidenko, Z., Badenhorst, P., Jones, T., Bi, X., and Mortin, M.A. (2001). Regulated nuclear export of the homeodomain transcription factor *Prospero*. Development *128*, 1359–1367.

Doe, C.Q., Chu-LaGraff, Q., Wright, D.M., and Scott, M.P. (1991). The *prospero* gene specifies cell fate in the *Drosophila* central nervous system. Cell *65*, 451–465.

Dudas, J., Papoutsi, M., Hecht, M., Elmaouhoub, A., Saile, B., Christ, B., Tomarev, S.I., Von Kaisenberg, C.S., Schweigerer, L., Ramadori, G., and Wilting, J. (2004). The homeobox transcription factor Prox1 is highly conserved in embryonic hepatoblasts and in adult and transformed hepatocytes, but is absent from bile duct epithelium. Anat. Embryol. (Berl.) *208*, 359–366.

Fraenkel, E., Rould, M.A., Chambers, K.A., and Pabo, C.O. (1998). Engrailed homeodomain-DNA complex at 2.2 Å resolution: a detailed view of the interface and comparison with other engrailed structures. J. Mol. Biol. *284*, 351–361.

Gehring, W.J., Affolter, M., and Bürglin, T.R. (1994). Homeodomain proteins. Annu. Rev. Biochem. *63*, 487–526.

Grant, R.A., Rould, M.A., Klemm, J.D., and Pabo, C.O. (2000). Exploring the role of glutamine 50 in the homeodomain-DNA interface: crystal structure of engrailed (Gln50 → Ala) complex at 2.0 Å. Biochemistry *39*, 8187–8192.

Hassan, B., Li, L., Bremer, K.A., Chang, W., Pinsonneault, J., and Vaessin, H. (1997). Prospero is a panneural transcription factor that modulates homeodomain protein activity. Proc. Natl. Acad. Sci. USA *94*, 10991–10996.

Hirata, J., Nakaghoshi, H., Nabeshima, Y., and Matsuzaki, F. (1995). Asymmetric segregation of the homeodomain protein Prospero during *Drosophila* development. Nature *377*, 627–630.

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

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

Kornberg, T.B. (1993). Understanding the homeodomain. J. Biol. Chem. *268*, 26813–26816.

Laughon, A. (1991). DNA binding specificity of homeodomains. Biochemistry *30*, 11357–11367.

Luscombe, N.M., Austin, S.E., Berman, H.M., and Thornton, J.M. (2000). An overview of the structures of protein-DNA complexes. Genome Biol. *1*, 1–10.

Mann, R.S., and Chan, S.K. (1996). Extra specificity from extradenticle: the partnership between HOX and PBX/EXD homeodomain proteins. Trends Genet. *12*, 258–262.

Mathias, J.R., Zhong, H., Jin, Y., and Vershon, A.K. (2001). Altering the DNA-binding specificity of the yeast Mat α2 homeodomain protein. J. Biol. Chem. *276*, 32696–32703.

Matsuzaki, F., Koizumi, K., Hama, C., Yoshioka, T., and Nabeshima, Y. (1992). Cloning of the Drosophila prospero gene and its expression in ganglion mother cells. Biochem. Biophys. Res. Commun. *182*, 1326–1332.

Oliver, G., Sosa-Pineda, B., Geisendorf, S., Spana, E.P., Doe, C.Q., and Gruss, P. (1993). Prox *1*, a prospero-related homeobox gene expressed during mouse development. Mech. Dev. *44*, 3–16.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Passner, J.M., Ryoo, H.D., Shen, L., Mann, R.S., and Aggarwal, A.K. (1999). Structure of the DNA-bound ultrabithorax-extradenticle homeodomain complex. Nature *397*, 714–719.

Read, R.J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Crystallogr. *A42*, 140–149.

Rossmann, M.G. (1972). The Molecular Replacement Method (New York: Gordon & Breach).

Ryter, J.M., Doe, C.Q., and Matthews, B.W. (2002). Structure of the DNA binding region of Prospero reveals a novel homeo-Prospero domain. Structure *10*, 1541–1549.

Spana, E., and Doe, C.Q. (1995). The prospero transcription factor is asymmetrically localized to the cell cortex during neuroblast mitosis in Drosophila. Development *121*, 3187–3195.

Tomarev, S., Sundin, O., Banerjee-Basu, S., Duncan, M., Yang, J., and Piatigorsky, J. (1996). Chicken homeobox gene Prox 1 related to *Drosophila* prospero is expressed in the developing lens and regina. Dev. Dyn. *206*, 354–367.

Vaessin, H.E., Grell, B., Wolff, E., Bier, L.Y., and Jan, Y.N. (1991). Prospero is expressed in neural precursors and encodes a nuclear protein that is involved in control of axonal outgrowth in Drosophila. Cell *67*, 941–953.

Wolberger, C. (1996). Homeodomain interactions. Curr. Opin. Struct. Biol. *6*, 62–68.

Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D., and Pabo, C.O. (1991). Crystal structure of a MATα2 homeodomain-operator complex suggests a general model for homeodomain-DNA interactions. Cell *67*, 517–528.

Accession Numbers

Coordinates and structure factors have been deposited in the Protein Data Bank with accession code 1XPX.