Structural and Dynamics Studies of Pax5 Reveal Asymmetry in Stability and DNA Binding by the Paired Domain

Cecilia Perez-Borrajero¹,², Mark Okon² and Lawrence P. McIntosh²,³,⁴

¹ - Genome Sciences and Technology Program, University of British Columbia, Vancouver V6T 1Z3, Canada
² - Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver V6T 1Z3, Canada
³ - Department of Chemistry, University of British Columbia, Vancouver V6T 1Z1, Canada
⁴ - Michael Smith Laboratories, University of British Columbia, Vancouver V6T 1Z4, Canada

Correspondence to Lawrence P. McIntosh: Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver V6T 1Z3, Canada. mcintosh@chem.ubc.ca
http://dx.doi.org/10.1016/j.jmb.2016.04.004
Edited by Charalampo Kalodimos

Abstract

The eukaryotic transcription factor Pax5 or B-cell specific activator protein (BSAP) is central to B-cell development and has been implicated in a large number of cellular malignancies resulting from loss- or gain-of-function mutations. In this study, we characterized the DNA-binding Paired domain (PD) of Pax5 in its free and DNA-bound forms using NMR spectroscopy. In isolation, the PD folds as two independent helical bundle subdomains separated by a conformationally disordered linker. The two subdomains differ in stability, with the C-terminal subdomain (CTD) being ~10-fold more protected from amide hydrogen exchange (HX) than the N-terminal subdomain (NTD). Upon binding DNA, the linker and an induced N-terminal β-hairpin become ordered with significantly dampened motions and increased HX protection. Both subdomains of the PD contribute to specific DNA binding, resulting in an equilibrium dissociation constant more than three orders of magnitude lower than exhibited by the separate subdomains for their respective half-sites (nM versus μM).

The isolated CTD binds non-specific DNA sequences with only ~10-fold weaker affinity than cognate sequences. In contrast, the NTD associates very poorly with non-specific DNA. We propose that the more stable CTD has evolved to provide relatively low affinity non-specific contacts with DNA. In contrast, the more dynamic NTD discriminates between cognate and non-specific sites. The distinct roles of the PD subdomains may enable efficient searching of genomic DNA by Pax5 while retaining specificity for functional regulatory sites.

© 2016 Elsevier Ltd. All rights reserved.

Introduction

The Paired box (Pax) family of proteins are a highly conserved group of transcription factors involved in embryonic development and cell differentiation (for reviews, see Refs. [1–4]). In mammals, nine Pax genes (Pax1 to Pax9) have been identified to encode proteins defined by an N-terminal DNA-binding region of ~128 amino acids known as the Paired domain (PD) [5,6]. Of these, Pax5 or B-cell specific activator protein (BSAP) was first identified for its DNA-binding ability and expression in immature B-cells [7]. In cooperation with Pax2, Pax5 is important in the formation of the central nervous system during mouse embryonic development [8–11]. In adult organisms, Pax5 drives differentiation and maturation of B-cells, serving both as an activator of B-cell-specific genes (e.g., CD19, mb-1, and Blnk) and a repressor of alternative developmental programs [12–15]. Not surprisingly, aberrant Pax5 activity has been implicated in a large number of B-cell malignancies and other non-hematopoietic cancers, where it can act as oncogene or tumor suppressor, depending on the cellular context (for review, see Ref. [15]). In the latter case, a reduction in Pax5 function can lead to B-cell acute lymphoblastic leukemia.
Asymmetry in DNA-binding by the Pax5 Paired domain

(B-ALL) [16–18], whereas reestablishment of Pax5 expression ameliorates the malignant phenotype [19]. X-ray crystallographic studies on Pax5 in a ternary complex with its partner transcription factor Ets-1 and DNA corresponding to the mb-1 promoter showed that the PD is composed of two helical bundle subdomains joined by a linker polypeptide of ~20 residues [20,21] (Fig. 1a). Helix-turn-helix motifs in the N-terminal subdomain (NTD, also called PAI) and C-terminal subdomain (CTD, also called RED) bind the major groove of DNA at sites separated by approximately one turn of the phosphodiester backbone. Additional minor groove contacts are provided by a short N-terminal β-hairpin and adjacent loops, as well as the linker. Although the mb-1 DNA sequence does not contain optimal binding sites for either of these two transcription factors, interactions between the β-hairpin of Pax5 and the ETS domain of Ets-1 result in cooperative binding to form a stable complex capable of transcriptional activation [20,22].

Previous studies have shown that the consensus DNA-binding site of the Pax5 PD is extended and somewhat degenerate [23–25] (Fig. 1a). This is consistent with the structure of the Pax5/Ets-1/DNA complex, as several regions of the PD can contribute to its net affinity for target sequences. A more recent study identified ~20,000 and ~15,000 DNA-binding sites for Pax5 in pro-B and mature B cells, respectively, corresponding to ~40% of the total number of nuclease hypersensitive regions identified in these cells [25]. However, only ~360 genes were found to change in expression level by more than 4-fold upon deletion of Pax5 [25]. These data suggest that, in most cases, DNA binding by Pax5 is not sufficient to significantly alter gene expression levels and

Fig. 1. The subdomains of Pax51-149 fold as independent helical bundles. (a) Boundaries of the protein fragments used in this study are shown with colored rectangles indicating the NTD and CTD. The DNA-binding consensus sequence of the PD according to the JASPAR CORE database [95] is also shown. Bases present in more than 50% of the sequences within this dataset are written explicitly, with those most conserved (~75%) underlined. The x represents less than 50% preference for any particular base. (b) Overlaid 15N-HSQC spectra of the three Pax5 fragments collected at pH 6.5 and 25 °C. Residues Arg50, Glu113, and Ile83 are labeled as examples. The close overlap of dispersed signals from the two subdomain fragments and the full-length PD indicates that the NTD and CTD are structurally independent. In contrast, the sharp signals with poorly dispersed 1HN shifts near 8 to 8.5 ppm arise from conformationally disordered amides flanking the helical bundles. See Supplementary Fig. S1 for full assignments. (c) In the absence of DNA, the Pax5 PD consists of two independent helical bundle subdomains flanked by conformationally disordered residues. Shown are normalized α-helical and β-strand propensities per residue based on the MICS analysis [40] of backbone 13Cα/13Cβ/15CO/15N/4H chemical shifts for Pax51-92 (yellow, left), Pax576-149 (red, right), and Pax51-149 (blue, bottom). The solid black lines denote the RCI-S2 values for each residue, with lower values indicating greater predicted conformational disorder. The top cartoon shows the secondary structure (black arrows for β-strands, rectangles for α-helices) of the PD observed in a crystallized Pax5/Ets-1/DNA complex (PDB ID: 1MDM).
other regulatory factors are needed. Partnerships with proteins such as Ets-1, Grg4, CBP, and the basal transcriptional machinery, along with upstream signaling events, must be critical for ensuring proper activation and repression of Pax5 gene targets [20,26–28].

To date, only the Paired domains of Pax6 and Pax8 have been structurally characterized in isolation. Early studies using circular dichroism spectroscopy indicated that these PDs had low α-helical content, which increased upon DNA binding [29,30]. In contrast, using NMR spectroscopy, Codutti et al. [31] demonstrated that the Pax8 DNA-binding domain adopts stable secondary structures on its own and folds in a manner similar to other PD proteins studied in complex with DNA [20,32,33]. However, the authors noted weak tertiary contacts in the helical subdomains of the PD and the absence of the N-terminal β-hairpin that is found in crystals of PD/DNA complexes [20,32,33]. Thus, we currently lack consistent insights into the stability and structural dynamics of the PDs in the Pax family and how these change upon formation of a complex with DNA. In addition, the relative contributions of the two subdomains to DNA binding and recognition is unclear and seems to be context dependent. For example, in the case of Paired from Drosophila melanogaster, the NTD appears to functionally dominate in the activation of genes involved in embryonic patterning, since deletion of the CTD has no phenotypic effect in vivo [34]. Furthermore, mutations in the DNA-binding domain affecting fly viability can be rescued with the NTD region of the protein alone [32]. Although a few reports agree on the dominance of the NTD in DNA binding by Paired [35,36], more recently, the CTD was found to be necessary for proper mating response in flies [37]. The importance of this subdomain in cooperating for target DNA recognition has also been recognized for Pax5 [23], Pax6 [38], and Pax8 [30,39]. Therefore, the individual subdomains may have different roles in DNA binding. However, given the similar structures and DNA recognition modes of the NTD and CTD subdomains, the reasons underlying these differences are currently unknown.

To gain a better understanding of the role of Pax proteins in transcriptional regulation, we used NMR spectroscopy to investigate the structure, dynamics, and stability of the free and DNA-bound forms of the PD of Pax5. We show that in the absence of DNA, the N-terminal ~30 residues, including those forming the β-hairpin, and linker regions are conformationally dynamic. In contrast, the subdomains adopt wellfolded three-helix bundles. However, the CTD is more stable than the NTD, as judged by amide hydrogen exchange (HX). Upon binding DNA, the helical regions do not undergo major structural changes, whereas residues encompassing the induced β-hairpin and the linker region become more ordered and protected from amide HX. As expected, the separate subdomains of Pax5 bind cognate DNAs weakly relative to the intact PD, which is required for high affinity complex formation. However, the CTD exhibits only modest differences in affinity for cognate versus non-specific DNAs, while the NTD interacts very poorly with non-specific DNAs. The different dynamic properties and relative affinities of the PD subdomains for specific versus non-specific DNA sequences suggest distinct asymmetric contributions toward promoter binding and transcriptional regulation by Pax5.

Results

The PD of Pax5 folds as two independent helical subdomains

To characterize the structural and dynamic properties of the Pax5 DNA-binding domain, we initially collected 15N-heteronuclear single quantum correlation (HSQC) spectra of three bacterially expressed fragments of this transcription factor: Pax51-149 containing the NTD and flanking regions, Pax576-149 containing the linker and CTD, and Pax51-149 spanning the entire PD (Fig. 1a). The two subdomain constructs showed dispersed amide peaks characteristic of folded structures, as well as sharp signals exhibiting random coil 1H N shifts (~8 to 8.5 ppm), diagnostic of conformational disorder (Fig. 1b). In addition, the corresponding amide 1H N, 15N signals of the smaller protein fragments (Pax51-149 and Pax576-149) overlaid closely with those of Pax51-149 (Fig. 1b), indicating that the NTD and CTD are structurally independent and can adopt the same fold whether in isolation or connected covalently to form the full-length PD. Consistent with this conclusion, the 15N-HSQC spectrum of 15N-labeled Pax51-92 was not perturbed upon addition of unlabeled Pax576-149 (not shown).

The assigned chemical shifts of backbone 1H N, 15N, 13C α, 13C β, and 13CO nuclei in Pax51-92, Pax576-149, and Pax51-149 (Supplemental Fig. S1) were used to calculate secondary structure propensities and random coil index-squared ordered parameters (RCI-S2) according to the MICS algorithm [40]. This analysis revealed that in solution, the NTD and CTD of Pax5 form well-ordered three-helix bundles (Fig. 1c). The predicted secondary structures of the individual subdomains did not change in the context of the full PD, and agree closely with those observed in the crystal structure of the Pax5/Ets-1/DNA complex (PDB ID: 1MDM) [21]. However, the N-terminal ~30 residues of Pax51-92 and Pax51-149 lacked any persistent secondary structure and exhibited low RCI-S2 values, indicative of conformational flexibility. Although residues 25–27 showed β-strand propensity, the overall chemical shift analysis indicates that the small β-hairpin
(residues 19–21 and 25–27) seen in the crystal structure of the Pax5/Ets-1/DNA complex is dynamic and not stably formed in either free Pax51-92 or Pax51-149. This conclusion is supported by amide HX and 15N relaxation studies presented in detail below. The linker residues in all three fragments were also found to be disordered, exhibiting random coil chemical shifts and hence low RCI-S2 parameters. Consistent with the spectral comparisons in Fig. 1b, these data confirm that the two subdomains do not interact with one another, but rather behave as “beads-on-a-string” to form the full PD.

The CTD of Pax51-149 is more stable than the NTD

Protein dynamics play a critical role in DNA binding by transcription factors [41,42]. Therefore, we used HX experiments to investigate the local and global stability of the Pax5 PD. Initially, lyophilized Pax51-149 was resuspended in D2O at pH* 7.0 and 25 °C and its protonation levels were monitored by NMR spectroscopy. In the first 15N-HSQC spectrum, recorded only ~10 min after resuspension, the vast majority of the amide signals were absent (Fig. 2a). Nearly all of the remaining 13 residues with detectable signals localized to the helical region of the CTD (Fig. 2b). Although amides within the unstructured N-terminal and linker regions were expected to exchange rapidly under these conditions, it was surprising that the entire NTD also exhibited little HX protection.

To better quantitate the HX behavior of Pax51-149, protium-deuterium exchange experiments were repeated at pH* 6.0 and 15 °C. The reduced sample pH* and temperature enabled measurements of HX rate constants for many amides that exchanged on the minutes–hours timescale. In parallel, we conducted magnetization transfer CLEANEX-PM experiments at pH 5.6, 6.3, and 8.0 (25 °C) to measure more rapid protium-protium HX (~seconds timescale). The fit exchange rate constants for each backbone amide, \( k_{\text{HX-obs}} \), were compared to those predicted for Pax51-149 (\( k_{\text{HX-pred}} \)) as a random coil polypeptide under the same conditions [43–45]. The calculated amide protection factors (PF) = \( k_{\text{HX-obs}} / k_{\text{HX-pred}} \) from HX and/or CLEANEX-PM experiments are shown in Fig. 2c. In accordance with their random coil chemical shifts and hence low RCI-S2 values, amides in the linker and terminal regions had PFs ~ 1. This indicates that these residues (which include those forming the β-hairpin and linker, increases (a) Pax51-149 was lyophilized and resuspended in D2O to monitor the decay in amide \( ^1\text{H}^{\text{N}} \)) signal intensity due to protium-deuterium exchange. Shown are overlaid 15N-HSQC spectra of Pax51-149 in H2O buffer (open blue, pH 6.5, 25 °C) and ~10 min after resuspension in D2O (solid cyan, pH* 7.0 and 25 °C). Peaks in the latter are assigned. (b) Amides that have not fully exchanged after this time localize primarily to the CTD and are highlighted in cyan on a cartoon representation of the PD derived from the Pax5/Ets-1/DNA crystal structure (PDB ID: 1MDM). (c) PFs for Pax51-149 obtained from either slow protium-deuterium exchange at pH* 6.0 and 15 °C or fast CLEANEX-PM protium-protium exchange at pH 5.6, 6.3, or 8.0 and 25 °C are plotted in blue or in cyan (the latter highlighting those most protected from HX in panel (a)). Also shown in gray are the PFs of Pax51-149 bound to CD19 DNA and measured from protium-deuterium exchange at pH* 6.60 and 25 °C. The arrow heads identify the lower PF limits for residues exchanging too slowly to obtain reliable \( k_{\text{HX-obs}} \) values. Missing data corresponds to prolines, residues with overlapping amide chemical shifts, or residues with exchange rate constants outside of the measureable experimental ranges of the two approaches. The top cartoon indicates the secondary structure of the PD as in Fig. 1c.
respectively. Thus, relative to well-folded globular proteins with PFs typically in the range of $10^4$ to $10^7$ [47–49], neither subdomain is highly stable against fluctuations leading to HX. Nevertheless, the CTD of Pax51-149 is more protected than the NTD.

Upon binding DNA, the Paired domain $\beta$-hairpin and linker region become ordered

To characterize the potential changes in structure and dynamics that occur upon DNA binding by the PD, we assigned the backbone chemical shifts of amide-protonated $^2$H/$^{15}$N/$^{13}$C-labeled Pax51-149 in complex with a 25 bp DNA duplex corresponding to the high affinity variant of the CD19 promoter, CD19-2_Ains (Table 1, Fig. 3a and Supplementary Fig. S1) [23]. The majority of the amides in the protein exhibited chemical shift perturbations (CSPs) in the presence of DNA (Fig. 3a and c). However, based on chemical shift analysis using the MICS algorithm, the $\alpha$-helical content of the complex remained very similar to that of the free protein (Fig. 3b). In contrast, the RCI-S$^2$ values of the linker and the N-terminal residues increased to match those in the helical bundles. This indicates that the protein backbone in these regions becomes more ordered when bound to DNA (Fig. 3b). In addition, the $\beta$-hairpin and linker regions showed slightly increased $\beta$-strand propensities. In the case of the linker, this most certainly reflects its restriction to an extended conformation and not the formation of a $\beta$-sheet. However, in the case of the N-terminal region, residues Gln22 and Asn29, which are part of the $\beta$-hairpin turn and loop leading to helix H1, respectively, showed large increases in turn propensities (Supplementary Fig. S2). This suggests that the $\beta$-hairpin and adjacent turns are stabilized upon DNA binding and support previous observations that this region makes crucial contacts with DNA [20,31–33].

Amide CSPs due to binding of the CD19 DNA were mapped onto the crystal structure of the PD in the Pax5/Ets1/DNA ternary complex (Fig. 3c, PDB ID:1MDM) [20,21]. The largest CSPs localized to amides in the $\beta$-hairpin and adjacent loops, the recognition helix (H3) of the NTD, the linker region, and the recognition helix (H6) of the CTD. This is in close agreement with the DNA contacts observed by X-ray crystallography. In particular, residues Gln22, Gly30, Val90, and Trp112, which exhibited very high CSPs, all contact the phosphate backbone or make hydrogen bonds with the bases of the DNA (Fig. 3c).

To gain complementary dynamic insights, we also performed HX experiments on the Pax51-149/CD19 complex at pH* 6.60 and 25 °C. Relative to the unbound state, amides throughout the protein became significantly more protected from exchange (Fig. 2c). Of particular note, residues in the $\beta$-hairpin and linker regions had PFs $\sim 10^3$ in the bound state, versus $\sim 1$ in the free state, indicating the stabilization of intra- or intermolecular hydrogen bonds.

### Table 1. Oligonucleotides used for DNA-binding studies.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD19 $^a$</td>
<td>5’- CGTTGGTCAACCTCAUTCCCCCAT 3’- CCACTGATGCTGGCTGTTATA</td>
<td>25</td>
</tr>
<tr>
<td>CD19-N $^a$</td>
<td>5’- GCTTGGTCAACCC 3’- CCACCAAGGTGCGGGTTA</td>
<td>12</td>
</tr>
<tr>
<td>CD19-C $^a$</td>
<td>5’- TCAGTCCCCCAT 3’- AGCTAACCAGGTTA</td>
<td>12</td>
</tr>
<tr>
<td>mb-1 $^b$</td>
<td>5’- GTCCGGAAAGTTGGCTCAATGCCTCT 3’- GCAAGCTCTATACCGGGATCCGGGA</td>
<td>27</td>
</tr>
<tr>
<td>mb-1-N $^b$</td>
<td>5’- GATGGGCTC 3’- ATCTAACGCGAG</td>
<td>12</td>
</tr>
<tr>
<td>mb-1-C $^b$</td>
<td>5’- GTCTCACTTGGCG 3’- CAGGCTACGGG</td>
<td>12</td>
</tr>
<tr>
<td>G(T)4(C)2(A)4(C)2T)4(C)2(A)4G</td>
<td>5’- GTTTTCCAAAATTTTTTCCAAAAAG 3’- AAAAAATTTTTGTGTTTTTC</td>
<td>24</td>
</tr>
<tr>
<td>G(T)4(A)4</td>
<td>5’- GTTTTTAAAAAC 3’- AAAATTTTTTGG</td>
<td>12</td>
</tr>
<tr>
<td>T(G)3(C)A</td>
<td>5’- TGCCCCCCCCACT 3’- ACCCCAAAAAGT</td>
<td>12</td>
</tr>
</tbody>
</table>

$^a$ Corresponding to CD19-2_Ains [23]. The approximate half-sites for the NTD and CTD are denoted as -N and -C, respectively. Methylated G residues that disrupt binding are highlighted in bold [23].

$^b$ Corresponding to the mb-1 promoter sequence in the Pax5/Ets-1/DNA complex [20]. Base pairs contacted directly in the major groove by Pax5 are highlighted in bold.
Asymmetry in DNA-binding by the Pax5 Paired domain

Fig. 3. Upon binding DNA, the β-hairpin and linker regions of the PD become ordered. (a) Comparison of 15N-HSQC spectra of Pax5-1-149 free (blue) and bound to the CD19 DNA (orange). Numerous amide signals, including those from residues flanking the helical bundles, change chemical shifts in the presence of DNA. (b) The backbone 13Cα/13Cβ/13C=O/15N/1HN chemical shifts of free and CD19-bound Pax5-1-149 were assigned (Supplementary Fig. S1). Shown are the normalized α-helical and β-strand propensities, as well as RCI-S2 values, per residue based on these shifts using the MICS algorithm [40]. The top cartoon indicates the secondary structure of the PD as in Fig. 1c. Although the helical bundle secondary structure of the PD does not change upon DNA binding, the β-hairpin and linker regions become more ordered as evidenced by their increased RCI-S2 values. (c) The amide 1H N, 15N chemical shift perturbations (CSPs, ppm) resulting from DNA binding are plotted (left) and mapped onto the cartoon representation of the PD (right, from PDB ID: 1MDM). Blank values correspond to prolines or amides with unassigned signals. Residues with values above 0.3 ppm (dashed line) are highlighted in green, and those for which there is no information or the CSP is below 0.3 are in blue. The largest CSPs localize to amides within the DNA recognition helices H3 and H6, the linker, and the β-hairpin region. Amide 15N (d) T2, (e) T1, and (f) heteronuclear NOE relaxation data confirmed that the linker and N-terminal residues are conformationally mobile in free Pax5-1-149, with unusually long T2 lifetimes and low or negative NOE values. However, these regions become ordered when DNA bound, with relaxation parameters similar to those of the NTD and CTD helical amides. The * denote clipped histogram bars for residues Asn148 and Gln149 with T2 (and NOE) values of 0.44 s (–0.6) and 0.80 s (<0.3) and 1.2 s (<0.8) for the free protein and 0.80 s (<0.3) and 1.2 s (<0.8) for the complex, respectively.

This is consistent with the structure of the PD in the Pax5/Ets1/DNA ternary complex. For example, the amides of Leu23 and Phe27 in the β-hairpin donate hydrogen bonds to Asn21 and Gly19, respectively, whereas those in the linker region (e.g., Gly85) interact with DNA. Furthermore, many amides

---

**Summary:**

- Upon binding DNA, the β-hairpin and linker regions of the Pax5 Paired domain become more ordered, as evidenced by changes in amide chemical shifts and increased RCI-S2 values.
- The largest CSPs are observed in regions involved in DNA recognition, such as helices H3 and H6, the linker, and the β-hairpin.
- Relaxed NOE data confirm the increased order of these regions upon DNA binding, indicating conformational changes.

**Note:** The text and diagrams are aligned with the figure, providing a comprehensive view of the structural changes and the data supporting these observations.
Throughout the NTD and CTD helical bundles of the Pax5\textsuperscript{1-149}/CD19 complex had PFs > 10\textsuperscript{7} when bound to CD19 DNA, representing dramatic increases of at least 10\textsuperscript{5}- or 10\textsuperscript{4}-fold, respectively, relative to the free protein. These data hint that the stabilities of the helical bundles become similar when bound to cognate DNA. However, due to the very slow exchange behavior of the PD/DNA complex, we can only provide minimum PFs for the most protected amides. Therefore, we cannot rule out that the subdomains have different stabilities in the bound state. In addition, exchange could occur within the bound state and/or via transiently unbound forms of the protein. Nevertheless, these HX measurements demonstrate that Pax5\textsuperscript{1-149} is dramatically stabilized upon binding a cognate DNA sequence.

**Amide \(^{15}\text{N}\) relaxation studies confirm the dynamic properties of free and DNA-bound Pax5\textsuperscript{1-149}**

To characterize the sub-nanosecond timescale dynamics of the PD, we collected amide \(^{15}\text{N}\) relaxation data \([T_1, T_2,\) and heteronuclear nuclear Overhauser enhancement (NOE)] of Pax5\textsuperscript{1-149} in the absence and presence of DNA (Fig. 3d–f). In its free form, amides throughout the helical bundle subdomains had relatively uniform \(T_1\) and \(T_2\) lifetimes and heteronuclear \(^{15}\text{N}-\text{NOE}\) values of ~0.75. This is indicative of well-defined helical structures and limited motions of the \(^1\text{H}^{15}\text{N}\) bonds on the sub-nanosecond timescale \([50]\). On the other hand, amides within the N-terminal ~30 residues and linker regions of free Pax5\textsuperscript{1-149} showed distinctly long amide \(T_2\) lifetimes and low or negative heteronuclear NOE values. Together with their random coil chemical shifts, low RCI-S\textsuperscript{2} values and PFs ~ 1, these data clearly demonstrate that these regions of Pax5\textsuperscript{1-149} are conformationally disordered in the absence of DNA. When bound to DNA, the relaxation properties of the linker and the N-terminal \(\beta\)-hairpin more closely match those of the helical subdomains. Thus, in contrast to the “beads-on-a-string” behavior of free Pax5\textsuperscript{1-149}, the entire PD becomes well ordered when in complex with DNA.

Fitting the relaxation data measurements \([T_1, T_2,\) \(^{15}\text{N}-\text{NOE}\)] for the most ordered amides in unbound Pax5\textsuperscript{1-149} to the model-free formalism using Tensor2 [51] yielded effective rotational correlation times for isotropic global tumbling of 8.9 and 8.2 ns for the NTD and CTD, respectively. Although the assumption of isotropic rotation is an oversimplification, these values are consistent with the similar masses of the NTD and CTD regions within the ~17 kDa Pax5\textsuperscript{1-149} protein fragment, and the fact that the two subdomains can tumble semi-independently due to their separation by a flexible linker. By way of comparison, the algorithm hydroNMR [52] predicts global rotational correlation times of 4.3 and 5.2 ns for the isolated NTD and CTD subdomains, respectively, and 19 ns for the full-length rigid Pax5\textsuperscript{1-149} PD. In the case of the CD19-bound Pax5\textsuperscript{1-149}, which is anisotropic with helical bundle domains joined by an ordered linker, analysis of the \(^{15}\text{N}\) relaxation data yielded a rotational diffusion tensor with \(D_{xx} = 4.9 \times 10^6\) s\(^{-1}\), \(D_{yy} = 5.6 \times 10^5\) s\(^{-1}\), and \(D_{zz} = 1 \times 10^7\) s\(^{-1}\). This corresponds to an effective tumbling time of 24 ns and is compatible with the increased mass of ~32 kDa for the 1:1 Pax5\textsuperscript{1-149}/DNA complex. The near axially symmetric diffusion tensor for the protein, with the \(z\)-axis parallel to the expected position of the long axis of the DNA double helix, is also consistent with Pax5\textsuperscript{1-149} adopting a well-defined, elongated structure when bound to DNA.

**Both subdomains of Pax5\textsuperscript{1-149} contribute to overall binding affinity towards CD19 DNA**

To determine the relative contribution of each subdomain to DNA binding, we measured the equilibrium dissociation constants \((K_d)\) values of Pax5\textsuperscript{1-149}, Pax5\textsuperscript{1-92}, and Pax5\textsuperscript{76-149} for the full CD19 DNA or its half-sites, CD19-N and CD19-C (Table 1). The latter were defined based on approaches including sequence comparisons, mutagenesis, and chemical modification studies from previous reports [23,53]. Using an electrophoretic mobility shift assay (EMSA), the \(K_d\) value of Pax5\textsuperscript{1-149} for CD19 DNA was determined to be 5 ± 2 nM (Table 2, Supplemental Fig. S3). In contrast, the individual Pax5\textsuperscript{1-92} and Pax5\textsuperscript{76-149} fragments bound their respective half-site DNAs with \(K_d\) values of 11 ± 0.4 μM and 16 ± 0.3 μM, respectively, as measured under similar experimental conditions by isothermal titration calorimetry (ITC) (Table 2, Supplemental Fig. S4). The change of more than three orders of magnitude in the \(K_d\) value for Pax5\textsuperscript{1-149} versus those for the two fragments indicates that both contribute to the overall affinity for the full-length CD19 DNA. This is expected for multivalent interactions involving the bipartite PD and an extended DNA sequence. Surprisingly, Pax\textsuperscript{1-92} did not measurably bind the CD19-C half-site, whereas Pax5\textsuperscript{76-149} bound both the CD19-N (12 ± 1 μM) and CD19-C (16 ± 0.3 μM) half-sites with similar moderate affinities (Table 2, Supplemental Fig. S4). As discussed below, the DNA-binding specificity of Pax5 appears to be set by the NTD, which discriminates cognate and non-specific binding sites.

DNA binding was also monitored by \(^{15}\text{N}-\text{HSQC}\) experiments. Pax5\textsuperscript{1-149} bound the CD19 DNA sequence in the slow exchange limit (bound spectrum shown in Fig. 3a and Supplementary Fig. S1). That is, the exchange rate constant between the free and bound states of the protein \((k_{on} = k_{off}[\text{DNA}] + k_{off})\) was much smaller than the chemical shift difference between these two states, \(|\Delta\omega|\) (reviewed in Ref. [50]). As a result, the \(^{1}\text{H}^{15}\text{N}\) signals of the free protein disappeared upon titration with DNA, while
new signals from the bound protein concomitantly appeared. Such tight binding effectively precludes the determination of a $K_d$ value by NMR spectroscopy. However, it is in agreement with that of 5 nM measured by EMSA.

Upon titration of Pax51–92 with its CD19-N half-site, many amides exhibited binding in the intermediate–slow exchange regime (Fig. 4a). With $k_{ex} \ll |\Delta \omega|$, this behavior is characterized by moderate shifting and significant broadening of the $^1H_N-^{15}N$ signals (in most cases, to the point of disappearance), followed by sharpening and reappearance with new chemical shifts over the course of the titration. Such exchange broadening also precluded the extraction of a $K_d$ value. However, given the weak affinity measured by ITC ($K_d \sim 11 \mu M$) for the Pax51–92/CD19-N interaction (Table 2), severe broadening was somewhat unexpected. As explained below, this likely resulted from a combination of large chemical shift changes ($|\Delta \omega|$) and relatively slow association and dissociation kinetics ($k_{on}$ and $k_{off}$).

In contrast, Pax576–149 bound the CD19-C half-site in the fast–intermediate exchange regime where $k_{ex} \gg |\Delta \omega|$ (Fig. 4b). This is characterized by moderate broadening and progressive changes of amide signals from their unbound position as DNA is initially added, followed by the sharpening of amide signals at the bound chemical shift position as saturation is reached. In reasonable agreement with ITC data, fitting of the NMR-monitored titration curves yielded a $K_d$ value of $26 \pm 5 \mu M$ for Pax576–149 with the CD19-C half-site (Table 2).

Our results also indicated that, whether separated or linked, the NTD and CTD bound the CD19 half-sites via the same general interfaces as identified in the Pax5/Ets-1/DNA ternary complex by X-ray crystallography (Fig. 5) [20]. Although we did not assign the spectrum of the Pax51–92/CD19-N complex, new dispersed signals appearing around 9 to 10 ppm upon addition of DNA matched those with largest CSP in the N-terminal region of the Pax51–149/CD19 complex (Figs. 3a, c and 4a, and Supplementary Fig. S1). Furthermore, a plot of the relative loss of amide signal intensities from the free state of Pax51–92 upon addition of CD19-N indicated that residues throughout the helical bundle and the preceding $\beta$-hairpin were perturbed (Fig. 5a). In the case of Pax576–149, most assignments for the bound state could be obtained by tracking $^1H_N-^{15}N$ signals over the course of the titration. Amides showing the largest CSPs upon binding CD19-C also matched those most affected in the complex of Pax51–149 with CD19 (Fig. 5b). Not surprisingly, the linker residues showed greater CSPs in the context of the full PD rather than the subdomains. For example, Lys87 had a CSP value of 0.82 ppm in the Pax51–149/CD19 complex (Fig. 3c), but only 0.49 ppm in the Pax576–149/CD19-C complex (Fig. 5b). Thus, the high affinity of the PD in Pax51–149 for CD19 DNA arises from the combined binding of the NTD and CTD to their respective half-sites, augmented by positioning the ordered linker to lie along the intervening DNA minor groove.

The $\beta$-hairpin contributes to DNA binding

The ordering of the $\beta$-hairpin upon DNA binding represents an entropic cost that must be offset by favorable enthalpic interactions. To assess the net
thermodynamic contribution of the N-terminal region of the PD to DNA affinity, we also expressed 15N-labeled Pax51-92 lacking the residues required to form this hairpin structure. The amide 1H-15N chemical shifts of this protein were almost identical to those of Pax5 1-92, indicating that the helical bundle remained stable in the absence of the disordered residues 1–31 (not shown). Upon addition of CD19-N DNA to Pax51-92, we observed similar residues being perturbed by DNA association as seen with Pax51-92 (Supplemental Fig. S5). However, the exchange behavior became faster, and fitting the NMR-monitored titration curves yielded a $K_d$ of 85 ± 30 μM (Table 2). This is ~ 8-fold larger than the value of 11 ± 0.4 μM measured by ITC for Pax51-92 to the CD19-N DNA, and thus the N-terminal β-hairpin region of the PD indeed contributes to its net DNA binding affinity.

The subdomains of Pax51-149 exhibit different binding properties for DNA half-sites. Shown are 15N-HSQC-monitored titrations of the CD19 half-sites into (a) 15N-labeled Pax51-92 and (b) Pax576-149, as well as the mb-1 half sites into (c) 15N-labeled Pax51-92 and (d) Pax576-149. Pax51-92 bound CD19-N in the slow–intermediate exchange regime, such that many amide signals initially shifted and broadened to disappearance and then reappeared with new chemical shifts, shown enclosed in dotted circles. In contrast, Pax51-92 bound the mb-1-N half-site weakly, exhibiting signal broadening, but no substantial chemical shift changes even after addition of more than three molar equivalents of DNA. Pax576-149 bound both half-sites in the fast exchange limit, allowing determination of the $K_d$ values listed in Table 2. The molar ratios of protein:DNA are indicated by the color codes, and selected assignments are provided. Solid arrows indicate the amide chemical shift changes from the free to the bound states. Dotted arrows are included for amides that initially disappear and reappear outside the spectral window shown.

The subdomains of Pax51-149 contribute differently towards binding mb-1 DNA

The mb-1 promoter is a well-characterized binding site for the partnership of Pax5 and Ets-1 [20,22,54–56]. Although this DNA sequence does not conform closely to the Pax5 consensus sequence [23], upon addition of the 27 bp mb-1 DNA duplex to Pax51-149, we also observed slow exchange behavior in 15N-HSQC spectra (bound spectrum shown in Supplementary Fig. S6). Thus, consistent with a $K_d$ ~ 2.4 nM reported by Fitzsimmons et al. [22], Pax51-149 can bind this sequence with high affinity, even in the absence of Ets-1. We also investigated the binding of the mb-1 half-sites to the Pax5 fragments by NMR spectroscopy (Table 1). Pax51-92 interacted very weakly with the mb-1-N half-site, showing only small CSPs even in the presence
Amides affected by DNA span the helical bundle of Pax51-92, in contacting the DNA.

The most perturbed amides still clustered near mb-1-N and palindromic sequences had more moderate unbound signals decreased to near baseline values. The different chemical shifts when bound to CD19-N, and hence including the β-hairpin. Signals from many amides had different chemical shifts when bound to CD19-N, and hence unbound signals decreased to near baseline values. The mb-1-N and palindromic sequences had more moderate effects, yet the most perturbed amides still clustered near helix H3. This is consistent with their behavior in the Pax51-149/CD19 complex (Fig. 3) and highlights the importance of these residues in helix H4 and the preceding linker, helix H5, and the recognition helix H6 showed the greatest CSPs, thus defining the DNA-binding interface. Addition of CD19-C caused the largest CSPs, followed by the palindromic sequence T(G)5(C)5A. During the titrations with these sequences, we were unable to track residues Val90, Trp112, Ile138, and Arg140 due to the large chemical shift differences of their free versus bound forms. This is consistent with their behavior in the Pax51-149/CD19 complex (Fig. 3) and highlights the importance of these residues in contacting the DNA.

Line broadening (Fig. 4d). A $K_d$ value of 380 ± 120 μM was obtained by fitting the titration data (Table 2). Amides showing the largest CSPs mapped to the same surface of Pax55-149, affected by the addition of the CD19-C half-site (Fig. 5b), indicating that the CTD uses a common binding interface to interact with the mb-1 and CD19 DNAs. Although the CD19-C and mb-1-C sequences are more similar to one another than are the CD19-N and mb-1-N sequences (Table 1), the overall behaviors of the NTD and CTD toward these half-site DNAs are strikingly different. In particular, the NTD shows greater DNA sequence discrimination than the CTD.

**The two subdomains of Pax55-149 differ in non-specific DNA binding**

Since the vast majority of genomic DNA does not correspond to Pax5 regulatory sites, we also investigated PD binding to non-specific DNA sequences by NMR spectroscopy. We initially used a 24 bp pseudo-palindrome (GTTTTCCAAAACCTT TTCCAAAAAG, Table 1) for studies with Pax55-149. Exchange broadening occurred over the course of the titration, demonstrating that Pax55-149 is able to bind DNA sequences that do not conform to a consensus site (Supplementary Fig. S6). This broadening also precluded spectral assignments and the estimation of a $K_d$ value (Table 2). Nevertheless, in the presence of a 3.6 molar excess of this pseudo-palindromic DNA duplex, several amides had dispersed chemical shifts similar to those observed in the spectra of the CD19 and mb-1 complexes (Supplemental Fig. S6). Although only tentatively assigned, these amides were mostly located in the CTD. In contrast, the dispersed signals from NTD amides seen with the two specific complexes were absent in the spectrum of the non-specific complex (Supplemental Fig. S6). This indicates that the CTD primarily mediates binding of Pax55-149 to the pseudo-palindromic DNA.

To further dissect the contribution of the subdomains to non-specific DNA binding, we chose two simple 12 bp palindromic sequences for titrations with Pax55-92 and Pax55-76-149 (Table 1). Surprisingly, Pax55-92 exhibited minimal CSPs upon addition of nearly 3-fold molar excess of either the 5'-G(T)5(A)5C3- or 5'-T(G)5(C)5A3 duplexes (Fig. 6a and b). Accordingly, we estimate $K_d$ values greater than 500 μM for Pax55-92 with either of these palindromes (Table 2). Nevertheless, residues throughout the helical bundle region showed patterns of reduced amide signal intensities similar to those seen with the CD19-N and mb-1-N DNAs (Fig. 5b). Therefore, Pax55-1-92 interacts with non-specific and specific DNA via the same canonical DNA-binding interface. This conclusion is also supported by the observation that the few amide signals with clear CSPs upon titration with...
either palindrome, such as Cys64, Ser61, and Leu69, map to this interface. Addition of the 
\[ \text{T(G)}_3 \text{A}_3 \] palindromic duplex to Pax51-92 lacking the \( \beta \)-hairpin region also did not result in any substantial spectral perturbations (Supplementary Fig. S5). Parenthetically, this eliminates the possibility that residues 1–31 in the NTD.

Unlike Pax51-92, Pax576-149 bound both palindromes in fast exchange and exhibited large amide CSPs, as shown in Fig. 5b. Solid arrows show the direction of the amide CSP from free to bound states. Dotted arrows indicate amide shifts that cannot be tracked or reappear outside the spectral window shown. Fitting the latter titrations yielded the \( K_d \) values in Table 2.

![Fig. 6](image)

Fig. 6. In contrast to the CTD, the NTD subdomain of Pax51-149 only weakly interacts with non-specific DNA. The addition of the two palindromic DNAs to (a, b) \( ^{15} \text{N-labeled Pax5}^{1-92} \) and (c, d) Pax576-149 was monitored by \( ^{15} \text{N-HSQC} \) experiments. The protein:DNA molar ratios are indicated by the color codes, and selected assignments are provided. Pax51-92 was only modestly perturbed by the presence of excess DNA with small amide intensity changes plotted in Fig. 5a. On the other hand, Pax576-149 bound both palindromes in fast exchange and exhibited large amide CSPs, as shown in Fig. 5b. Solid arrows show the direction of the amide CSP from free to bound states. Dotted arrows indicate amide shifts that cannot be tracked or reappear outside the spectral window shown. Fitting the latter titrations yielded the \( K_d \) values in Table 2.

Discussion

Structure, stability, and dynamics of the un-bound Pax5 PD

Based on backbone chemical shift, \( ^{15} \text{N} \) relaxation, and amide HX measurements, Pax51-149 folds as two independent three-helix bundles separated by a conformationally disordered linker. Similar to Pax8 [31], free Pax5 lacks any stable secondary structure in the region encompassing the first ~30 residues at the N-terminus of the protein, including the \( \beta \)-hairpin.
This result is also consistent with molecular dynamics simulations of Pax6 and DNA, which predict that the \( \beta \)-hairpin structure requires stabilizing contacts provided by DNA [31]. Given the small number of residues involved in the formation of the \( \beta \)-hairpin and the lack of tertiary contacts or disulfide bonds, this result is not unexpected.

Although well folded, we found that the helical bundles of the PD exhibited small PFs in the range of only \( ~10^{-5} \) to \( 10^{-3} \). Thus, the NTD and CTD are dynamic and readily undergo conformational fluctuations detectable by HX. More importantly, the helical bundles differ in their stability, with the CTD helices having PFs \( ~10 \) times greater than the NTD. Assuming that the slowest amides exchange through global fluctuations, this means that the free energy of unfolding of the NTD is approximately 2/3 that of the CTD. The reasons for this difference in stabilities are not obvious, as both subdomains have similar helical bundle folds and charge distributions. However, the NTD and CTD do not show any obvious sequence similarity, and their tertiary structures do not superimpose closely (backbone RMSD of \( \sim 2 \) Å in PDB ID:1MDM). In addition, the helices that form the NTD versus CTD bundles differ in length in all PD structures reported to date. In particular, helix H2 is relatively short and predominantly polar, and does not contribute substantially to the hydrophobic core of the NTD subdomain. In contrast, residues Ile99, Ile114, and Leu118, which were found to be among the most protected from HX in Pax5[149], make hydrophobic and van der Waals contacts that likely stabilize the tertiary structure of the CTD. Additionally, the loops separating the helices in each subdomain are longer in the CTD, which may allow for better positioning of the helices relative to each other in order to maximize tertiary contacts and minimize exposed hydrophobic surfaces. Another possible explanation for the difference in stability is the position of the side chain of Arg38 in the context of the Pax5/Ets-1/DNA ternary complex. This arginine is invariant in the NTD of all nine human Pax proteins and is almost fully buried in the Pax5/Ets-1/mb-1 crystal structure, with only a small portion of its guanidinium group exposed to interact with the DNA phosphodiester backbone. In addition, the positioning of its positive charge in close proximity to the side chain of Arg71 is expected to be unfavorable. Recognizing that it is difficult to predict stability from protein sequence or structure, and that the structure of the free PD of Pax5 has not been determined, it is certainly plausible that these differences contribute to the \( ~10 \)-fold lower PFs for the NTD versus CTD subdomain.

Finally, it is worth mentioning that Pax5[149], which carries two exposed cysteine residues in helices H1 and H2 near the DNA-binding interface, was found to be very sensitive to oxidation and prone to aggregation (not shown). Addition of dithiothreitol (DTT) reducing agent partly reversed this aggregation and restored the original protein fold. In contrast, Pax5[76-149], which has one cysteine residue in the loop between helices H5 and H6, did not readily aggregate under similar conditions. Several studies have linked PD proteins, including Pax5, to redox regulation of transcriptional activity, whereby only the reduced forms of the protein can bind to DNA. In its oxidized state, intramolecular disulfide bonds prevent formation of the DNA complex [31,57,58]. In the case of Pax5, this process seems to be mediated by the redox modulator APE/Ref-1 [58]. It is conceivable that the dynamic nature of the NTD plays a role in this mechanism, allowing the cysteine residues to be more susceptible to redox-induced modifications.

The Pax5 subdomains contribute differently to DNA binding

The PD makes extensive contacts with DNA. Upon binding the CD19 duplex, amides within the recognition helices of the NTD and CTD showed substantial \( ^1H^N,N \) chemical shift changes. The N-terminal \( \beta \)-hairpin and flanking loops, as well as the interdomain linker, also exhibited large amide shift perturbations accompanied by reduced subnanosecond timescale mobility, indicative of structural ordering. In addition, amides within the \( \beta \)-hairpin, linker region, and both helical bundles became markedly more protected from HX. These changes are consistent with the X-ray crystallographic structure of the Pax5/Ets-1/mb-1 ternary complex, in which the two recognition helices dock within the major groove of DNA, while the \( \beta \)-hairpin/loop and intervening linker residues provide minor groove contacts.

Our studies also showed that both the NTD and CTD contribute to the net affinity of Pax5 for the CD19 DNA. The separate subdomains exhibited similar (~10 to 20 \( \mu \)M) dissociation constants for their respective CD19-N and -C half-sites. As expected for multivalent interactions, the binding affinity is much stronger (5 nM) for the intact PD with the full-length CD19 DNA. Surprisingly, the NTD interacted very weakly with all other DNA sequences tested. Because most DNA-binding modules have at least some weak affinity for random sequences due to electrostatic interactions with the phosphodiester backbone, this result was rather unexpected [59]. In contrast, the CTD bound the two CD19 half-sites with ~10 \( \mu \)M affinities and bound the mb-1-C and two palindromic DNAAs with \( K_d \) values between ~150 and 400 \( \mu \)M. Thus, unlike the NTD, the CTD showed only modest sequence specificity and was able to measurably interact with all DNA sequences tested. In agreement with these results, previous studies noted more variability in the DNA sequences bound...
by the CTD relative to the NTD [23,25]. It is also interesting to note that one of the structures of the Pax5/Ets-1/DNA complex determined by X-ray crystallography (PDB ID: 1K78) contained an extra CTD bound to a pseudo-consensus site present in the DNA duplex [20].

The reason behind the disparity in DNA recognition by the NTD, which binds only specific DNA sequences, and the CTD, which associates with DNA rather indiscriminately, is unclear. Both subdomains have similar predicted isoelectric points (10.2 and 10.0 for Pax51-92 and Pax576-149, respectively) [60] and a similar number of positive charges at their DNA-binding interfaces. Also, in the Pax5/Ets-1/mb-1 complex, both subdomains provide comparable number of base-specific contacts, such as those involving His62 and Asn29 in the N-terminal region and Ser133 and Arg137 in the C-terminal region. Most other contacts are to the DNA phosphodiester backbone, hinting that indirect readout of the sequence-dependent DNA shape may be important for specificity [61]. The β-hairpin and following loop, which make key contacts with DNA, are also likely involved in setting the specificity of the N-terminal fragment of the PD.

Besides the differences in affinity for various DNA sequences, our data suggest that the two subdomains exhibit different binding kinetics ($k_{on}$ and $k_{off}$). Although we did not measure this directly, we found that both Pax51-92 and Pax576-149 had similar dissociation constants for their respective CD19 half-sites (~11 and 16 μM, respectively), yet exhibited markedly different exchange behaviors between their free and bound states. In the case of the NTD-containing fragment, binding occurred in the intermediate–slow exchange regime ($k_{ex} \leq |\Delta \omega|$), whereas the CTD-containing fragment exhibited mostly fast exchange ($k_{ex} \geq |\Delta \omega|$). This cannot be attributed to differences in $|\Delta \omega|$ values alone because several amide signals in the NTD (e.g., Arg50 and Leu69) exhibited small chemical shift changes, yet displayed slow exchange behavior. Since the $k_{d}$ values ($= k_{off}/k_{on}$) are comparable for both fragments, this indicates that the exchange rate constant $k_{ex} (= k_{on}[DNA] + k_{off})$ must be smaller for Pax51-92 than Pax576-149. Overall, these results indicate that NTD of the PD associates and dissociates from DNA more slowly than the CTD. As elaborated upon in the following paragraph, this could reflect a conformational change in the NTD, but not CTD, required for DNA binding. Part of this change could involve the disorder–order transition of the β-hairpin residues. However, upon deletion of these N-terminal residues, the $k_{d}$ value of the NTD increased from ~11 μM to ~85 μM, yet this subdomain still showed exchange broadening in the presence of CD19-N (Supplementary Fig. S5). Thus, the postulated conformational changes likely also occur in the helical bundle of the NTD.

The relationship between protein stability and DNA-binding specificity is worth examining. The less stable NTD undergoes more facile structural fluctuations detectable by amide HX measurements than the CTD and also exhibits greater binding specificity. A recent study comparing the structures of free and DNA-bound proteins in 90 cases where high-resolution data was available for both states found that larger conformational changes upon complex formation were associated with greater specificity [62]. In the case of the PD of Pax5, the increased dynamics and flexibility of the NTD may reduce the time spent in a DNA binding-competent conformation. In turn, this would impact both the thermodynamics and kinetics of DNA complex formation. In the case of the CTD subdomain, greater stability may be important in maintaining the orientation of positively charged residues involved in making non-specific contacts. In support of this idea, the helical regions of the NTDs of DNA-bound Pax5 (X-ray crystallography, PDB ID: 1MDM) and highly homologous DNA-free Pax8 (NMR spectroscopy, PDB ID: 2K27) do not align, as closely as their CTD helices. In addition, Pax8 has a poorly defined helix H3 that adopts only ~1.5 turns in its free state [31]. Taken together, this suggests that the N-terminal regions of PDs generally have more structural heterogeneity than their C-terminal regions, which in turn impacts their DNA binding properties.

Finally, our results also show that DNA binding, regardless of the strength of the interaction and DNA sequence, occurs through the same canonical protein interfaces. In the presence of the CD19 and mb-1 DNA duplexes, the $^{1}H-^{15}N$ spectra of Pax51-149 showed similar changes in the amide signal positions. Additionally, amides in Pax51-92, Pax576-149, and Pax576-149 exhibiting signal intensity or chemical shift changes upon titration with CD19, mb-1, and non-specific palindromic half-sites mapped to similar
Asymmetry in DNA-binding by the Pax5 Paired domain

Implications in transcriptional regulation and disease

Protein dynamics and weak non-specific interactions between transcription factors and the DNA phosphodiester backbone are critical for the function of these proteins in the nuclei of eukaryotic cells [74,75]. A recent genome-wide study of Pax5 occupancy revealed that it was associated with a large fraction of accessible DNA sites in cells of the B-lineage [25]. This finding challenges the view that a small number of transcription factors must rapidly scan a vast amount of DNA in order to reach their cognate regulatory sites. Instead, it appears that Pax5 is constantly associated with a substantial portion of available DNA sites in pro-B and mature B-cells. Given that the CTD can bind DNA indiscriminately, this subdomain may be crucial in facilitating the widespread association of Pax5 with open chromatin. In contrast, the NTD seems to function in ensuring that high affinity complexes are formed only in the presence of a binding site closely resembling the consensus site and/or in the presence of protein partners that stabilize the ternary complex.

Due to its central role in B-cell differentiation, Pax5 is closely linked to oncogenic processes. Roughly 20 different fusion proteins involving this protein have been found in association with various types of cancers [76]. The majority of these chimeric products retain the PD, and many have been shown to maintain its nuclear localization and DNA-binding capacity [16,77–83]. This highlights the crucial role of the PD in tumorigenesis. Furthermore, a large number of Pax5 mutations associated with B-ALL localize to residues in the NTD and linker regions. These include Gly30, Ser66, and Gly85, which are involved in making contacts with DNA [16,18,84,85]. Other mutations in this region could affect binding to co-factors required for transcriptional regulation, as well as the structural integrity of the dynamic NTD helical bundle. In contrast, relatively few B-ALL-related changes have been identified in the CTD. This suggests that this subdomain may be less susceptible to missense mutations and is consistent with its greater structural stability and higher affinity for non-specific DNA sequences relative to the NTD.

In closing, our study provides mechanistic details of DNA binding by the PD and sheds light into the role of the two subdomains in forming a functional unit to regulate transcription. Given the high degree of conservation in the PD of Pax genes across the animal kingdom, the differences we observed in subdomain stability and DNA-binding modes likely hold true for other members of this transcription factor family.

Materials and Methods

Protein expression and purification

The genes encoding Pax51-149, Pax51-92, Pax532-92, and Pax576-149 were cloned from the full-length Pax5 gene (NCBI Gene ID: 5079) into the pET28-MHL vector (Addgene, plasmid #26096) using NdeI and HindIII restriction sites. This vector encodes an N-terminal His6 affinity tag followed by a TEV cleavage site. Unlabeled proteins were expressed in Escherichia coli BL21 (DE3) cells grown in LB media, whereas isotopically labeled proteins were produced using M9 minimal media supplemented with 3 g/L 13C6-glucose and/or 1 g/L 15NH4Cl as the sole carbon and nitrogen sources, respectively. Uniformly 2H13C15N-labeled Pax51-149 was produced in M9 minimal media, using a protocol modified from that published for preparing deuterated proteins [86]. Briefly, a 25 mL starter culture was grown to OD600 ~ 0.6 in LB media (H2O) at 37 °C. The cells were then collected by centrifugation and resuspended in 75 mL of M9 media prepared with 99% D2O. Protonated additives for the M9/D2O media were dissolved in D2O and lyophilized prior to use. The bacterial culture was allowed to reach OD600 ~ 0.6 and diluted 4-fold with fresh M9/D2O media. This was repeated until reaching the final culture volume of 1 L.

Protein expression was induced at OD600 ~ 0.6 with 0.5 mM IPTG, followed by growth at 30 °C for 16 hours. After centrifugation, the cell pellet was frozen at -80 °C, then later thawed, resuspended in denaturing buffer [4 M guanidinium HCl, 20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole (pH 7.4)], and sonicated to ensure complete lysis. Denaturation also led to full amide protonation in the otherwise uniformly 2H13C15N-labeled Pax51-149. The cleared supernatant was applied to a Ni2+-NTA HisTrap HP column (GE Healthcare). After binding, the column was washed with 20 mM sodium phosphate, 0.5 M NaCl, 20 mM imidazole at pH 7.4 to allow on-column refolding, followed by elution with 20 mM sodium phosphate, 0.5 M NaCl, and 1 M imidazole at pH 7.4. The appropriate fractions were pooled and the His6 affinity tag...
cleaved with TEV protease during a dialysis step against 20 mM Tris-HCl, 2 mM DTT at pH 8.0. Two (or three in the case of Pax5\(^{1-149}\) and Pax5\(^{52-92}\)) non-native amino acid residues (Gly-His-Met) remained at the N-terminus of each construct. Uncleaved protein and cleaved tag were removed using a HisTrap HP column. A subsequent size-exclusion chromatography (Superdex 75, GE Healthcare) step was used to increase sample purity and for buffer exchange. Unless noted otherwise, the final buffer used for NMR experiments contained 20 mM MES, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT, and 6 mM MgCl\(_2\) at pH 6.5 for 10 min, followed by slow cooling to room temperature. The resulting double-stranded DNAs were purified using size-exclusion chromatography with sample buffer (Superdex 75, GE Healthcare) to remove single-stranded DNA and adjust the salt concentration. In the case of the CD19 duplex, were collected using the Bruker Avance III 600 MHz spectrometer with standard pulse sequences [93]. In the case of the DNA complex, the NaCl concentration was lowered to 20 mM to improve spectral quality and increase cryoprobe sensitivity. The \(T_1\) and \(T_2\) curves for well-resolved amide signals (peak intensity \textit{versus} relaxation time) were fit to single exponential decays using Sparky [89]. Errors were estimated using a Monte Carlo approach.

The heteronuclear \(^1H-^{15}N\) NOE values were calculated as the ratios of the peaks intensities in the NOE spectrum (5 s relaxation delay followed by 3 s of \(^1H\) irradiation) \textit{versus} a control reference spectrum (6 s delay without \(^1H\) irradiation). The resulting data were fit with Tensor2 [51] to obtain global tumbling correlation times and, in the case of CD19-bound Pax5\(^{1-149}\), an anisotropic rotational diffusion tensor. Selected structural coordinates were taken from the PDB file 1MDM for fitting of the diffusion tensor and for the calculations of rotation correlation times with hydroNMR [52].

**HX rates and protection factors**

To obtain the HX rates of slowly exchanging amides, Pax5\(^{1-149}\) in 500 \(\mu\)L of NMR buffer (pH 6.5) was lyophilized. The dry protein sample was then resuspended in 500 \(\mu\)L of \(D_2O\) and immediately placed in the spectrometer. \(^{15}N\)-HSQC spectra were collected at 25 °C in succession every 5 min, starting \~5 min after resuspension. However, only a few amide signals were detected in the first spectrum collected. Therefore, the experiment was repeated using NMR buffer at pH 5.5 and 15 °C to decrease the exchange rate and enable quantitation of the HX rates for a greater number of amides. Proton-deuterium exchange rate constants for residues with well-resolved \(^1H-N\) signals were obtained using Sparky [89]. Peak intensities were fit to the equation \[I_T = I_0 e^{-k_{HX-obs}t}\] where \(I_T\) is the observed peak height at time \(t\) after resuspension in \(D_2O\), \(I_0\) is the fit initial height, and \(k_{HX-obs}\) is the fit exchange rate constant. In the case of the Pax5\(^{1-149}/CD19\) complex, a similar procedure was performed at pH\* 6.60 and 25 °C. The initial 5-min HSQC spectrum was collected \~12 min after resuspension. The last spectrum was collected approximately a month later and still contained many amide signals that had not decayed sufficiently to obtain a reliable \(k_{HX-obs}\). Therefore, only a lower PF limit is provided in Fig. 2c for these residues with \(k_{HX-obs} < 1.8 \times 10^{-2}\) s\(^{-1}\).

To obtain \(k_{HX-obs}\) values for rapidly-exchanging amides for the free PD, we employed a CLEANEX-PM sequence [94] using mixing times from 4 to 160 ms at 25 °C. Pax5\(^{1-149}\) protein samples in NMR buffer at pH 5.6, 6.3, and 8.0 were used to maximize the number of residues with detectable proton-proton exchange. The resulting growth curves were fit to the equation

\[
\frac{I_T}{I_0} = \frac{k_{HX-obs}}{k_{IX-obs} + R_1} \left(1 - e^{-k_{IX-obs} + R_1}t\right)
\]

where \(I_T\) is the amide peak height for transfer time \(t\), \(I_0\) is the corresponding height in the reference spectrum (without

**DNA oligonucleotides**

The sequences of oligonucleotides used in this study are summarized in Table 1. All single-stranded DNA oligonucleotides were purchased from Integrated DNA Technologies. Non-palindromic complementary strands were mixed in a 1:1 ratio based on the quantities reported by the vendor. All duplexes were annealed by heating to \~100 °C in sample buffer (20 mM MES, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT, 6 mM MgCl\(_2\) at pH 6.5) for 10 min, followed by slow cooling to room temperature. The resulting double-stranded DNAs were purified using size-exclusion chromatography with sample buffer (Superdex 75, GE Healthcare) to remove single-stranded DNA and adjust the salt concentration. In the case of the CD19 duplex used for EMSA, one strand (5′CGGTGGTCACG) was 5′-labeled with Alexa Fluor 647 for detection. The two strands were mixed and adjusted the salt concentration. In the case of ITC and NMR-monitored DNA binding titrations, 6 mM MgCl\(_2\) was included, and the NaCl concentration of each construct. Uncleaved protein and cleaved tag were removed using a HisTrap HP column. In the subsequent size-exclusion chromatography (Superdex 75, GE Healthcare) step was used to increase sample purity and for buffer exchange. Unless noted otherwise, the final buffer used for NMR experiments contained 20 mM MES, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT, and 6 mM MgCl\(_2\) at pH 6.5 for 10 min, followed by slow cooling to room temperature. The resulting double-stranded DNAs were purified using size-exclusion chromatography with sample buffer (Superdex 75, GE Healthcare) to remove single-stranded DNA and adjust the salt concentration. In the case of the CD19 duplex used for EMSA, one strand (5′CGGTGGTCACG) was 5′-labeled with Alexa Fluor 647 for detection. The two strands were mixed and annealed as described above. Concentrations of the purified dsDNA were determined by ultraviolet absorbance at 280 nm using predicted molar absorptivities [60].

**NMR spectroscopy**

NMR experiments were performed using cryoprobe-equipped Bruker Avance III 500, 600, or 850 MHz spectrometers. Unless noted otherwise, protein samples were concentrated to 0.3–0.8 mM in 95% NMR sample buffer [20 mM MES, 200 mM NaCl, 2 mM DTT, 0.5 mM EDTA (pH 6.5)] with 5% lock D\(_2O\) and data were collected at 25 °C. The spectra were processed and analyzed using NMRPipe [88] and Sparky [89]. The backbone (\(^{15}C\)α, \(^{13}C\)\(\gamma\), \(^{13}C\)O, \(^{1}N\), and \(^{15}N\)) chemical shifts were assigned using standard \(^1H\)-\(^13\)C-\(^15\)N scalar correlation experiments [90]. In the case of amide-protonated \(^1H\)-\(^15\)C/\(^15\)N-labeled Pax5\(^{1-149}\) in complex with the 25 bp CD19 DNA duplex, the NaCl concentration was lowered to 20 mM and transverse relaxation optimized spectroscopy (TROSY)-based pulse sequences with \(^2H\)-decoupling were employed [91,92]. Secondary structure propensity and RCI-S\(^2\) calculations were carried out using MICS [40].

**Amide \(^{15}N\) relaxation**

Amide \(^{15}N\) relaxation data (\(T_1\), \(T_2\), heteronuclear NOE) for Pax5\(^{1-149}\), both free and in complex with the CD19 DNA duplex, were collected using the Bruker Avance III 600 MHz spectrometer with standard pulse sequences [93]. In the case of the DNA complex, the NaCl concentration was lowered to 20 mM to improve spectral quality and increase cryoprobe sensitivity. The \(T_1\) and \(T_2\) curves for well-resolved amide signals (peak intensity \textit{versus} relaxation time) were fit to single exponential decays using Sparky [89]. Errors were estimated using a Monte Carlo approach.

The heteronuclear \(^1H\)-\(^{15}N\) NOE values were calculated as the ratios of the peaks intensities in the NOE spectrum (5 s relaxation delay followed by 3 s of \(^1H\) irradiation) \textit{versus} a control reference spectrum (6 s delay without \(^1H\) irradiation). The resulting data were fit with Tensor2 [51] to obtain global tumbling correlation times and, in the case of CD19-bound Pax5\(^{1-149}\), an anisotropic rotational diffusion tensor. Selected structural coordinates were taken from the PDB file 1MDM for fitting of the diffusion tensor and for the calculations of rotation correlation times with hydroNMR [52].
The NMR spectral assignments for Pax51-149 alone and in complex with CD19 DNA have been deposited in the database.
BioMagResBank under accession numbers 26730 and 26731, respectively.

Acknowledgments

We thank Dr. James Hagman (National Jewish Health, Denver, CO) for providing the human Pax5 gene, Genevieve Desjardins and Kevin Zhang for making the original expression plasmids used for this study, members of the Van Petegem laboratory for assistance with ITC experiments, and the Center for Drug Research and Development for use of the Typhoon Imager. C.P.B. was supported through the Genome Sciences and Technology Program by the Natural Sciences and Engineering Research Council (NSERC) of Canada CREATE grant. This research was funded by the Canadian Cancer Society Research Institute (CCSRI 2011-700772 to L.P.M.) and the Canadian Institutes of Health Research (CIHRMOP-136834 to L.P.M.). Instrument support was provided by the Canada Foundation for Innovation (CFI), the British Columbia Knowledge Development Fund (BCKDF), the UBC Blusson Innovation (CFI), the British Columbia Knowledge Society Research Institute (CCSRI 2011-700772 to L.P.M.) and the Michael Smith Foundation for Health Research (MSFHR).

Appendix A. Supplementary Data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jmb.2016.04.004.

Received 16 January 2016;
Received in revised form 1 April 2016;
Accepted 1 April 2016
Available online 8 April 2016

Keywords:
transcription factor;
B-cell development;
protein dynamics;
hydrogen exchange;
NMR spectroscopy

Abbreviations used:
Pax, Paired box; PD, Paired domain; NTD, N-terminal subdomain; B-ALL, B-cell acute lymphoblastic leukemia; CTD, C-terminal subdomain; HX, hydrogen exchange; HSQC, heteronuclear single quantum correlation; RCI-S2C, random coil index-squared order parameter; pH, pH meter reading without correction for isotope effects; PF, protection factor; CSP, chemical shift perturbation; NOE, nuclear Overhauser enhancement; EMSA, electrophoretic mobility shift assay; ITC, isothermal titration calorimetry; TROSY, transverse relaxation optimized spectroscopy.

References


[53] B. Adams, P. Dorfler, A. Aguzzo, Z. Kozmin, P. Urbanek, I. Maurer-Fogy, M. Busslinger, Pax-5 encodes the transcription factor BSAP and is expressed in B lymphocytes, the developing CNS and adult testis, Genes Dev. 6 (1992) 1589–1607.


Asymmetry in DNA-binding by the Pax5 Paired domain


SUPPLEMENTAL INFORMATION

Structural and dynamics studies of Pax5 reveal asymmetry in stability and DNA binding by the Paired domain

Cecilia Perez-Borrajero\textsuperscript{1,2}, Mark Okon\textsuperscript{2}, and Lawrence P. McIntosh\textsuperscript{2,3,4*}

\textsuperscript{1} Genome Sciences and Technology Program, University of British Columbia, Vancouver, Canada V6T 1Z3
\textsuperscript{2} Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, Canada V6T 1Z3
\textsuperscript{3} Department of Chemistry, University of British Columbia, Vancouver, Canada V6T 1Z1
\textsuperscript{4} Michael Smith Laboratories, University of British Columbia, Vancouver, Canada V6T 1Z4
Supplemental Fig. S1. Assigned $^{15}$N-HSQC spectra of $^{15}$N-labeled (a) Pax$^{1-92}$, (b) Pax5$^{76-149}$, and (c) Pax5$^{1-149}$ in their unbound forms, and (d) $^2$H/$^{15}$N/$^{13}$C-labeled Pax5$^{1-149}$ bound to CD19 DNA. All spectra were collected at pH 6.5 and 25 °C. A TROSY-based HSQC pulse sequence was used for the DNA complex. Regions within dashed boxes are expanded for clarity.
Supplemental Fig. S2. The β-hairpin structure is stabilized upon binding CD19 DNA. (a) Type I and (b) type II turn propensities obtained from a MICS analysis of backbone chemical shifts of Pax4<sup>1-149</sup> show that residues Gln22 and Asn29 flanking the β-hairpin adopt chemical shifts consistent with the ordering and stabilization of this small secondary structure. Gln22 is located in the loop between the two β-strands and Asn29 forms a turn into helix 1 of the NTD.
Supplemental Fig. S3. Electrophoretic mobility shift assay (EMSA, left) used to quantitate the interaction between Pax5\textsuperscript{1-149} and full-length CD19 DNA. The lower band corresponds to free Alexa Fluor 647 fluorescently-labeled CD19 DNA and the upper band corresponds to the Pax5\textsuperscript{1-149}/CD19 complex. The resulting binding curve derived from this gel is shown on the right. The fraction bound was calculated as the ratio of band intensity of the Pax5\textsuperscript{1-149}/DNA complex to total DNA band intensities, and plotted as a function of total protein concentration (right). Two independent data sets, each measured twice on separate gels, were fit to a simple 1:1 binding model and the results were averaged to yield the reported $K_d$ value ± standard deviation of 5 ± 2 nM (see Methods and Table 2). The assay was carried out in 20 mM MES, 100 mM NaCl, 6 mM MgCl\textsubscript{2}, 6 mM DTT, 0.2 mM EDTA, 200 µg/mL bovine serum albumin and 10% glycerol at pH 6.5 and 4 °C.
(a) Pax5^{1-92}

- **Time (min)**
  - 0, 10, 20, 30, 40

- **µcal/sec**
  - 0.00, -2.00, -4.00, -6.00, -8.00, -10.00

- **kcal mol^{-1} of injectant**
  - 0.00, -2.00, -4.00, -6.00, -8.00, -10.00

- **Molar Ratio**
  - 0.0, 1.0, 2.0, 3.0, 4.0, 5.0, 6.0

- **K_d = 10.8 +/- 0.4 µM**
- **N = 2.94 +/- 0.01**
- **ΔH = -10230 +/- 70 cal/mol**
- **ΔS = -11.6 cal/mol/deg**

(b) Pax5^{76-149}

- **Time (min)**
  - 0, 10, 20, 30, 40

- **µcal/sec**
  - 0.00, -2.00, -4.00, -6.00, -8.00, -10.00

- **kcal mol^{-1} of injectant**
  - 0.00, -2.00, -4.00, -6.00, -8.00, -10.00

- **Molar Ratio**
  - 0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0

- **K_d = 16.3 +/- 0.3 µM**
- **N = 1.09 +/- 0.00**
- **ΔH = -8220 +/- 45 cal/mol**
- **ΔS = -5.67 cal/mol/deg**

- **K_d = 12 +/- 1 µM**
- **N = 1.16 +/- 0.02**
- **ΔH = -6650 +/- 140 cal/mol**
- **ΔS = 0.293 cal/mol/deg**
Supplemental Fig. S4. The DNA-binding subdomains of Pax5 have similar affinities for their corresponding CD19 half-sites. However, the CTD can associate with both half-sites, whereas the NTD only binds its cognate CD19-N sequence. The equilibrium dissociation constants for (a) Pax5\textsuperscript{1-92} and (b) Pax5\textsuperscript{76-149} towards the CD19-N and CD19-C half-sites were measured by ITC experiments. The top panels contain the raw buffer-corrected data with the heats produced at each injection of concentrated protein into DNA solution. The bottom panels show the corresponding Wiseman plots of integrated heats. The curves were fit with Origin to a simple binding model and the resulting values for dissociation constant (K\textsubscript{d}), stoichiometry (N), enthalpy change (ΔH), and entropy change (ΔS) are shown (see Table 2). Addition of Pax5\textsuperscript{1-92} to CD19-C DNA yielded very small heat changes (note the different scale relative to CD19-N) and the titration data were not fit. The fit stoichiometry value of 2.94 for the binding of Pax5\textsuperscript{1-92} with CD19-N may result from the propensity of Pax5\textsuperscript{1-92} to aggregate at high concentrations. These measurements were carried out in 20 mM MES, 100 mM NaCl, 0.5 mM EDTA, 2 mM DTT, and 6 mM MgCl\textsubscript{2} at pH 6.5 and 25 °C.
Supplemental Fig. S5. Deletion of the β-hairpin weakens DNA binding by the NTD. Shown are $^{15}$N-HSQC-monitored titrations of $^{15}$N-labeled Pax5$^{32-92}$ with (a) CD19-N and (b) non-specific palindromic DNA. Pax5$^{32-92}$ bound CD19-N in the near fast exchange limit, allowing the determination of a $K_d$ value of 85 ± 30 µM (averaged from the fit titration curves of 10 residues; Table 2). In contrast, Pax5$^{1-92}$ bound this specific half-site DNA with higher affinity (11 ± 0.4 µM) and in the intermediate-slow exchange regime (Figure 4). Both Pax5$^{1-92}$ and Pax5$^{32-92}$ only weakly interacted with the non-specific DNA, and thus the β-hairpin does not impair binding. In these spectra, the molar ratios of protein:DNA are indicated by the color codes, and selected assignments provided. Solid arrows indicate the amide chemical shift changes from the free to the bound states. Dotted arrows are included for amides that initially disappear and reappear outside the spectral window shown.
Supplemental Fig. S6. Pax5\textsuperscript{1-149} binds specific and non-specific DNAs. Shown are overlaid \textsuperscript{15}N-HSQC spectra of \textsuperscript{15}N-labeled Pax5\textsuperscript{1-149} free or in the presence of the DNA duplexes CD19, mb-1, and the 24 bp pseudo-palindrome G(T)\textsubscript{4}(C)\textsubscript{2}(A)\textsubscript{4}(C)\textsubscript{2}(T)\textsubscript{4}(C)\textsubscript{2}(A)\textsubscript{4}G (Table 1). The DNA:protein molar ratios were 1:1, 3:1 and 3.6:1, respectively. Binding occurred in the slow exchange limit for specific complexes involving the CD19 and mb-1 DNAs, whereas exchange broadening was observed over the course of the titration with the pseudo-palindromic sequence (Table 2). The spectra of the specific complexes are similar, containing well dispersed amide signals in the 9 to 10 ppm region that are not seen with the free protein. However, the pseudo-palindromic complex also showed chemical shift perturbations. This demonstrates that Pax5\textsuperscript{1-149} associates with this non-specific DNA sequence, albeit more weakly than with specific DNAs. Although only
the spectrum of Pax5\textsuperscript{1-149}/CD19 was assigned, representative signals within the dashed ovals likely arise from the same amides in all three complexes (Pax5\textsuperscript{1-149}/CD19, Pax5\textsuperscript{1-149}/mb-1, and Pax5\textsuperscript{1-149}/pseudo-palindrome) and localize to the CTD. On the other hand, \textsuperscript{1}H\textsuperscript{15}N signals within boxes are similar in the specific Pax5\textsuperscript{1-149}/CD19 and Pax5\textsuperscript{1-149}/mb-1 complexes, but absent in the non-specific Pax5\textsuperscript{1-149}/pseudo-palindrome complex. These correspond mostly to amides in the NTD and linker regions. Thus, the CTD binds all three DNAs similarly, whereas the NTD discriminates between specific and non-specific DNAs. All spectra were collected at reduced ionic strength (20 mM MES, 20 mM NaCl, 2 mM DTT, 0.5 mM EDTA, pH 6.5).