Identification of the Single-stranded DNA Binding Surface of the Transcriptional Coactivator PC4 by NMR*

(Received for publication, August 10, 1998, and in revised form, November 9, 1998)

Sebastiaan Werten‡‡, Rainer Wechselberger‡, Rolf Boelens‡, Peter C. van der Vliet‡, and Robert Kaptein¶¶

From the ‡Laboratorium voor Fysiologische Chemie, Universiteit Utrecht, Stratenum, Universiteitsweg 100, 3584 CG Utrecht, The Netherlands and ¶¶Bijvoet Centrum voor Biomoleculair Onderzoek, Universiteit Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

The C-terminal domain of the eukaryotic transcriptional cofactor PC4 (PC4CTD) is known to bind with nanomolar affinity to single-stranded (ss)DNA. Here, NMR is used to study DNA binding by this domain in more detail. Amide resonance shifts that were observed in a ¹H¹⁵N-HSQC-monitored titration of ¹⁵N-labeled protein with the oligonucleotide dT₁₈ indicate that binding of the nucleic acid occurs by means of two anti-parallel channels that were previously identified in the PC4CTD crystal structure. The β-sheets and loops that make up these channels exhibit above average flexibility in the absence of ssDNA, which is reflected in higher values of T₁pop, reduced heteronuclear nuclear Overhauser effects and faster deuterium exchange rates for the amides in this region. Upon ssDNA binding, this excess flexibility is significantly reduced. The binding of ssDNA by symmetry-related channels reported here provides a structural rationale for the preference of PC4CTD for juxtaposed single-stranded regions (e.g. in heteroduplexes) observed in earlier work.

The eukaryotic general transcriptional cofactor PC4 is known to enhance activated in vitro and in vivo transcription from various RNA polymerase II promoters, in concert with members of all major classes of transcriptional activators (1–3). The activator-dependent stimulation of transcription by PC4 has been shown to originate mostly from increased recruitment of basal transcription factors during the early stages of preinitiation complex (PIC) formation (4). It is likely that this effect is caused by the interactions that have been reported between PC4 and the activation domains of promoter-binding activator proteins on the one hand, and basal factors (including TFIIA) on the other (1, 3). Thus, PC4 appears to act as a bridging factor that stabilizes the preinitiation complex.

Recently it has become clear that PC4 has several additional functions. In the first place, the protein was found to be able to repress transcription under specific conditions (5–7). Repression at moderate PC4 concentrations is observed in minimal in vitro transcription systems, that normally do not require the basal factor TFIIH (5, 6), and with respect to specific transcription initiation by RNA polymerase II from DNA ends and unwound regions (6). Repression is not observed in full-factor in vitro transcription, because TFIIH is able to overcome repression by PC4 (5, 6). Thus, PC4 appears to act as what could be called a fidelity-enhancing factor for RNA polymerase II transcription initiation by repressing adventitious transcription from incomplete PICs on the one hand, and on the other hand enhancing transcription from fully assembled PICs, i.e. those that include TFIIA, generally believed to be the last factor to enter the PIC during assembly (8). Surprisingly, PC4 has recently also been found to associate in vitro with the RNA polymerase III transcription factor TFIIIC and to stimulate transcription by RNA polymerase III by increasing the rate of reinitiation (9). Interestingly, copurification of PC4 with the human replication protein A (RPA) and effects of PC4 observed in a reconstituted SV40 replication system suggested that the protein may play a further role in DNA replication (10). Thus, it is emerging that PC4 is an extremely versatile factor, involved in the processing of genetic information at multiple levels.

Although stimulation of RNA polymerase II transcription by PC4 is probably largely effected by protein-protein interactions depending on the N-terminal half of the protein, the C-terminal half of the protein in addition comprises a powerful ssDNA-binding domain (3, 4, 11, 12). The physiological role of this domain has long remained enigmatic. Recently, binding of single-stranded DNA was found to be entirely dispensable for transcription activation by PC4 but strictly required for repression of transcription in the absence of TFIIH (6). Earlier analysis of PC4CTD (12) has shown that it binds with nanomolar affinity to juxtaposed stretches of ssDNA running in opposite directions, such as present in partially melted double-stranded DNA or heteroduplexes. Optimal binding to such structures requires approximately 8 unpaired nucleotides in each of the two strands. Single-stranded oligonucleotides can be bound with almost equal affinity, provided these molecules contain a minimum of 16–20 nucleotides. These observations support the idea that ssDNA has to be bent 180° to create an arrangement of anti-parallel strands similar to the topology of a heteroduplex for PC4 to bind.

Earlier, the crystal structure of PC4CTD was solved (13). This structure revealed a novel homodimeric fold in which each monomer consists of a curved four-stranded β-sheet and a 45°-kinked α-helix, both contributing to a complex dimerization interface. A prominent feature of the structure is the presence of two channel-like depressions running in anti-parallel directions. These channels, formed by the curved β-sheets, are lined both by positively charged and aromatic residues, suggesting involvement in ssDNA binding in a manner similar to the DNA binding mode of other single-stranded DNA-binding proteins.

* This work was supported by the Netherlands Foundation for Chemical Research (SON) with financial support from the Netherlands Organization for Scientific Research (NWO). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Bijvoet Centrum voor Biomoleculair Onderzoek, Padualaan 8, 3584 CH Utrecht, The Netherlands. Tel.: 31-30-2533787; Fax: 31-30-2537623; E-mail: kaptein@nmr.chem.uu.nl.

‡‡ The abbreviations used are: PIC, preinitiation complex; RPA, replication protein A; ss, single-stranded; NOE, nuclear Overhauser effect.
such as RPA (14). To investigate whether the channels that were identified indeed constitute the DNA binding surface of PC4
CTD, we have studied this protein domain and its interaction with a single-stranded oligonucleotide by means of heteronuclear NMR.

**EXPERIMENTAL PROCEDURES**

**NMR Sample Preparation**—PC4
CTD was overexpressed in *Escherichia coli* strain BL21 (DE3) using pET-11a (Novagen) constructs. The PC4
CTD construct encodes amino acids 63–127 of the full-length protein, preceded by Met-Ala, originating from the vector sequence. Bacteria were grown at 37 °C in synthetic medium (6.0 g/liter Na2HPO4, 2 H2O, 3.0 g/liter KH2PO4, 0.5 g/liter NaCl, 1.0 mM MgSO4, 0.2 g/liter FeCl3, 20 μM CaCl2, 0.2 mg/liter thiamine) containing 0.5 g/liter 15NH4Cl as the sole nitrogen source and either 4.0 g/liter 12C-glucose, or 1.0 g/liter 13C-glucose as the only carbon source. Induction took place at an A600 nm of 0.5, by the addition of isopropyl-β-D-thiogalactopyranoside to a final concentration of 1 mM. After 4 h the cells were harvested, resuspended in buffer A (20 mM Tris-HCl, pH 7.3, 10% glycerol, 1 mM EDTA, 5 mM dithiothreitol, 10 mM Na2S2O5, 1 mM phenylmethylsulfonyl fluoride, 10 μg/ml aprotinin, 20 μg/ml leupeptin) containing 500 mM KCl and lysed by freeze-thawing (once) and sonication for a total of 5 min (in bursts of 20 s). The lysate was then centrifuged at 20,000 × g for 20 min at 4 °C. The supernatant was diluted to 200 mM KCl by adding buffer A without KCl and immediately loaded onto a heparin-Sepharose column (Amersham Pharmacia Biotech). The column was washed using buffer A at 200 mM KCl and eluted with buffer A containing 500 mM KCl. Purest fractions, selected on the basis of Coomassie-stained SDS-polyacrylamide gel electrophoresis gels, were pooled and, after dilution with buffer A without KCl to a final concentration of 75 mM KCl, applied to an SP-Sepharose column (Pharmacia), followed by washing with buffer B (20 mM potassium phosphate, pH 5.5, 5 mM dithiothreitol) containing 75 mM KCl and elution by means of a linear salt gradient (75–1,000 mM KCl) in buffer B. Peak fractions were further purified using a Superdex 75 column (Pharmacia in the same buffer containing 400 mM KCl and subsequently concentrated by means of Centricon spin dialysis tubes (molecular mass cut-off 10 kDa, Amicon). Protein samples at this stage appeared as single bands on silver-stained polyacrylamide gels. Protein concentrations were determined using Bio-Rad protein assays (Bio-Rad) employing bovine gammaglobulin as a standard. Deuterated glycine was added to the final samples to an end concentration of 2 M, because this resulted in better long-term stability of the protein at

![Fig 1. 1H15N-HSQC spectrum (600 MHz) of PC4
CTD, in which backbone assignments have been indicated. Side chain amide groups are indicated by horizontal lines but have not been assigned.](image-url)
Identification of the ssDNA Binding Surface of PC4

PC4<sub>CTD</sub> backbone assignments

Backbone assignment for PC4<sub>CTD</sub> (in ppm). The first Aα residue is not part of the PC4 sequence but originates from the expression vector. For several residues (Ile-66, Glu-93, Glu-95, and Met-96), multiple sets of signals could be observed.

| Amino Acid | Free Protein | Complex |
|------------|--------------|---------|
| N<sub>H1</sub> | C<sub>H1</sub> | CO | N<sub>H2</sub> |
| Ala | 126.6 (8.99) | 52.0 (3.94) | 172.7 | 126.8 (8.90) |
| Met-63 | 121.0 (8.50) | 54.7 (5.06) | 174.8 | 121.0 (8.21) |
| Phe-64 | 121.0 (9.25) | 56.6 (ND) | 174.7 | 121.2 (9.28) |
| Glu-65 | 128.2 (9.36) | 57.7 (ND) | 176.7 | 128.0 (9.73) |
| Ile-66 | 118.3 (8.34) | 60.6 (4.50) | 173.0 | 118.2 (8.35) |
| Glu-67 | 118.7 (8.20) | ND (ND) | ND (ND) |
| Gly-68 | 117.8 (8.20) | ND (ND) | ND (ND) |
| Gly-69 | 118.4 (8.20) | ND (ND) | ND (ND) |
| Lys-70 | 116.6 (7.50) | 44.8 (ND) | 176.5 | 116.8 (7.54) |
| Lys-71 | 120.6 (8.52) | 58.9 (ND) | ND (ND) |
| Met-69 | ND (ND) | ND (ND) | ND (ND) |
| Arg-70 | ND (ND) | ND (ND) | 174.2 | ND (ND) |
| Tyr-71 | 119.9 (8.81) | 56.4 (5.21) | 175.1 | 119.8 (8.89) |
| Val-72 | 121.7 (9.40) | 60.4 (5.34) | 175.5 | 121.9 (9.45) |
| Ser-73 | 122.5 (9.59) | 55.8 (5.40) | 173.6 | 122.5 (9.67) |
| Val-74 | 124.8 (8.39) | 61.3 (4.86) | 175.2 | 124.7 (8.35) |
| Arg-75 | 124.5 (8.56) | 54.5 (4.66) | 172.8 | 124.5 (8.51) |
| Asp-76 | 119.6 (8.41) | 52.6 (5.07) | 175.9 | 119.5 (8.44) |
| Phe-77 | 126.1 (9.16) | 56.9 (4.76) | 175.4 | 126.5 (9.23) |
| Lys-78 | 122.0 (9.00) | 57.0 (5.00) | 176.1 | 119.6 (9.17) |
| Gly-79 | 110.5 (8.55) | 46.0 (3.98) | 170.8 | 110.4 (8.51) |
| Met-79 | ND (ND) | ND (ND) | ND (ND) |
| Asp-81 | 118.1 (8.39) | 50.3 (4.31) | 174.3 | 118.2 (8.43) |
| Leu-91 | 122.9 (8.58) | 55.6 (5.14) | 173.9 | 122.9 (8.63) |
| Ser-92 | 122.5 (9.67) | 55.8 (5.40) | 173.6 | 122.5 (9.67) |
| Val-93 | 121.3 (9.43) | 60.4 (5.34) | 175.5 | 121.9 (9.45) |
| Arg-94 | ND (ND) | ND (ND) | ND (ND) |
| Gly-95 | 118.3 (8.34) | 60.6 (4.50) | 173.0 | 118.2 (8.35) |
| Ala | 126.8 (8.90) | 52.0 (3.94) | 172.7 | 126.8 (8.90) |

Higher temperatures.

The oligonucleotide dT<sub>18</sub> was purchased from Carl Roth GmbH & Co. The DNA was dissolved in buffer B containing 400 mM KCl and 2 M deuterated glycine to an oligonucleotide concentration of 11 mM.

NMR spectroscopy—<sup>15</sup>N-{sup}</sup>H<sup>-</sup>NMR spectra were recorded at 325 K (32°C) on Bruker AMX500, AMX600, Varian UnityPlus 500, and UnityPlus 750 spectrometers. The protein concentration was 1.5 mM in all experiments.

Protein-DNA titrations were carried out by repeated addition of small aliquots of an 11 mM solution of dT<sub>18</sub> to a 1.5 mM PC4<sub>CTD</sub> sample. Formation of the protein-DNA complex was observed through the recording of a 500 MHz <sup>1H</sup>-<sup>15</sup>N-HSQC spectrum after each addition.<sup>15</sup>N-T<sub>r</sub> and <sup>15</sup>N-T<sub>r</sub> relaxation and heteronuclear NOE experiments were carried out at 500 MHz as described earlier (15–17), using pulsed-field gradients for coherence selection in combination with the sensitivity enhancement scheme (18–20). The longitudinal <sup>15</sup>N relaxation rates were determined from a series of 7 spectra with delays of 32, 64, 128, 256, 512, 768, and 1024 ms. The transverse in-phase <sup>15</sup>N relaxation rates (21) were determined from a series of 7 spectra with delays of 5, 16, 32, 48, 64, 96, and 128 ms, using a 1<sup>5</sup>N-spin lock with a field strength of 1.4 kHz. Heteronuclear cross relaxation constants were derived from two 500 MHz spectra recorded with and without 2.5 s of saturation of the amide protons.

To measure the exchange rate of amide protons, the protein was first lyophilized and then redissolved. Almost all amide protons, followed by the recording of a series of <sup>1H</sup>-<sup>15</sup>N-HSQC spectra, up to 200 h after redissolving the protein. All spectra were processed on Silicon Graphics O2 workstations using the software package NMR Pipe (22). Longitudinal and transversal in-phase relaxation rates and the errors in these parameters were obtained by curve-fitting of a mono-exponential function through the peak intensities according to the Levenberg-Marquardt algorithm using in-house developed software, as described by Vis et al. (17).

RESULTS AND DISCUSSION

The 600 MHz <sup>1H</sup>-<sup>15</sup>N-HSQC spectrum of PC4<sub>CTD</sub> is shown in Fig. 1. A nearly complete backbone assignment (listed in Table I) was obtained using double and triple resonance methods, essentially in the manner described by Vis et al. (23). Thus, <sup>15</sup>N and <sup>1</sup>H frequencies were obtained from a high-resolution 600 MHz <sup>1H</sup>-<sup>15</sup>N-HSQC spectrum, followed by the collection of corresponding intra- and inter-residual C<sup>b</sup>- frequencies (24) and, if detectable, side chain resonances obtained from a 750 MHz total correlation spectroscopy-HSQC spectrum. Nearly all H<sub>b</sub> frequencies could also be obtained from this total correlation spectroscopy-HSQC spectrum. Remaining ambiguities in the sequential assignment were resolved using short range NOE distances from a 750 MHz NOE spectroscopy-HSQC.

To investigate which parts of PC4<sub>CTD</sub> are involved in the interaction with ssDNA, a titration was performed with the oligonucleotide dT<sub>18</sub>. The choice of an 18-mer for our titration experiments was based on earlier experiments that indicated that PC4<sub>CTD</sub> interacts with approximately 16–20 residues of ssDNA (12). To investigate which parts of PC4<sub>CTD</sub> are involved in the interaction with ssDNA, a titration was performed with the oligonucleotide dT<sub>18</sub>. The choice of an 18-mer for our titration experiments was based on earlier experiments that indicated that PC4<sub>CTD</sub> interacts with approximately 16–20 residues of ssDNA (12). To investigate which parts of PC4<sub>CTD</sub> are involved in the interaction with ssDNA, a titration was performed with the oligonucleotide dT<sub>18</sub>. The choice of an 18-mer for our titration experiments was based on earlier experiments that indicated that PC4<sub>CTD</sub> interacts with approximately 16–20 residues of ssDNA (12). To investigate which parts of PC4<sub>CTD</sub> are involved in the interaction with ssDNA, a titration was performed with the oligonucleotide dT<sub>18</sub>. The choice of an 18-mer for our titration experiments was based on earlier experiments that indicated that PC4<sub>CTD</sub> interacts with approximately 16–20 residues of ssDNA (12).
maxima in the HSQC spectra with and without an equimolar amount of dT18. Although small shifts (around 10 Hz) in peak positions are observed for residues throughout the protein, all perturbations exceeding 20 Hz map to the \( \beta \)-sheet region. Especially large shifts of 30 Hz or more are seen for the outermost parts of \( \beta \)-sheets, \( \beta_2 \) and \( \beta_3 \), and the loop that connects them (residues Phe-77, Lys-78, Gly-79, Lys-80, and Val-81), the “\( \beta \)-ridge” that separates the two anti-parallel channels (residues Trp-89, Met-90, Asp-91, and Lys-101), and the region that connects sheet \( \beta_4 \) to the \( \alpha \)-helix (residues Leu-105 and Asn-106). In addition, a surprisingly large perturbation was found for the N-terminal Ala residue, which originates from the expression vector sequence but is located in close proximity to the ends of the anti-parallel channels in the crystal structure.
Fig. 4. Values of $^{15}$N-$T_1$, $^{15}$N-$T_{1r}$, heteronuclear $1H^{15}$N-NOE and hydrogen exchange rate for PC4-CTD backbone amides. The exchange rates were classified as follows. Fast (F), no peak was observed in the first HSQC spectrum after redissolving the lyophilized protein in D$_2$O; medium (M), peaks were still observable in the first spectrum, but the intensity significantly decreased in 200 h; slow (S), no significant decrease in peak intensity could be detected in 200 h. The secondary structure of PC4-CTD has been indicated schematically below the graph.

Fig. 5. $^{15}$N-$T_{1r}$ values for the amides in free (gray) and bound PC4-CTD (black). The values obtained from curve fitting of peak intensities are shown with error bars indicating single standard deviations. The secondary structure of PC4-CTD has been indicated schematically below the graph.
Fig. 3, the perturbations of Fig. 2 are shown in the PC4 CTD crystal structure by means of color coding, with colors ranging from green (indicating no perturbation) through white to red (corresponding to a perturbation of 30 Hz or more). As this figure shows, the data support a model in which the ssDNA is held in between the $\beta_2\beta_3$-arms and the $\beta$-ridge of the protein.

Consistent with current views concerning protein-ssDNA interactions (25), we find a large number of positively charged and aromatic residues at the binding surface with their side chains exposed to solution (Lys-68, Arg-70, Tyr-71, Arg-75, Phe-77, Lys-80, Arg-86, Tyr-88, Trp-89, Lys-97, Arg-100, and Lys-101). Interestingly, superposition of one of the PC4 anti-parallel channels onto the ssDNA binding channel found in the RPA-ssDNA cocrystal structure revealed that several of these residues (Arg-75, Phe-77, Arg-86, Tyr-88, Trp-89, Lys-97, Arg-100) occupy positions with respect to the channel that correspond well to the positions of similar residues in the two RPA subunits (13, 14). This indicates that PC4 and RPA may have similar ways of contacting ssDNA even though the $\beta$-strand topology is markedly different. For residues Tyr-71, Arg-75, Phe-77, Lys-78, Lys-80, Arg-86, Tyr-88, Trp-89, and Lys-101, large ssDNA-induced amide resonance perturbations were observed in our titration. We expect these residues to make side chain contacts with the DNA. For Trp-89, involvement in binding has recently been confirmed by mutagenesis (6). For residues Lys-68, Arg-70, Tyr-71, Lys-78, Lys-80, Arg-86, Tyr-88, Trp-89, Lys-97, and Arg-100, no assignments could be obtained, but the position of these residues with respect to the binding channels strongly suggests that their side chains are also involved in binding.

Comparison with other single-stranded DNA-binding proteins suggests a particularly important role in ssDNA binding for the ends of $\beta$-strands, $\beta_2$ and $\beta_3$, and the loop $\beta_2\beta_3$ that connects them (residues 77–80). This part of the protein is reminiscent of typical ssDNA-binding loops found in many other single-stranded DNA-binding proteins, like the $\Lambda_{20}$-loop of the OB-fold proteins (26). It contains an aromatic residue (Phe-77) and several positively charged residues (Lys-78 and Lys-80, with Arg-75 located nearby in $\beta$-strand $\beta_2$). In the two RPA subunits, the corresponding strands and the connecting loop ($\beta_4\beta_5$) are seen to bind ssDNA through stacking of the aromatic residue Phe-269 (in subunit A) or Phe-386 (subunit B) onto a DNA base, as well as through interactions of the positively charged residue Lys-263 (A) or Arg-382 (B) with the phosphate backbone. The NMR data are in agreement with a similar role of the $\beta_2\beta_3$-loop of PC4 in ssDNA binding, because the ssDNA-induced amide resonance perturbations of all residues in this region are among the largest observed. The importance of the $\beta_4\beta_5$-loop region for ssDNA binding is further
supported by a recent mutagenesis study (6).

Temperature factors from the crystallographic structure determination of PC4_{CTD} (13) suggested that the β2β3-loop and the β-ridge exhibit above average mobility. To assess the flexibility of these regions in solution, we carried out relaxation measurements (15N-T_{1p}, 15N-T_{2}, and H15N-NOE) for the backbone amides. Results are shown in Fig. 4. Although 15N-T_{1p} did not show significant variation along the protein sequence, 15N-T_{2p} values are indeed higher in this region than in the rest of the protein. Considerably lower heteronuclear 1H15N-NOE values were observed in the β2β3-loop and the β-ridge compared with the rest of the molecule. Consistent with these relaxation data, fast hydrogen exchange was observed in both the β2β3-loop and the β-ridge (Fig. 4, lower panel), whereas most of the amides in the core of the protein exhibit either medium or slow exchange. Thus, the regions of PC4 that mediate ssDNA contacts are relatively flexible in solution. Upon formation of the complex with dT_{18}, T_{1p} values for amide nitrogens throughout the protein are reduced, most likely because of the larger size of the complex (Fig. 5). However, the reduction of signals from residues in the β2β3-loop region is significantly larger than this uniform decrease in signal intensity. This effect may be caused by attenuation of the higher mobility in this region, which can be understood in terms of stabilization by ssDNA contacts.

DNA binding by the two channels as demonstrated here is consistent with and provides support for existing biochemical data in several ways. In the first place, high affinity for heteroduplex structures and the apparent requirement for two juxtaposed single-stranded regions (12) are readily explained by the present model: each binding channel can contact one of the strands running in opposite directions, as illustrated in Fig. 6. Moreover, the binding site size (approximately 8 nucleotides in each of the two single-stranded regions) as determined from the affinities for both heteroduplexes and single-stranded oligonucleotides of increasing size (12), corresponds closely to what would be expected on the basis of the channel length in the PC4 structure: if it is assumed that ssDNA bound by PC4 is similar in conformation to ssDNA bound by RPA, 8 nucleotides should span a distance of nearly 40 Å, which is also approximately the length of the channels (Fig. 6).

The backbone resonances of the residues Met-69, Arg-70, Tyr-88, and Lys-97–Arg-100 of PC4_{CTD} could not be assigned because the corresponding signals were missing in all spectra. Interestingly, several other residues (Glu-93–Met-96) that are in close proximity to the ones that were unassigned give rise to multiple sets of signals, suggesting conformational heterogeneity. Because all of these residues are located in the region of the PC4_{CTD} structure that connects the tips of the β-ridge (which themselves produce well defined nondegenerate resonances) to the core of the protein dimer, it is tempting to speculate that the orientation of the tips with respect to the protein core is ill-defined in solution. Interestingly, the tips contain several negatively charged residues (Asp-91, Glu-93, and Glu-95) that seem to interfere with the interaction of PC4_{CTD}, with an 8-nucleotide heteroduplex in the hypothetical model published earlier (see Ref. 13 and Fig. 6B), because these charges would be in close proximity to the phosphate backbone of the DNA. However, such DNA molecules are bound with very high affinity (12), suggesting that steric hindrance does not occur. A possible explanation for this is that the tips may be able to adopt alternative conformations in which they have moved “inward” compared with the crystal structure (i.e. toward each other). Such alternative conformations may allow the β-ridge to be inserted into the heteroduplex opening more easily. Conformational heterogeneity would be consistent with the NMR degeneracy that is found for the amino acids in the hinge region and the fact that crystallographic temperature factors for this region are significantly higher than the average for PC4_{CTD} (13). So far only a single protein (other than the direct homologues of PC4 in various species) has been characterized that contains a region homologous to the PC4 C terminus. This protein, the yeast factor Sub1/Tsp1 (27, 28), is also an RNA polymerase II transcriptional coactivator but differs from PC4 in several important respects. In the first place, Sub1/Tsp1 is much larger than PC4 (33 kDa) and bears no homology to PC4 outside the region corresponding to the PC4 ssDNA-binding domain. Furthermore, the reported mechanisms for coactivation reported for Sub1/Tsp1 (28) and PC4 (4) differ considerably. Nevertheless, the homology between PC4_{CTD} and Sub1/Tsp1 (140–105) is very high (73% conserved, 49% identical), suggesting a similar fold and function. Indeed, Sub1/Tsp1 has been reported to bind to ssDNA (27). The fact that two otherwise quite different transcriptional cofactors both contain this type of ssDNA binding domain suggests that interaction with ssDNA or heteroduplexes is part of a general mechanism that helps regulate transcription in eukaryotes. Indeed, direct evidence for a role of ssDNA binding by PC4 in regulation of transcription was recently provided (6). The identification of DNA-interacting residues described here and the construction of point mutants that do not interact with ssDNA (6) will be helpful in the further elucidation of the precise role of the PC4 C-terminal domain and the homologous domain in Sub1/Tsp1, in transcription and other processes.

Acknowledgments—Marco Tessari is acknowledged for help with computer programs. We thank Petra Dux for critical reading of the manuscript and discussions.

REFERENCES

1. Ge, H., and Roeder, R. G. (1994) Cell 78, 513–523
2. Guernah, M., Malik, S., and Roeder, R. G. (1998) Mol. Cell. Biol. 18, 3234–3244
3. Kretschmar, M., Kaiser, K., Lottspeich, F., and Meisterernst, M. (1994) Cell 78, 525–534
4. Kaiser, K., Stelzer, G., and Meisterernst, M. (1995) EMBO J. 14, 3520–3527
5. Malik, S., Guernah, M., and Roeder, R. G. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 2392–2397
6. Werten, S., Stelzer, G., Goppelt, A., Langen, F. W. M., Gros, P., Timmers, H. Th. M., Van der Vliet, P. C., and Meisterernst, M. (1998) EMBO J. 17, 5103–5111
7. Wu, S. Y., and Chiang, C. M. (1998) J. Biol. Chem. 273, 12492–12498
8. Roeder, R. G. (1996) Trends Biol. Sci. 21, 327–335
9. Wang, Z., and Roeder, R. G. (1996) Mol. Cell 1, 749–757
10. Pan, Z-Q., Ge, H., Amin, A. A., and Hurwitz, J. (1996) J. Biol. Chem. 271, 22111–22116
11. Ballard, D. W., Philbrick, W. M., and Bothwell, A. L. M. (1988) J. Biol. Chem. 263, 8450–8457
12. Werten, S., Langen, F. W. M., Van Schaik, R., Timmers, H. Th., Meisterernst, M., and Van der Vliet, P. C. (1998) J. Mol. Biol. 276, 367–377
13. Brandsen, J., Werten, S., Van der Vliet, P. C., Meisterernst, M., Koon, J., and Gros, P. (1997) Nat. Struct. Biol. 4, 900–903
14. Bochkarev, A., Plüetzner, R. A., Edwards, A. M., and Frapprier, L. (1997) Nature 385, 176–181
15. Dayie, K. T., and Wagner, G. (1994) J. Magn. Reson. A 111, 121–126
16. Farron, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Biochemistry 33, 5984–6003
17. Vis, H., Vergne, C. E., Wotton, K. S., Kaptein, R., and Boelens, R. (1998) J. Biomol. NMR 11, 265–277
18. Cavanagh, J., Palmer, A. G., III, Wright, P. E., and Rance, M. (1991) J. Magn. Reson. 91, 429–436
19. Davis, A. L., Keeler, J., Laue, D. E., and Moskau, D. (1992) J. Magn. Reson. 98, 207–216
20. Kay, L. E., Keifer, P., and Saarinen, T. (1992) J. Am. Chem. Soc. 114, 10663–10665
21. Peng, J. W., and Wagner, G. (1994) Methods Enzymol. 239, 563–596
22. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) J. Biomol. NMR 6, 277–293
23. Vis, H., Mariani, M., Vergne, C. E., Wilson, K. S., Kaptein, R., and Boelens, R. (1995) Q. Magn. Res. Biol. Med. 2, 107–117
24. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) Protein NMR Spectroscopy. Principles and Practice, 1st edition, Academic Press, San Diego
25. Suck, D. (1995) Nat. Struct. Biol. 1, 161–165
26. Morin, A. G. (1995) EMBO J. 12, 861–867
27. Henry, N. L., Bushnell, D. A., and Kornberg, R. D. (1996) J. Biol. Chem. 271, 21842–21847
28. Knaus, K., Pollock, R., and Guarente, L. (1996) EMBO J. 15, 1933–1940